

Asset Management by Derivation

From the asset manager's questions to probabilistic inference and decisions

Draft · 465 pages

Ricardo Marquez, Ph.D.

ricardo-marquez.com

Asset Management by Derivation

Volume I

From the asset manager's questions to probabilistic inference and decisions

Ricardo Marquez

November 1, 2026 (work in progress: chapters are added as they are written)

How to read this book

This book is written for the asset manager, the maintenance or reliability planner and the condition-monitoring engineer. Its questions are the ones an asset's life raises: what condition it is in today, how fast it is changing, when it will need attention, what to do about it and when, and what the monitoring itself is worth. A trend chart answers them from raw performance: the approach temperature, the return temperature, the efficiency, plotted month by month with a line through it. Raw performance mixes aging with load and weather, so the chart reads a summer as degradation and a light load as a recovery. It cannot say when the coil will no longer hold the hall, whether it will last to the planned shutdown, or whether restoring it is worth more than replacing it. Each of these is a probability or an expected consequence over the unknowns of one model of the plant, and the book derives and computes each one.

The health index in this book is a parameter of the physics: a coil's conductance, a battery's capacity and internal resistance, read against the value the commissioning test found. Load and weather enter the model as known inputs, so the parameter's trend is aging and nothing else. Time enters as rows that tie one month's parameter to the next, so a trend is more rows in the same fit rather than a second tool, and every date and every decision is read from the posterior that fit produces.

The plant grows with the book. Part I works on one equipment hall, a state of two and then three entries, so that every object of the method can be seen whole. From Part II the hall has two cooling coils and three racks, and it is followed from its commissioning test through three years of monthly records. Two smaller assets join it: the site's UPS battery string and a fleet of fifty coils of the same design.

Every chapter that answers a question follows the same progression, and the progression is the method.

Step 0: the plant graph. The asset and the plant around it, before any question is asked: its nodes, paths and sources; which quantities are pinned and which are free; the physical relations that must hold among them. The free quantities form the state vector $\mathbf{z}$.

Step 1: the question. The asset manager's question in their own words, then as a probability or an expectation of a function of the state, with the unknown on the left of the conditioning bar and everything treated as known on the right.

Step 2: the graph extended for the question. The windows of the record and the trend rows that tie them; candidate explanations when something may have changed; the actions when a decision is asked; and the records attached as rows.

Step 3: the posterior of the state. Derived and computed once, before any question is answered: the MAP state, the information, the covariance, and the exact posterior where one exists.

Step 4: the answers. Each question read from that posterior down to its number: an entry of the state or the gradient of a derived quantity; a trend and the date it crosses a limit; a ratio of evidences when explanations compete; or an expected consequence when actions do.

Step 5: evaluation and checks. The gaps between the reported values and the exact values or draws collected, the approximations tested and, because the record is constructed from a stated truth, the answers tested against that truth.

Step 6: the report. The date, the probability, the band or the action the asset manager acts on, with the exact value beside it where one exists.

Step 7: the next measurement. If the band is too wide, the measurements that could be added ranked by the width each would remove from the answer that matters.

Five kinds of question cover what the book asks, distinguished by the mathematical object that answers them rather than by the word the maintenance office uses: *estimation*, *prediction under a threshold*, *hypothesis comparison*, *diagnosis* and *design*. Beneath them sit five computations: the *posterior* of the state, the *trend* of a parameter across windows, the *evidence* of a drawing of the plant, the *preposterior* after a measurement not yet taken, and the *expected consequence* of an action. Asset management is mostly trend and expected consequence. The notation chapter that follows fixes the symbols, and the appendix collects the asset manager's questions in an atlas that maps each to its mathematical form, its computation and the section that answers it.

The framework the book teaches is physics-constrained probabilistic modeling: a plant written as a graph, its unknowns as a state vector, every reading as a row, and every question as a probability over that state. TERRA is the author's own implementation of this framework; the book names the framework and the computations, not the software, and every number in it is reproducible from the derivations as written. The method and its notation are those of the companion books *Reliability Engineering by Derivation* and *Operations and Control by Derivation*, which ask the reliability engineer's and the operator's questions of the same hall.

Each chapter closes with a statement of what to retain, a table of its terms, and exercises at three levels, formulation, derivation and application, each followed by its answer.

Contents

How to read this book **ii**

Notation and vocabulary **viii**

I Thinking in Probabilities **1**

1 Physical assets as graphs **2**
1.1 One asset and several questions . . . 2
1.2 The posterior of the state (step 3) . . . 15
1.3 The answers (step 4) . . . 21
1.4 The checks (step 5) . . . 27
1.5 The report (step 6) . . . 29
1.6 The next measurement (step 7) . . . 31
1.7 What to retain . . . 35
1.8 Exercises . . . 38

2 The language of asset-management questions **40**
2.1 Asset-management questions and their mathematical form . . . 40
2.2 The grammar of probability statements . . . 43
2.3 The general forms and the five computations . . . 45
2.4 Estimation, prediction and joint events . . . 59
2.5 Model comparison, diagnosis and decisions . . . 70
2.6 Measurements, classification and workflow . . . 74
2.7 What to retain . . . 80
2.8 Exercises . . . 83

3 Uncertainty and conditional probability **86**
3.1 Uncertainty and evidence in asset-management questions . . . 86
3.2 Conditional probability and Bayes' theorem . . . 92
3.3 Bayesian diagnosis of fouling from a condition test . . . 94
3.4 Marginalization and dependence . . . 97
3.5 Conditioning in a physical model . . . 101
3.6 Predictive variance and the posterior covariance . . . 110
3.7 What to retain . . . 117
3.8 Exercises . . . 120

4 The anatomy of probabilistic inference **122**
4.1 From a plant graph to an acceptance probability . . . 122
4.2 The posterior of the state (step 3) . . . 131
4.3 The answers (step 4) . . . 144

4.4 The checks (step 5) . . . 149
4.5 The report (step 6) . . . 155
4.6 The next measurement (step 7) . . . 156
4.7 What to retain . . . 167
4.8 Exercises . . . 170

II Assessing Present Condition 173

5 Health indices from physical parameters 174
5.1 The dashboard's health index and the coil's . . . 174
5.2 The posterior of the state (step 3) . . . 180
5.3 The answers (step 4) . . . 185
5.4 The checks (step 5) . . . 192
5.5 The report (step 6) . . . 195
5.6 The next measurement (step 7) . . . 196
5.7 What to retain . . . 199
5.8 Exercises . . . 200

6 Hidden degradation under control 202
6.1 The held set point and the effort behind it . . . 202
6.2 The posterior of the state (step 3) . . . 208
6.3 The answers (step 4) . . . 210
6.4 The checks (step 5) . . . 216
6.5 The report (step 6) . . . 219
6.6 The next measurement (step 7) . . . 220
6.7 What to retain . . . 223
6.8 Exercises . . . 224

7 Degradation or sensor drift 226
7.1 A rise in the approach with two explanations . . . 226
7.2 The posterior of the state (step 3) . . . 233
7.3 The answers (step 4) . . . 237
7.4 The checks (step 5) . . . 246
7.5 The report (step 6) . . . 250
7.6 The next measurement (step 7) . . . 251
7.7 What to retain . . . 255
7.8 Exercises . . . 256

8 Battery state of health 258
8.1 The state-of-health tile and the string's circuit . . . 258
8.2 The posterior of the state (step 3) . . . 265
8.3 The answers (step 4) . . . 270
8.4 The checks (step 5) . . . 276
8.5 The report (step 6) . . . 280
8.6 The next measurement (step 7) . . . 281
8.7 What to retain . . . 285
8.8 Exercises . . . 286

9 Fleet pooling and hierarchical models 288
9.1 The fleet chart and the coil's place in it . . . 288
9.2 The posterior of the state (step 3) . . . 294
9.3 The answers (step 4) . . . 301
9.4 The checks (step 5) . . . 306
9.5 The report (step 6) . . . 309
9.6 The next measurement (step 7) . . . 310
9.7 What to retain . . . 314
9.8 Exercises . . . 315

III Estimating Degradation Rates 316

10 Degradation rates from condition records 317
10.1 The dashboard's trend line and the coil's rate . . . 317
10.2 The posterior of the state (step 3) . . . 323
10.3 The answers (step 4) . . . 327
10.4 The checks (step 5) . . . 336
10.5 The report (step 6) . . . 342
10.6 The next measurement (step 7) . . . 344
10.7 What to retain . . . 347
10.8 Exercises . . . 348

11 Accelerating degradation 350
11.1 A bend in the trend chart . . . 350
11.2 The posterior of the state (step 3) . . . 356
11.3 The answers (step 4) . . . 359
11.4 The checks (step 5) . . . 367
11.5 The report (step 6) . . . 373
11.6 The next measurement (step 7) . . . 374
11.7 What to retain . . . 378
11.8 Exercises . . . 379

12 Change-point detection 381
12.1 A step on the dashboard and a pump change . . . 381
12.2 The posterior of the state (step 3) . . . 386
12.3 The answers (step 4) . . . 389
12.4 The checks (step 5) . . . 397
12.5 The report (step 6) . . . 403
12.6 The next measurement (step 7) . . . 404
12.7 What to retain . . . 408
12.8 Exercises . . . 409

13 Seasonal effects and aging 411
13.1 The summer drop on the dashboard . . . 411
13.2 The posterior of the state (step 3) . . . 418
13.3 The answers (step 4) . . . 419
13.4 The checks (step 5) . . . 431
13.5 The report (step 6) . . . 434
13.6 The next measurement (step 7) . . . 436
13.7 What to retain . . . 439
13.8 Exercises . . . 440

A Atlas of asset-management questions **442**

Notation and vocabulary

Every symbol below is used with the same meaning in every chapter. The unknown always stands on the left of a conditioning bar and the evidence on the right. Vectors and matrices are bold. Coil subscripts are $u \in \{1, 2\}$, rack subscripts $j \in \{1, 2, 3\}$ and fleet subscripts $i = 0, \ldots, 49$; windows are months, indexed t.

Symbol	Unit	Meaning
The plant graph and its state		
$\mathbf{z}$, z_i	varies	State vector of the plant graph: every free quantity estimated at once; its i-th entry.
$\mathbf{z}_t$, t	varies, –	The state of window t; the window index, $t = 0$ the commissioning test and $t = 1, \ldots, 36$ the months from January of year 1 to December of year 3.
$\mathbf{y}$, y	varies	The evidence: all readings, test results, counts and prior information treated as known; a single reading, subscripted by what it reads.
$r_k(\mathbf{z})$	varies	Row k of the model: a physical relation that must hold ($r_k = 0$), a reading compared with the state, a trend row, or a prior. As a subscript k is this row (and in Chapter 24 a channel's rows); in Chapters 2, 3 and 15, whose bare k is the count of failures, the row is ℓ.
w_k, σ_k	varies	Weight (precision) of row k and the standard deviation it corresponds to, $w_k = 1/\sigma_k^2$.
σ_p	varies	Tolerance of a physics row; an exact constraint is the limit $\sigma_p \to 0$.
The equipment hall of Part I		
T, T_c	°C	Temperature of the hall air; the supply air held by the cooling unit's discharge controller, pinned.
Q, G	kW, kW/K	Heat load of the hall; effective conductance of its air path to the supply air.
g, G_{ref}	–, kW/K	Log-conductance $\ln(G/G_{\text{ref}})$, the entry of the state that carries the conductance; the design conductance on the design sheet, 6 kW/K.
h, h^*, h_F	–	Part I's health index $h = G/G_{\text{ref}} = \mathrm{e}^g$, the conductance against design; the index a requirement needs; the functional limit, at which the asset stops doing its job. From Part II the health index is against commissioning, $h_{u,t}$ below.
μ_g, σ_g	–	Mean and spread of the handbook prior on g, $\mu_g = \ln(7.5/6)$.

Symbol	Unit	Meaning
g_t, $\hat{g}_0$	–	The log-conductance in window t; the commissioned value, the fit of the commissioning test (window 0).
$\sigma_{\hat{g}_0}$, $\hat{G}_0$	–, kW/K	Spread of the commissioning row on g (the commissioning fit carried forward); the commissioned conductance $G_{\text{ref}}\text{e}^{\hat{g}_0}$.
G_{test}, h_{test}	kW/K, –	The conductance and index one condition test reports when its reading is taken as exact, $Q/(y_T - T_c)$.
g^*, g_F	–	A requirement level on g, $\ln h^*$; from Chapter 19, the level on each coil's log-derate at which two equal coils give G^*, $\ln(G^*/(\bar{G}_1 + \bar{G}_2)) = -0.26313$; the functional limit, $\ln h_F$, in Chapter 20 $\ln(5/6) = -0.18232$ on each fleet coil's log-derate against rating.
$T_{\max}$, T^*	°C	A temperature limit: 24°C now and 27.5°C at 60 kW (the owner's request, Chapter 1), 28°C (the commissioning record's acceptance limit, Chapter 4); a threshold on a test temperature.
Q_2, T_2	kW, °C	The requested load and the hall temperature it produces, a scenario node.
ρ, μ_0, σ_0	K/kW	Thermal resistance of the air path, $1/G$; the commissioning record's pre-test estimate of it and its spread (Chapter 4).
b_T	K	Calibration offset of the hall sensor, a nuisance quantity or a bias entry.
Δ_{m}, σ_{m}	kW	A model discrepancy of the physics row and its spread (Chapter 1).
C_{th}, ϑ, ϑ_{th}, ΔT	kJ/K, s or h, s, K	Thermal capacitance of the hall; elapsed time within a record (seconds in Chapter 1's transient, hours into a battery test in Chapter 8, hours within a window in Chapter 24); the hall's time constant C_{th}/G; the departure from steady state.
y_T, σ_T	°C, K	The hall sensor's reading and the standard deviation of its error.
The two-unit hall of Parts II–VI		
$T_s(t)$, $Q_o(t)$	°C, kW	Chilled-water supply temperature and the hall's other heat in window t: known inputs that follow the season.
Q	kW	IT heat load of the hall, an entry of the state read by the hall load meter.
G_u, $\bar{G}_u$	kW/K	Effective conductance of coil u; its rated value, 7 kW/K.
g_u, d_u	–	Log-derate $\ln(G_u/\bar{G}_u)$, the entry that carries coil u's conductance against its rating; the derate factor $d_u = \text{e}^{g_u}$, one at the rating.
$g_{u,t}$, $d_{u,t}$, Q_t	–, –, kW	The log-derate, derate factor and IT load in window t.
$h_{u,t}$	–	Health index of coil u in window t against its commissioned value, $G_{u,t}/G_{u,0} = \text{e}^{g_{u,t}-g_{u,0}}$ (Part II on).
$A_{u,t}$, h_u^{raw}	K, –	Approach of unit u in window t, $T_{d,u} - T_s(t)$; the dashboard's raw-performance index $A_{u,0}/A_{u,t}$, the naive baseline.
$G_{u,\text{test}}$	kW/K	Conductance of coil u with the window's readings taken as exact.

Symbol	Unit	Meaning
σ_W	kW	Standard deviation of a water-side heat meter's reading on a coil's duty.
$g_{u,0}$	–	The commissioned value of g_u, the fit of window 0.
C_u	kW/K	Air capacity rate of unit u, pinned at its fixed fan speed; read by an airflow station only in Chapter 23's candidates.
T_r, T_c, $T_{d,u}$	°C	Return air, cold aisle (the supply air of Part I's hall, now made by two coils), and discharge air of unit u.
T_j, β_j	°C, –	Inlet temperature of rack j; its recirculation fraction.
G^*, f^*	varies	A requirement level: G^* the conductance a duty needs (in Part I, 5.0 kW/K for 24°C at 30 kW and 6.316 kW/K for 27.5°C at 60 kW; from Part II, the total $G_1 + G_2$ of the capacity scenario); f^* a level on any derived quantity. The requirement holds on the safe side of f^*: $f \geq f^*$ for a conductance, $f \leq f^*$ for a temperature.
b_j	varies	Bias of reading j, an entry freed when a sensor is suspected of drifting.
$\check{X}$	as X	The value of X at the scenario node of the commissioned model: today's load and inputs with the commissioned coils ($\check{T}_r$, $\check{D}_u$, $\check{A}_u$; Chapter 6).
$D_{u,t}$	kW	Duty of unit u in window t, the heat its coil passes to the chilled water, $G_{u,t}(T_{r,t} - T_s(t))$.
$\eta_{A,u}$, $\eta_{D,u}$, η_s	K, kW, K	Excess approach, duty shift and supply penalty against the commissioned model at today's conditions (Chapter 6).
$A_{r,t}$, T_c^{sp}	K, °C	Return approach $T_{r,t} - T_s(t)$; a cold-aisle set point.
$G_{\mathrm{tot},t}$, G_t^{app}, $\bar{Q}_o$	kW/K, kW/K, kW	The hall's total conductance $G_{1,t} + G_{2,t}$; its apparent conductance, the metered IT load over the return approach, $y_{Q,t}/(y_{T_r,t} - T_s(t))$, a raw figure; the mean other heat, 8.00 kW (Chapter 13).
ζ_0, s_{season}, s_{load}	K/kW, K, K	The cold aisle's rise per kW at the commissioned coils; the season's and the load's terms of the cold aisle's change since commissioning; from Chapter 13, the season's and the load's terms of a raw figure's change between any two windows.
$b_{d,1}$, σ_b, $\sigma_{\ln h}$	K, K, –	Drift of unit 1's discharge sensor since commissioning (a bias entry); the prior spreads of a freed bias and of a freed change $\ln h_{u,t}$ (Chapter 7). From Chapter 10 the drift is a function of the window, $b_{d,1}(t)$.
$\dot{b}_{d,1}$, t_b, $\sigma_{\dot{b}}$	K/yr, window, K/yr	Drift rate of unit 1's discharge sensor, one entry for the record, with $b_{d,1}(t) = \dot{b}_{d,1}\,(t - t_b + 1)\,\Delta t$ for $t \geq t_b$ and zero before; the drift's first window, $t_b = 20$; the drift rate's prior spread, 2.0 (Chapter 10).
M_{neither}, M_{foul}, M_{drift}, M_{both}	–	The four drawings of Chapter 7: neither term, the derate term, the bias term, both.
K_{kk}, r_k^{N}, n_r	–	Leverage of row k, the diagonal of the hat matrix $\mathbf{K} = \mathbf{J}\boldsymbol{\Sigma}\mathbf{J}^{\mathsf{T}}$, with redundancy number $1 - K_{kk}$; the normalized residual $r_k/\sqrt{1 - K_{kk}}$; the number of rows.

Symbol	Unit	Meaning
ψ_k, $s_{d,1}$, y_b	–, K, K	The ratio in the closed-form Bayes factor of a bias on reading k (Chapter 12's ψ extends it to a step's column); the shift of unit 1's discharge reading the two suspects make; a side-by-side check's reading of a bias.
Time, trend and dates		
Δt	yr	The spacing of two windows, one month.
γ, γ_u, $\dot{\gamma}_u$	yr^{-1}, yr^{-1}, yr^{-2}	Trend (degradation rate) of a parameter; coil u's rate, the trend entry of its log-derate path, read as $100(\mathrm{e}^{\gamma_u} - 1)\,\%/\mathrm{yr}$ (Chapter 10), and in a drawing that frees the curvature the rate at the record's center $\bar{t}_\mathrm{y}$; its curvature (acceleration), the change of the rate per year (Chapter 11). Chapter 9's γ_i indexes the fleet's coils instead: its γ_1 is fleet coil 1, not the hall's coil 1.
t_y, $\bar{t}_\mathrm{y}$, $\gamma_u(t_\mathrm{y})$	yr, yr, yr^{-1}	Time since commissioning, $t_\mathrm{y} = t\,\Delta t$ (Chapter 11); the record's center, 1.5 yr (window 18); coil u's rate at time t_y, $\gamma_u + \dot{\gamma}_u(t_\mathrm{y} - \bar{t}_\mathrm{y})$, so that $\gamma_u(3)$ is the rate today.
x_t, $\dot{g}_{u,t}$, $\sigma_{\dot{\gamma}}$	yr, yr^{-1}, yr^{-2}	The curvature's lever on the increment from window t to $t+1$, $(t+\frac{1}{2})\Delta t - \bar{t}_\mathrm{y}$; the path's monthly rate $(g_{u,t+1} - g_{u,t})/\Delta t$; the curvature's prior spread, 0.05 (Chapter 11).
t_knee	window	The knee of a rate, after which it changes by $\dot{\gamma}_u$ per year: 12 or 24 in Chapter 11's sensitivity forms.
Γ	yr^{-1}	The hall's conductance rate, $\ln(G_{\mathrm{tot},36}/G_{\mathrm{tot},0})/(36\,\Delta t)$ (Chapter 13). With an argument, $\Gamma(\cdot)$ is the gamma function (Chapter 2).
α	–	Level entry of a straight-line trend at the record's center $\bar{t}$.
σ_q	–	Spread of a trend row: one monthly window's departure of a parameter from its straight line, $\sigma_q\sqrt{n_\mathrm{w}}$ across n_w windows without a reading; the straight line is the limit $\sigma_q \to 0$.
$r_{q,u,t}$, $\mathbf{D}_t$, $\mathbf{E}$, $\mathbf{F}$	–, varies	Coil u's trend row from window t to $t+1$; the trend rows' blocks of the stacked Jacobian in window t's columns and in the trend entries (γ_1, γ_2) (Chapter 10), and in the curvatures (Chapter 11).
n_w, n_z, t_as	–, –, window	The number of monthly increments a rate is read over, 36 for the whole record, or of unread windows a trend row spans (Chapters 4 and 10); the number of entries of the state (Chapters 3, 4, 7 and 12); the last window of an as-of fit, the record cut at window t_as.
s_u^raw, $\dot{s}_u^\mathrm{raw}$	yr^{-1}, yr^{-2}	Slope of the dashboard's least-squares line through the 37 tiles $h_{u,t}^\mathrm{raw}$ (Chapter 10); the dashboard's bend, twice the quadratic coefficient of the least-squares quadratic through them (Chapter 11).
$\tilde{\mathbf{z}}^0$, $\tilde{\mathbf{z}}^g$	varies	The stacked state with every window's coils at the commissioned values and no drift; with the fitted coils and no drift: the steps of Chapter 10's sequential split of the dashboard's slope.

Symbol	Unit	Meaning
γ_u°	yr^{-1}	Pseudo-true value of γ_u: the drawing's fit to the truth's noise-free readings (step 5, Chapter 10).
M_0, M_1, M_2, M_{12}	–	The constant-rate drawing (Chapter 10's base fit) and the drawings that free coil 1's, coil 2's and both curvatures (Chapter 11). Chapter 12 writes the first M_{none}.
δ, δ_j, t_0	–; –, K or kW; window	Size of a step change in a parameter (with an argument, $\delta(\cdot)$ is the Dirac delta, Chapter 3); a step of kind j (Chapter 12): δ_1, δ_2 a coil's, δ_{12} both coils', δ_{T_r}, $\delta_{T_{d,1}}$, $\delta_{T_{d,2}}$, δ_{T_1} a temperature sensor's, δ_Q the load meter's; the window at which it occurs.
s, σ_δ, Π_s	window, as δ_j, –	A candidate window of a step, 1 to 36; the step's prior spread, 0.15 on a coil, 1.0 K on a temperature sensor, 10 kW on the meter; the location prior, the prior probability of a step's window s (Chapter 12).
M_{none}, $M_{j,s}$, $\mathbf{a}_{j,s}$	–, –, varies	The drawing with no step (Chapter 10's base fit); the drawing with one step of kind j into window s; the step's column in the normalized rows, $-1/\sigma_q$ on a coil's trend row into s and $1/\sigma_k$ on a sensor's readings from s on (Chapter 12).
$B_{j,s}$, $\bar{B}_s$, $\bar{B}$, $B_{\max}$	–	Bayes factor of $M_{j,s}$ against M_{none}; its mean over the eight kinds at window s; its mean over all 288 candidates; the best candidate's (Chapter 12).
ψ, ψ_a	–	A step's prior variance over its variance from the other rows, $\sigma_\delta^2\,\mathbf{a}^\mathsf{T}(\mathbf{I}-\mathbf{K})\mathbf{a}$; the same after a candidate measurement a, $\sigma_\delta^2/\sigma_{\text{dat},a}^2$ (Chapter 12).
D_h, D_γ, $\mathbf{g}_{D_\gamma}$, $\mathbf{g}_{\text{split}}$	–, yr^{-1}, –, mixed	The difference $\ln h_{2,18} - \ln h_{1,18}$ of the two coils' indices (Chapter 5); the difference $\gamma_1 - \gamma_2$ of the two rates; its gradient; the gradient of coil 2's two parts of loss, $36\,\Delta t\,\gamma_2 + \delta_2$ (Chapter 12).
σ_{dat}^2, $\sigma_{\text{dat},a}^2$, σ_{new}^2	–	A step's data-only variance from the present record, with the prior taken out; the same after a candidate measurement a; that of the candidate's rows alone (Chapter 12, step 7).
$\tilde{\delta}$, $\tilde{\delta}_0$, $\tilde{\delta}_{\text{new}}$, $\tilde{\mu}$, $\tilde{v}$	–	The mean of a step's likelihood, the estimate the rows alone give; the present record's; the candidate rows'; the combined estimate's mean and variance under the posterior predictive (Chapter 12, step 7).
t_1, t_2	window	The two windows of a raw figure's change; for the summer drop, July of year 3 (window 31) and today (window 36) (Chapter 13).
ϕ_t, mo_t	rad, –	The season's phase of window t, $2\pi(\mathrm{mo}_t - 0.5)/12$, with $\cos\phi_t = 1$ between January and February; the month index of window t, 0 for January. Window 0 takes zero seasonal regressors (Chapter 13).

Symbol	Unit	Meaning
Y_t, y_t, $\check{Y}_t$, $\check{Y}_t^{\mathrm{m}}$	as the figure	The model value of a raw figure ($T_{r,t}$ or $\ln G_t^{\mathrm{app}}$) at window t; its reading; its value with the coils at the commissioned ones, at the window's own load and inputs; that value at the season's mean inputs, window 0's (Chapter 13).
s_{coils}, s_{err}	K or –	The coils' term and the reading errors' term of a raw figure's change from window t_1 to t_2, beside s_{season} and s_{load}; in bold, each term's column over the windows (Chapter 13).
$\mathbf{X}$, $\mathbf{X}^+$, $\hat{\mathbf{k}}$, k_c, k_s	varies	The design of a least-squares regression on the windows' times (Chapters 2, 11 and 13); its pseudo-inverse $(\mathbf{X}^{\mathsf{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathsf{T}}$; the fitted coefficients; a seasonal regression's cosine and sine coefficients (Chapter 13).
$\mathbf{y}^{\mathrm{raw}}$, $\mathbf{w}$, $\mathbf{w}^{\mathrm{line}}$, $\mathbf{w}^{\mathrm{seas}}$	as the figure, yr^{-1}	The readings of a raw figure in windows 0 to t_{as}; the slope row of a regression, the second row of $\mathbf{X}^+$, without and with an annual harmonic (Chapter 13).
s^{reg}, s^{line}, s^{seas}, $s^{\mathrm{seas}}_{\mathrm{drop}}$	varies	The slope of a regression; of the line and of the seasonal regression through a raw figure, per year; the seasonal regression's reading of the summer drop, its fitted trend over the five months plus the residuals' change (Chapter 13).
$\hat{\nu}_t$, $\hat{\sigma}$, $\mathrm{se}(\cdot)$	as the figure	A regression's residual at window t; the residuals' standard deviation; the practitioner's standard error from the residual scatter (Chapter 13).
$\mathbf{\Sigma}_{31,36}$, $\mathbf{g}_{31,36}$	–	The covariance block of $(g_{1,31}, g_{2,31}, g_{1,36}, g_{2,36})$; the gradient of the coils' term of the summer drop on those entries (Chapter 13).
g_u^c, g_u^s, $\mathbf{g}^{\mathrm{sea}}$, σ_{sea}	–	Cosine and sine amplitudes of coil u's seasonal conductance, on top of its trend; the four amplitudes as one vector; their prior spread, 0.05 (Chapter 13).
M_{base}, M_{sea}	–	The base drawing (Chapter 10's base fit) and the seasonal-conductance drawing (Chapter 13).
t^*, t_c	window	The window at which a parameter's path crosses its requirement level: the remaining useful life is t^* less the present window. From Chapter 14 it is the first passage of $G_{\mathrm{tot},t}$ to G^* on the extended state, read in continuous windows by interpolation between $t_c - 1$ and t_c, with $t_c = \lceil t^* \rceil$ the first window at or below the limit, the date; from Chapter 17, $t_c(\mathbf{z}; \mathcal{P}_{\mathrm{v}})$ is counted from a visit.
t_{sd}	window	A planning date: the October shutdown of year 4 (window 46).
Failures and the fleet		
λ, τ	yr^{-1}, yr	A constant failure rate; the length of a mission window (Chapters 2, 3, 15 and 16).
k, T_{tot}	–, unit-yr	A count of failures or corrective work orders; the exposure behind it.
ω, λ_{ref}	–, yr^{-1}	Log-rate $\ln(\lambda/\lambda_{\mathrm{ref}})$, the entry that carries a rate; its reference, 0.1 per year.

Symbol	Unit	Meaning
σ_ω, T_f, K_τ	–, yr, –	Posterior spread of the log-rate; time to a unit's next corrective failure; count of corrective work orders in a period τ.
F, H, S	–	Events of a diagnosis: the coil fouled; the condition flag fires; an inspection grade, of value s in $\Pr(S = s)$ (Chapter 3).
F_A, F_B, q_S, q_A, q_B	–	Two findings or replacements; the probabilities of a shared cause and of each one's own wear.
a, b, a_λ, b_λ	–, yr	Shape and rate of a gamma prior on a rate (primed after the record); the context tells them from an action a. From Chapter 15, where a is the age, they are written a_λ, b_λ: the prior is a count row of a_λ failures in b_λ unit-years.
$\lambda(a)$, λ_0, θ_a	yr^{-1}, yr^{-1}, yr^{-1}	Failure rate at age a, $\lambda_0 \mathrm{e}^{\theta_a a}$, the rate of occurrence of failures under minimal repair (Chapter 16); its value at age zero; its log-slope in age, a scalar.
n_a, E_a	–, coil-yr	Failures and exposure of the fleet in age bin a, the ages a to $a+1$ (Chapter 16).
μ_γ, τ_γ	yr^{-1}	Mean and spread of the fleet's degradation rates, the shared prior of partial pooling. In Chapter 9 $\tau_\bullet$ stands for either of the fleet's spreads, τ_γ or τ_c, and $n_\mathrm{coil} = 50$ counts the coils.
μ_c, τ_c	–	Mean and spread of the fleet's commissioning values.
$g_{i,0}$, γ_i, y_{ij}, a_{ij}, σ_test, k_i	–, yr^{-1}, –, yr, –, –	Commissioning value and fouling rate of fleet coil i; its j-th annual condition test and the age at it; the test's spread (0.04); its number of tests.
ℓ_c, ℓ_γ, $\varepsilon_{\mathrm{c},i}$, $\varepsilon_{\gamma,i}$	–	$\ln \tau_\mathrm{c}$, $\ln \tau_\gamma$, the entries that carry the spreads; the standardized deviations of the non-centered parametrization.
$\mathbf{\Omega}$, $\mathbf{\Lambda}_i$, $\mathbf{\Lambda}_\mu$, $\mathbf{\Lambda}_\mathrm{pr}$, $\mathbf{\Sigma}_i$	varies	Population covariance $\mathrm{diag}(\tau_\mathrm{c}^2, \tau_\gamma^2)$; coil i's information given the spreads and means; the information on the means; their prior information; coil i's covariance with the means integrated (Chapter 9).
Ψ, ϖ_m	–	The Gaussian rows' sum of squares at the conditional minimum; the weight of node m of a grid on the state: the spreads' grid (Chapters 9 and 20), the log-rate's (Chapter 15), the age slope's (Chapter 16).
π_i, $\mathcal{K}_i$, $\tilde{\gamma}_0$	–	Fleet position: coil i's percentile in the design's population of rates and its rank among the fifty (1 the fastest); coil 0's standardized rate.
γ_new, $\hat{\gamma}_0^\mathrm{chart}$	yr^{-1}	The rate of a new coil of the design (the posterior predictive); the fleet chart's slope of coil 0.
The UPS battery string		
$C_{\mathrm{bat},t}$, $R_{\mathrm{bat},t}$	Ah, mΩ	Usable capacity and ohmic resistance of the string in test t; $\bar{C}_\mathrm{bat}$, $\bar{R}_\mathrm{bat}$ their ratings, 100 Ah and 40 mΩ.
ξ_t, h_C, ΔR	–, –, mΩ	Log-capacity $\ln(C_{\mathrm{bat},t}/\bar{C}_\mathrm{bat})$; capacity against commissioning; resistance rise since commissioning.

Symbol	Unit	Meaning
v_0, Δv, h_{ch}, N	V, V, V, –	A block's loaded curve at the usable-empty point, its rise to full, and the charge curve's offset above it; the number of blocks, 40.
I_d, I_c, SOC	A, A, –	Discharge and recharge currents; state of charge.
$V_{d,t}$, $\dot{V}_{d,t}$, $V_{c,t}$, $\dot{V}_{c,t}$	V, V/h	Level and slope of test t's discharge line and of its recharge line.
$T_{\mathrm{run},t}$, V_{cut}, $\mathrm{SOC}_{\mathrm{cut},t}$, $V_{\mathrm{h},t}$	h, V, –, V	Run time to the cutoff; the cutoff on the string, 420 V; the state of charge there; the loaded head $N(v_0 + \Delta v) - V_{\mathrm{cut}} - I_d R_{\mathrm{bat},t}/1000$.
σ_V, $\sigma_{\mathrm{run},t}$, $\mathbf{s}$, $\mathbf{n}$	V, h, mixed	The voltmeter's spread; the run-time row's spread; the two discharge lines as coordinates; the ridge (null direction) of the discharge rows.
Probability		
$\Pr(A)$, $\Pr(A \mid B)$	–	Probability of event A; conditional on event B.
$p(\cdot)$	varies	A probability density; $p(\mathbf{z} \mid \mathbf{y})$ is the posterior of the state, $p(\mathbf{y} \mid \mathbf{z})$ the likelihood, $p(\mathbf{z})$ the prior.
$\mathbb{E}$, Var, Cov	varies	Expectation, variance, covariance.
$\mathcal{N}(\mu, \sigma^2)$	–	Normal distribution with mean μ and variance σ^2; Φ its standard cumulative distribution function, φ its density.
Inference objects		
$\mathcal{L}(\mathbf{z})$	–	The objective, minus the log posterior up to a constant.
$\hat{\mathbf{z}}$	varies	The maximum a posteriori (MAP) state, the minimizer of $\mathcal{L}$.
$\mathbf{J}$	varies	Jacobian of the normalized rows with respect to the state, evaluated at $\hat{\mathbf{z}}$.
$\mathbf{\Lambda}$	varies	Information matrix (Gauss–Newton Hessian of $\mathcal{L}$), $\mathbf{J}^{\mathsf{T}}\mathbf{J}$ for normalized rows.
$\mathbf{\Sigma}$, Σ_{ij}	varies	Posterior covariance of the state under the Laplace approximation, $\mathbf{\Lambda}^{-1}$; its entries. The posterior spread of entry i is written $\sqrt{\Sigma_{ii}}$.
$f(\mathbf{z})$, $\mathbf{g}$	varies	A derived quantity of the state and its gradient $\partial f/\partial \mathbf{z}$ at $\hat{\mathbf{z}}$.
$\mathbf{N}$, $\mathbf{J}_{\mathrm{red}}$	varies	Null-space basis of the physics rows; the Jacobian of the reduced state (Chapter 4).
f_L, f_U	varies	The ends of the credible interval of a derived quantity.
ϱ, $\boldsymbol{\varrho}$, Φ_2	–	A correlation coefficient (Chapters 2 and 3; in Chapter 24 that of two consecutive logged readings' errors); the posterior's correlation matrix (Chapter 7); the distribution function of two standard normals with correlation ϱ.
υ, $\mathbf{\Upsilon}$, S_{in}	–	Weight of a reading; the gain of the correction form; the innovation variance of one reading (Chapter 4).
x_j, w_j	–	Gauss–Hermite nodes and weights.
ϖ_j, N_{eff}, p_{L}	–	Weight of draw j in the reweighted draws; their effective sample size $1/\sum_j \varpi_j^2$; the Laplace normal used as the proposal (Chapter 5).

Symbol	Unit	Meaning
ℓ_t, ε_q	–	Magnitude of the slope of window t's reading row; the summed departure of unread windows from the line (Chapter 4).
y_{new}, $\tilde{y}_a$	varies	A future reading (the predictive) and the outcome of a candidate measurement a before it is taken (the preposterior).
U_{p}, $U(a)$	varies	A quantity of the posterior after $\tilde{y}_a$; its expectation over the outcomes, which ranks the candidates (Chapter 2's design form).
M, M_j, B_{AB}	–	A drawing (model) of the plant; the members of a family of candidates; the Bayes factor $p(\mathbf{y} \mid M_A)/p(\mathbf{y} \mid M_B)$.
M_{clean}, M_{fouled}, M_{bias}, B_{fc}, B_{fb}	–	Chapter 2's drawings of a warm condition test: the air path clean, the air path fouled, the hall sensor with a bias; the Bayes factors of fouled against clean and of fouled against bias.
M_{one}, M_{two}, $B_{\text{two,one}}$	–	Chapter 4's alternatives for the hall's air path, a single thermal resistance and two distinct heat-removal paths (beside the sensor bias M_{bias}); the Bayes factor between them.
$\mathbf{J}_a$, $\mathbf{\Lambda}_a$, $\mathbf{\Sigma}_a$	varies	The Jacobian rows of a candidate measurement, its information $\mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a$, and the preposterior covariance $(\mathbf{\Sigma}^{-1} + \mathbf{\Lambda}_a)^{-1}$.
Decisions		
a, $L(a, \mathbf{z})$	–, R$	An action (repair, replace, inspect, wait); its loss when the state is $\mathbf{z}$.
$\mathcal{R}(a)$, c_a, C_e	R$	The expected loss of action a, $\mathbb{E}[L(a, \mathbf{z}) \mid \mathbf{y}]$; its fixed cost; the consequence of event e.
c, c^*, c_1^*	varies, –, –	A decision level on a monitored parameter (the work-order trigger), from Chapter 19 a level on the probability $q_{t_{\text{as}}}$; the level of least expected loss with a refit every month, 0.0197; the one-look level $C_{\text{false}}/(C_{\text{false}} + C_{\text{miss}})$, 0.0980 (Chapter 19).
κ	–	Assurance level of an acceptance rule: the posterior probability a requirement must reach (0.95 in Chapter 22's verdict; Chapter 21's break-even p_{be} replaces it with a level the costs set).
EVSI, EVPI, EVPPI	R$	Expected value of sample information: what a measurement is worth before an action. Expected value of perfect information, $\mathcal{R}(a^*) - \mathbb{E}[\min_a L(a, \mathbf{z}) \mid \mathbf{y}]$: what knowing the whole state would be worth, the most any measurement can be (Chapter 17, its general form; Chapter 20 specializes it to the fleet's orders). Expected value of partial perfect information: the same with one open question known instead of the whole state (Chapter 17).

Vocabulary

Plant graph The physical system as nodes (places where a conserved quantity balances), edges (paths along which it flows) and sources, with the relations that must hold among them.

State vector The collection of free quantities the fit estimates at once.

Pinned, free, known input A quantity treated as known (a boundary condition, or an input such as the supply temperature read by a precise instrument) against one carried in the state.

Residual, row A relation written so that it equals zero when satisfied: physics rows constrain the state, reading rows compare it with the records, trend rows tie one window's parameter to the next, prior rows carry handbook or commissioning knowledge.

Window The span of records fitted as one steady condition: one representative hour of a month here.

Health index A physical parameter of the model read against its commissioned value; never a score built from raw performance.

Condition monitoring Following the health index through the windows of a record.

Degradation rate The trend of the health index per year.

Remaining useful life The time until the health index crosses the level at which the asset can no longer do its job, as a date with a band.

P–F interval The time from the first detectable change of the health index to its crossing of the functional limit.

Condition-based maintenance Maintenance triggered by the health index crossing a decision level, rather than by the calendar.

Partial pooling Estimating each unit of a fleet with a shared prior whose mean and spread are themselves estimated, so that the fleet informs the unit and the unit informs the fleet.

Change-point detection Locating a step in a parameter's path by the evidence of a step freed at each candidate window.

State of health The capacity and internal resistance of a battery against their values when new.

Laplace approximation Replacing the posterior by the Gaussian whose covariance is the inverse curvature of the objective at the MAP state.

Delta method Linearizing a derived quantity around the MAP state so that its variance is $\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}$.

Posterior predictive The distribution of a future reading obtained by averaging its model over the posterior.

Evidence, Bayes factor The normalizing constant of one model's posterior; the ratio of two models' evidences.

Preposterior The posterior that would result from a measurement not yet taken; the object that values a measurement before it is made.

The five computations Posterior, trend, evidence, preposterior and expected consequence: the objects every asset-management question is answered from.

Part I

Thinking in Probabilities

Chapter 1

Physical Assets as Graphs

1.1 One asset and several questions

The owner of a data center equipment hall asks its asset manager whether the hall can take twice its present load, from 30 to 60 kW, with the hall air then held below 27.5°C. The hall's air path, from the hall to the supply air that the cooling unit's discharge controller holds at 18°C, has a short history: a design sheet, which specifies its conductance as 6 kW/K, and a commissioning test, at which the one hall sensor, with an error of about half a kelvin, read 23.2°C at 30 kW. The health-index dashboard takes that reading as exact: at 23.2°C under 30 kW the path conducts 5.77 kW/K, an index of 0.962 against the design sheet, and the dashboard reports that the path was commissioned 3.8 % below design. At that conductance the doubled load puts the hall at 28.4°C, over the limit. Read from the dashboard alone, the path was handed over short of design and must be brought up to design before the load doubles.

The handbook tile beside it tells another story. The design handbook gives the conductance of a path of this kind as 7.5 kW/K, right to within a factor of about 1.35, an index of 1.25; heat leaves the hall in proportion to its excess over the supply air, so at that conductance the hall sits at 22°C under 30 kW and at 26°C under 60 kW, a comfortable pass. Neither tile carries a band, and neither is weighed against the other. One test is one window of the path's record, and it carries no degradation rate and no age (Chapter 4 adds a second window). What the decision needs is a probability that the path carries 60 kW within 27.5°C, and, before an upgrade is ordered, the probability that the path was commissioned below design at all. The question of this chapter is therefore

how a physical asset is represented so that one representation serves every question the asset manager asks of it (its condition, its capacity, the next measurement), and what must be fixed before any inference is performed.

Jargon: plant graph, node, path, source, pinned and free quantities, state vector, residual, row, observation, prior, posterior, design value, health index, commissioning test, commissioned value, model discrepancy, scenario node, Jacobian, normal equations, information matrix, covariance, Laplace approximation, factor graph

A *plant graph* represents a physical system as *nodes*, places where a conserved quantity such as energy or mass must balance; *paths* (edges), along which the quantity flows; and *sources*, which supply it. A quantity is *pinned* when it is treated as known, as a boundary condition, and *free* when it is unknown and must be inferred. The free quantities, collected in one vector, form the *state vector* $\mathbf{z}$. A *residual* is a relation written so that it equals zero when it is satisfied; each residual is a *row* of the model. A physical row states a relation that must hold among entries of the state; an *observation* (a reading) is a measured value related to an entry of the state with an error, and its row compares the two. The *prior* expresses what is known about the state before the readings, for example a handbook value of a conductance; the *posterior* $p(\mathbf{z} \mid \mathbf{y})$ expresses what is known after the evidence $\mathbf{y}$. The *design value* of a parameter is the value its design sheet specifies. The *health index* is a parameter of the physics read against a reference value, here the conductance of the air path as a fraction of its design value. A *commissioning test* (acceptance test) is the first test of the asset in service, and its fit, the *commissioned value*, is the baseline against which later condition tests are read. A *model discrepancy* is the difference between the physical system and the model of it that would remain if every entry of the state were known exactly. A *scenario node* is a node added to the graph for an operating condition that has not occurred, so that a question can be asked about it. The *Jacobian* $\mathbf{J}$ is the matrix of derivatives of every row with respect to every entry of the state, one row per row of the model and one column per entry. The *normal equations* $(\mathbf{J}^\mathsf{T}\mathbf{J})\,\Delta\mathbf{z} = -\mathbf{J}^\mathsf{T}\mathbf{r}$ give the step that best reduces all the rows at once when each is replaced by its tangent. The *information matrix* $\mathbf{J}^\mathsf{T}\mathbf{J}$ states how sharply the rows together fix the state, direction by direction, and its inverse is the posterior *covariance* $\boldsymbol{\Sigma}$, whose diagonal holds the variances of the entries and whose off-diagonal entries hold how they vary together. The *Laplace approximation* replaces the posterior by the normal distribution centered on its most probable state with that covariance. A *factor graph* draws the entries of the state and the rows that act on them, one node each.

Key idea

A graph describes the asset and the plant around it; a probabilistic question states what is to be learned about it. The plant graph, its pinned and free quantities and its physical rows are built once, and each question is a probability over the state of that one graph, with the unknown on the left of the conditioning bar and the evidence on the right. The posterior of the state is carried by three objects, the most probable state $\hat{\mathbf{z}}$, the covariance $\boldsymbol{\Sigma}$ and the Jacobian $\mathbf{J}$ of the rows, and every answer is an entry of the state read off them or a function of the state carried through them.

1.1.1 Several questions about one equipment hall

An asset manager responsible for the hall might ask any of the following:

1. What temperature is the hall running at now?
2. How far is the as-commissioned conductance of the air path from its design value?
3. Can the path carry the doubled load within its limit?
4. Is the hall sensor drifting, or is the path fouled?
5. Would another sensor sharpen the health index?
6. What happens if the path loses half its conductance at once?

The questions are different. The first two ask for the value of an unknown quantity; the third asks for the probability that a future temperature stays within a limit; the fourth asks which of two descriptions of the record is true; the fifth asks what a measurement not yet taken would be worth; the sixth asks about an operating condition that has not occurred. All six, however, concern the same physical asset: the same hall, the same air path, the same sensor.

Constructing a separate mathematical model for every question would repeat the same physics six times, and the six models could disagree about it. The method of this book begins instead with one common representation of the equipment and its physical relations, the *plant graph*. It is step 0 of the method stated in *How to read this book*: the system as it is, before any question is asked, built once and kept for every question that follows. Section 1.1.6 writes the questions that this chapter answers over the graph.

1.1.2 The system, the question and the evidence

Three things must be fixed before any probabilistic inference is performed, and the plant graph is where they are fixed. The first is the division of the quantities: which are known and pinned (the supply-air temperature, the heat load), which are unknown and free (the hall temperature, the conductance), and which are observed (the reading of the commissioning test). The field often separates the unknowns into *states*, such as temperatures and flows, and *parameters*, such as conductances and rates; in this book both kinds are entries of one state vector, and the inference treats them alike. The second is the set of physical relations that must hold among the quantities, such as conservation of energy at the hall. The third is the evidence: the readings, the records and the prior knowledge, attached to the graph as rows of their own.

The question is a separate object. It is stated over the graph after the graph is built, and it decides what is computed, not what the system is. The separation makes the asset manager's assessments systematic: when a new question arrives, the graph is already in hand, and the work is to write the question over its state.

> **Principle 1.1.** *The physical model is reusable; the probabilistic question determines what is calculated from it. A plant graph is built once from the equipment and its physics, and every question is a probability or an expectation over the state of that graph.*

1.1.3 Larger plant graphs and factor graphs

The equipment hall of this chapter has one thermal path and two unknown quantities, the smallest graph with every kind of element. The same construction serves a plant of many components, and the graph on which an inference operates is a second object, drawn from the plant graph and the evidence attached to it.

A cooling network. Real plants may include hundreds or thousands of components, each with its own equations and uncertainties. Consider a cooling network with two chillers, two pumps, a common electrical supply and several temperature sensors. Its plant graph may contain temperatures, pressures, flows and electrical states at its nodes; physical components represented by paths or by component relations; heat, mass and power conservation equations; sensor measurements and historical observations; failure and repair states of equipment; and shared parameters and common-cause dependencies.

Figure 1.1 draws a reduced version. The green boxes are free nodes: the two chillers and the two pumps, each carrying its status and capacity or flow, the supply header with its water temperature T_s and flow, and the hall with its temperature T. The gray box on the left is the pinned outdoor air, to which the chillers reject heat; the white box in the middle is the common electrical supply, which feeds all four machines; the white box at the upper right is

the heat load of the hall. Solid green arrows are physical paths of heat and water; thin gray arrows are electrical feeds; dashed arrows are observations, two temperature readings and the maintenance record of a chiller. The common supply is a single node on which four components depend: if it fails, all four stop together, and a graph that omitted it would treat their failures as independent.

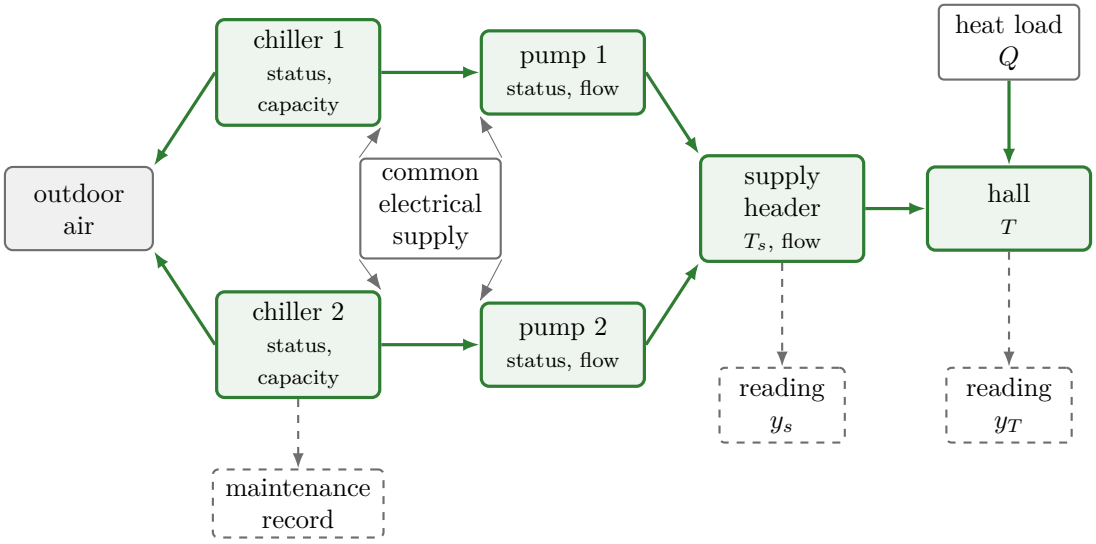


Figure 1.1: Plant graph of a reduced cooling network (illustrative; return piping omitted). Free: the status and capacity of each chiller, the status and flow of each pump, the header water temperature T_s and flow, and the hall temperature T. Pinned: the outdoor air, to which the chillers reject heat, and the heat load Q; the common electrical supply feeds all four machines and is the shared dependency. Observed: T_s and T through two readings, and chiller 2 through its maintenance record.

The mathematical principle is the one the single hall shows below. The network is a collection of physical relations,

$$r_1(\mathbf{z}) = 0, \quad r_2(\mathbf{z}) = 0, \quad \ldots, \quad r_m(\mathbf{z}) = 0,$$

one per balance and per component relation, over a state vector $\mathbf{z}$ that now has many entries. Observations and probability models for the uncertain quantities are attached as rows of their own. What is ultimately sought is the joint posterior of the state,

$$p(\mathbf{z} \mid \mathbf{y}, M),$$

where M represents the assumed system model, including its structure: which components exist, how they are connected and which dependencies are represented. Writing M explicitly matters because the structure itself may be uncertain, and a question about the structure is answered by comparing two such models, not by the posterior inside either.

Physical graphs and factor graphs. A physical graph and a probabilistic factor graph are related, but they are not the same object. The physical graph represents equipment and connections: halls, paths, chillers, pumps, sensors. A *factor graph* represents the variables of the inference and the mathematical relations among them, including the observations, the priors and the dependencies: each entry of the state is one node, each row of the model is one factor node, and each factor is linked to the entries its row involves. It is constructed from the physical graph and the evidence attached to it, and it shows which row acts on which unknown. The distinction becomes important when competing system structures and diagnosis are studied in later chapters. Two physical drawings of the same plant, one with a single air path and

one with two, give two factor graphs with different variables and rows, and comparing them is comparing two models M. A sensor bias adds a variable to the factor graph without changing the equipment at all.

1.1.4 The plant graph (step 0)

The simplest plant graph with every kind of element has one free node, one pinned node, one path, one source and one observation, and the equipment hall is that graph. It is built before any question is asked and before any number is fitted.

Nodes, a path, a source and an observation. Consider a simplified equipment hall with one heat-removal path, the hall's air path to the supply air that the cooling unit delivers. Figure 1.2 draws it as a plant graph. The green box on the left is the hall, a free node whose air temperature T is unknown; the gray box on the right is the supply air, a pinned node whose temperature $T_c = 18°\mathrm{C}$ is held by the cooling unit's discharge controller. The solid green arrow between them is the air path, the asset of this chapter, characterized by its effective thermal conductance G. The box above the hall is the heat load Q supplied by the servers, a source. The dashed box below is the hall sensor, read at the commissioning test, and the dashed arrow from the hall to it is an observation: the sensor reports a value related to T, not T itself. Solid connections are physical relations; the dashed connection is evidence.

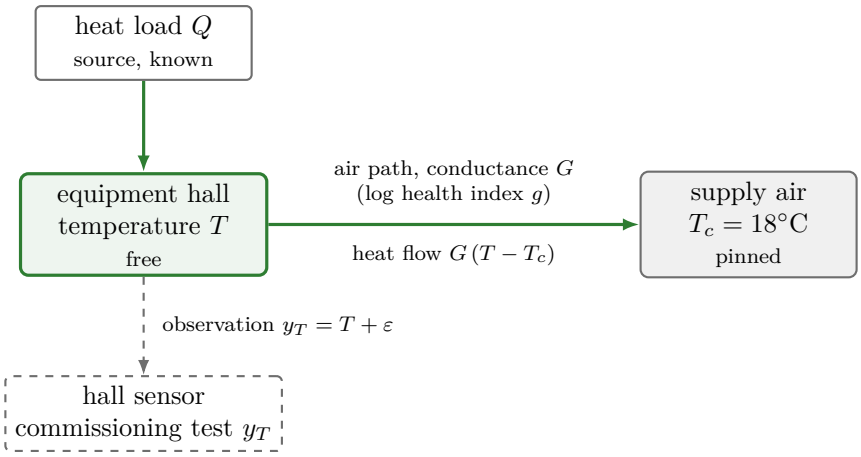


Figure 1.2: Plant graph of the equipment hall in steady state. Free: the hall air temperature T and the conductance G of the air path, carried as the log health index $g = \ln(G/G_{\mathrm{ref}})$ (Section 1.1.4). Pinned: the supply-air temperature $T_c = 18°\mathrm{C}$, held by the cooling unit's discharge controller, and the heat load Q. Observed: T, through the commissioning-test reading y_T with error ε. Solid arrows are physical relations; the dashed arrow is an observation.

Table 1.1 lists the elements of the graph. The first column names each element, the second its role in the graph, the third its symbol and the fourth its unit; the last column states how the element enters this first example.

The modeling choices of the graph. Three choices fix the graph before any question is asked, and the effective conductance is then the one uncertain property of the path. The hall is taken in steady state through the commissioning test, storing no more and no less heat from one moment to the next: this is what makes one reading at one load a statement about the conductance, and it fails in the period after a change of load or of conductance (Section 1.6.5). The heat load is pinned at its 30 kW: this is what leaves the conductance as the one unknown of the path, and it fails if the load is not known. The supply air is pinned at $T_c = 18°\mathrm{C}$ because

Table 1.1: The elements of the plant graph of Figure 1.2.

Element	Role	Symbol	Unit	In the example
Equipment hall	free node	T	°C	air temperature unknown
Supply air	pinned node	T_c	°C	held at 18°C by the discharge controller
Heat load	source	Q	kW	thermal power supplied by the servers, known
Air path	path (edge)	G	kW/K	effective conductance, uncertain; the asset
Hall sensor	observation	y_T	°C	the commissioning-test reading of the hall temperature

the discharge controller holds it at its set point: this is what makes it a boundary rather than an unknown, and it fails if the unit saturates and can no longer hold it (Section 1.6.5). The conductance lumps every mechanism by which the hall's heat reaches the supply air, the airflow through the cooling unit and the mixing of the hall air alike, into one number; how well that lumping represents the hall is a modeling assumption, examined in Section 1.6.5. This is the starting physical model.

Conservation of energy at the hall node. A graph becomes useful for computation when the mathematical relations between its elements are specified. In the steady state of the modeling choices the hall's temperature is constant, and conservation of energy at the hall node requires the heat entering to equal the heat leaving,

$$Q_{\text{in}} = Q_{\text{out}},$$

where Q_{in} and Q_{out} are thermal powers in kilowatts. The heat entering is the load, $Q_{\text{in}} = Q$. The heat leaving is modeled as proportional to the temperature difference across the path,

$$Q_{\text{out}} = G\,(T - T_c),$$

the defining relation of the effective conductance: G is the heat flow per kelvin of difference between the hall and the supply air. Combining the two,

$$Q = G\,(T - T_c), \tag{1.1}$$

where Q is the heat generated by the servers in kW, G the effective thermal conductance in kW/K, T the hall air temperature in °C and T_c the supply-air temperature in °C. A temperature difference has the same value in kelvin and in degrees Celsius, so the units balance: kW/K times K is kW.

Equation (1.1) is a *physical relation*, not a probability statement. It states what must hold if the simplified thermal model is correct, whatever is or is not known about its quantities. Since $G > 0$, dividing by G and adding T_c gives

$$T = T_c + \frac{Q}{G}, \tag{1.2}$$

which gives the temperature when the other three quantities are specified. The hall is warmer than the supply air by Q/G: more load or less conductance raises it.

The forward calculation. The hall is operated at the load $Q = 30\,\mathrm{kW}$, and its air path was designed for a conductance of $G = 6\,\mathrm{kW/K}$, the design value on its design sheet. With $T_c = 18°\mathrm{C}$, equation (1.2) gives

$$T = 18 + \frac{30\,\mathrm{kW}}{6\,\mathrm{kW/K}} = 18 + 5 = 23°\mathrm{C},$$

the steady-state temperature of the hall with a path built exactly to its design. The calculation used no probability, estimated no uncertainty and used no observational evidence. It is a deterministic *forward calculation*: from pinned inputs and a specified conductance to the temperature they imply. Every later chapter needs it, because every inference compares temperatures computed this way with what the sensor reports. The temperature falls as the conductance grows, with the slope $\partial T/\partial G = -Q/G^2$, which at the design point is $-30/36 = -0.833\,\mathrm{K}$ per kW/K: an error of $1\,\mathrm{kW/K}$ in the conductance moves the computed temperature by almost a kelvin. Doubling the load doubles the excess $T - T_c$ at every conductance: at $G = 6\,\mathrm{kW/K}$ the hall at $60\,\mathrm{kW}$ would be at 28°C.

The state vector. The forward calculation required the conductance to be known. In the inference problems of this book it is not, and neither is the hall temperature. The *state vector* is the collection of the unknown quantities that the inference estimates at once. In the equipment hall the unknowns are the hall temperature and the thermal conductance, and one possible representation is

$$\mathbf{z} = \begin{bmatrix} T \\ G \end{bmatrix}.$$

The state holds everything the inference treats as free; the pinned quantities T_c and Q are not in it, and neither is the reading y_T, which is evidence. The concept recurs throughout the book: the graph fixes the state, the evidence constrains it, and each question is a statement about it. Quantities that are entries of the state, such as T here, are answered directly by the inference; quantities that are functions of the state, such as the heat flow through the path, require one further transformation.

The log-conductance. A conductance is positive, and an estimate of it must be positive too. In practice the conductance is therefore often represented by a logarithmic entry,

$$g = \ln\!\left(\frac{G}{G_{\mathrm{ref}}}\right),$$

where $G_{\mathrm{ref}} > 0$ is a reference conductance in kW/K and g is dimensionless. Inverting,

$$G = G_{\mathrm{ref}}\,\mathrm{e}^{g},$$

which is positive for every real g: an inference free to move g anywhere on the real line can never produce a negative or zero conductance. A change in g is a relative change in G, since $\mathrm{d}G = G_{\mathrm{ref}}\,\mathrm{e}^{g}\,\mathrm{d}g = G\,\mathrm{d}g$, so that $\mathrm{d}g = \mathrm{d}G/G$; uncertainty in g reads directly as a fractional uncertainty in G. The reference value only sets the origin of g. In this chapter it is taken as $G_{\mathrm{ref}} = 6\,\mathrm{kW/K}$, the design conductance on the design sheet and in the forward calculation, so that $g = 0$ there, $G = 7.5\,\mathrm{kW/K}$ corresponds to $g = \ln 1.25 = 0.2231$, and a conductance 10 % above the reference to $g = \ln 1.1 = 0.0953$. The design value is chosen as the origin and not attached as a row: it states what the path was specified to deliver, not what it delivers, so it carries no evidence about the conductance, and the fit would be the same with another origin. Read against the same reference, the exponential of the entry is the *health index* of the path (in the framework, a parameter of the physics read against its design value),

$$h = \frac{G}{G_{\mathrm{ref}}} = \mathrm{e}^{g}, \tag{1.3}$$

the fraction of its design conductance the path has: $h = 1$, $g = 0$, is a path built to design, and the handbook's 7.5 kW/K is $h = 1.25$. The state vector becomes

$$\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} T \\ g \end{bmatrix}, \tag{1.4}$$

with entries $z_1 = T$ in °C and $z_2 = g$, dimensionless. It is the one state of this chapter: every question below is a statement about these two entries.

The physics row. The thermal relation can now be written in terms of the state. Substituting $G = G_{\text{ref}}\,\mathrm{e}^{g}$ into equation (1.1),

$$Q = G_{\text{ref}}\,\mathrm{e}^{g}\,(T - T_c),$$

or, moving everything to one side,

$$r_{\text{phys}}(\mathbf{z}) = G_{\text{ref}}\,\mathrm{e}^{g}\,(T - T_c) - Q = 0. \tag{1.5}$$

The function $r_{\text{phys}}(\mathbf{z})$ is a *residual*, in kW. A residual measures how far a candidate state is from satisfying a relation: it is zero when the relation holds exactly and grows as the candidate departs from it. With $G_{\text{ref}} = 6$ kW/K, $T_c = 18$°C and $Q = 30$ kW, the candidate $\mathbf{z} = (23°\text{C},\ 0)$ gives $r_{\text{phys}} = 6 \times 1 \times 5 - 30 = 0$: it is the forward calculation. The candidate $\mathbf{z} = (23.2°\text{C},\ 0)$ gives $6 \times 5.2 - 30 = 1.2$ kW: at that temperature and conductance the path would remove 1.2 kW more than the servers supply, so the hall could not be in steady state. The candidate $(23.2°\text{C},\ \ln(30/31.2)) = (23.2°\text{C},\ -0.0392)$ again gives zero: the temperature 23.2°C is consistent with the physics if the conductance is $30/5.2 = 5.769$ kW/K. The zero set of r_{phys} is the curve $T = T_c + Q/(G_{\text{ref}}\,\mathrm{e}^{g})$ in the plane of (g, T); the physics alone allows every point of that curve and no point off it.

The same row has a second form, in the units of the quantity it fixes. Dividing equation (1.5) by $G_{\text{ref}}\,\mathrm{e}^{g} = G$, which is positive, leaves the zero set unchanged and gives the row in kelvin, the departure of a candidate temperature from the temperature the physics implies. Divided further by a *model tolerance* σ_p, the temperature error in kelvin that the physics is allowed, it becomes a dimensionless row,

$$r_1(\mathbf{z}) = \frac{T - T_c - (Q/G_{\text{ref}})\,\mathrm{e}^{-g}}{\sigma_p}, \tag{1.6}$$

which counts the violation of the physics in multiples of its tolerance. This is the form in which the physics row enters every computation of the book. The formulation extends to systems of any size. A larger physical system is represented by a collection of residuals, one per relation: an electrical network has rows for Kirchhoff's current and voltage laws, a fluid network rows for mass conservation at its junctions and for the pressure losses along its pipes, and a cooling plant rows for the energy balance at each of its nodes. Each row has the form of equation (1.5), and the inference treats them alike.

Exact constraints and model discrepancy. Physical rows and reading rows should not automatically be treated as equivalent kinds of noisy measurement. Conservation of energy is a law: if the hall is in steady state and the model has captured every path, equation (1.5) holds exactly, and it can be imposed as an *exact constraint*, $r_{\text{phys}}(\mathbf{z}) = 0$, which restricts the state to the zero set of the row. In a computation the exact constraint is carried as the physics row (1.6) with a model tolerance σ_p far smaller than every other spread of the problem, and exactness is the limit $\sigma_p \to 0$; the limit is checked by showing that the results stop moving as σ_p shrinks, never by using the row to eliminate an entry of the state. A reading row, by contrast,

is meant to be violated by an amount of the order of the sensor's error: the reading is not the temperature.

Approximate physics is a third case. The single effective conductance of equation (1.1) is an idealization; the real hall has several paths, a conductance that varies with airflow, and stored heat. When the relation is only approximately true, the honest model is

$$G_{\mathrm{ref}}\, \mathrm{e}^{g}\, (T - T_c) - Q = \Delta_{\mathrm{m}},$$

with an explicit *model discrepancy* Δ_{m} in kW, an entry of its own with a stated prior, for example $\Delta_{\mathrm{m}} \sim \mathcal{N}(0, \sigma_{\mathrm{m}}^2)$. Algebraically this is a physics row with the weight $1/\sigma_{\mathrm{m}}^2$, and in the objective it looks like a reading row. It behaves differently. A measurement error is new at each reading and averages out over many readings; a discrepancy is a property of the model and repeats wherever the same operating condition repeats, so that many readings at one condition do not reduce it. Treating a discrepancy as measurement noise therefore makes the inference overconfident, and treating a law as noisy lets the inference trade physics against data. The choice between an exact constraint and an explicit discrepancy is made when the graph is built, and it is stated with the model. In this chapter conservation of energy is exact.

1.1.5 The record and the handbook value

The record of the hall is one reading, taken at the commissioning (acceptance) test of the air path: window 0 of the path's record. The hall sensor, the one temperature sensor installed in the hall, reported at the load of 30 kW

$$y_T = 23.2^{\circ}\mathrm{C}.$$

The sensor is not perfectly accurate. Its measurement model is

$$y_T = T + \varepsilon, \qquad \varepsilon \sim \mathcal{N}(0, \sigma_T^2),$$

where ε is the measurement error in kelvin, assumed normal with zero mean, and σ_T its standard deviation, which the sensor's specification gives as

$$\sigma_T = 0.5\,\mathrm{K}.$$

The reading is the only observation: the supply-air temperature and the load are pinned, and nothing measures the conductance. The *commissioned value* $\hat{g}_0$ is the fit of this test, the most probable value of the log-conductance under the posterior of step 3, written $\hat{g}$ in this chapter of one window; it is the baseline against which later condition tests are read. Chapter 4 adds a later condition test of the same path at window 24, and Chapter 5 reads each coil of a larger hall against its own commissioning fit.

Knowledge of the conductance that does not come from the reading is also in hand. A design handbook, or earlier tests of halls of the same kind, give an estimate of the effective conductance with an uncertainty; for this hall (illustrative) the handbook gives 7.5 kW/K, with the true conductance within a factor of about 1.35 of it with probability of about 0.68. The handbook value is not a measurement of this hall, and it is not the design value either: the design sheet states what the path was specified to deliver, the handbook what paths of its kind are found to deliver. It is information held before the reading, and it enters the model as a prior.

1.1.6 The questions (step 1)

The asset manager's questions can now be written over the graph, each as a probability or an expectation over the state, with the unknown on the left of the bar and the evidence on the

right. Table 1.2 lists the five that this chapter answers: the first column is the question in the asset manager's words; the second the general form that every question of its category takes, written in the book's notation for a quantity $f(\mathbf{z})$ derived from the state, a threshold f^* on it and the posterior density $p(\mathbf{z} \mid \mathbf{y})$; the third its mathematical form for the hall; the fourth its category among the five named in *How to read this book*, which Chapter 2 develops; and the last the subsection where it is answered. The first two are estimation, of an entry of the state and of a quantity derived from one, the health index $h = G/G_{\text{ref}} = \mathrm{e}^{g}$ of equation (1.3), with the probability that it lies below one, below design. The third and fourth are predictions under a threshold, tail probabilities of the posterior: the third at the present load, the fourth at a load of 60 kW at which the hall has never run, whose temperature T_2 is a quantity the graph acquires in step 2. The fifth is a design question about a measurement not yet taken; it asks nothing new of the present evidence, and it is answered as the next measurement for the first four.

Table 1.2: The questions asked of the equipment hall at its commissioning test, the general form of their category, their mathematical form for the hall, their category and where each is answered. h is the health index of equation (1.3), T_2 the hall temperature at the load $Q_2 = 60\,\text{kW}$ (Section 1.1.7); $\mathbf{\Sigma}_a$ the covariance expected after a candidate row with Jacobian row $\mathbf{J}_a$.

Asset manager's question	General form	Mathematical form	Category	Answered in
What temperature is the hall running at now?	$\mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}]$, $\mathrm{Var}(f(\mathbf{z}) \mid \mathbf{y})$	$p(T \mid \mathbf{y})$, $T = z_1$	estimation	1.3.1
How far is the as-commissioned conductance from its design value?	$\mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}]$, $\mathrm{Var}(f(\mathbf{z}) \mid \mathbf{y})$, $\Pr(f(\mathbf{z}) \le f^* \mid \mathbf{y})$	$p(h \mid \mathbf{y})$, $h = G/G_{\text{ref}} = \mathrm{e}^{z_2}$; $\Pr(h < 1 \mid \mathbf{y})$	estimation of a derived quantity	1.3.1
Is the hall within its 24°C limit now?	$\Pr(f(\mathbf{z}) < f^* \mid \mathbf{y})$	$\Pr(T < 24^\circ\text{C} \mid \mathbf{y})$	prediction under a threshold	1.3.2
Can the path carry 60 kW within 27.5°C?	$\Pr(f(\mathbf{z}) < f^* \mid \mathbf{y})$	$\Pr(T_2 < 27.5^\circ\text{C} \mid \mathbf{y})$, $T_2 = T_c + (Q_2/G_{\text{ref}})\,\mathrm{e}^{-z_2}$	prediction under a threshold	1.3.2
Which added sensor would narrow the health index?	preposterior covariance $(\mathbf{\Sigma}^{-1} + \mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a)^{-1}$	$\mathbf{\Sigma}_a$ for each candidate row	design	1.6.1

The questions differ in what they compute and agree in what they are computed over: the same two entries of the state, under the same posterior. Two of the asset manager's six questions are not in the table. Whether the hall sensor is drifting or the path is fouled compares two complete models of the same reading, the plant with a fouled path read by a true sensor and the plant with a bias entry on the reading row, and its answer is a ratio of their evidences rather than a property of either posterior; when every sensor of a larger plant is a candidate, the same comparison run over the candidates is a diagnosis. Chapter 7 answers it, and the residual check of Section 1.4.3 takes a first, coarse look at it. What happens if the path loses half its conductance at once is a question about a transient that the steady-state graph cannot represent, and Section 1.6.5 states what the graph would need for it.

1.1.7 The evidence as rows (step 2)

Step 2 attaches the reading and the handbook value to the graph as rows on the state $\mathbf{z} = (T, g)^{\mathsf{T}}$, and extends the graph with the operating condition the fourth question names. The three rows,

the physics row of Section 1.1.4 among them, are Gaussian, and each is written normalized, with a residual and a Jacobian row.

The modeling choices. The energy balance is taken as exact, the limit $\sigma_p \to 0$ of the physics row: this is what confines the posterior to the physics curve, so that a reading of the temperature informs the conductance, and it fails if heat leaves the hall by a mechanism the single conductance does not lump or if the hall is not in steady state (Section 1.6.5); Section 1.4.1 checks that the tolerance used has reached the limit.

The reading row. The reading of Section 1.1.5 becomes the normalized observation residual

$$r_{\text{obs}}(\mathbf{z}) = \frac{T - y_T}{\sigma_T}, \tag{1.7}$$

which is dimensionless: the difference between the candidate temperature and the reading, in units of the sensor's standard deviation. At $T = 23.2°\text{C}$ it is zero; at the design temperature $T = 23°\text{C}$ it is $(23 - 23.2)/0.5 = -0.4$. The row carries the probability of the reading. Given the state, the reading differs from T only by the error, so its density is the normal density of ε evaluated at $y_T - T$:

$$p(y_T \mid \mathbf{z}) = \frac{1}{\sqrt{2\pi}\,\sigma_T} \exp\left[-\frac{(y_T - T)^2}{2\sigma_T^2}\right] = \frac{1}{\sqrt{2\pi}\,\sigma_T} \exp\left[-\tfrac{1}{2}\, r_{\text{obs}}(\mathbf{z})^2\right].$$

Read as a function of the state for the fixed reading, this is the *likelihood* of the state. A candidate state with a small $|r_{\text{obs}}|$ makes the reading probable; the ratio of the likelihoods of $T = 23°\text{C}$ and $T = 23.2°\text{C}$ is $\mathrm{e}^{-0.4^2/2} = \mathrm{e}^{-0.08} = 0.923$, so the reading alone hardly distinguishes the two. The likelihood does not depend on g: the sensor sees the temperature, and it informs the conductance only through the physics row that ties the two together.

The prior row. The handbook value becomes a row of the same form. Written as a normal distribution on the log-conductance, $g \sim \mathcal{N}(\mu_g, \sigma_g^2)$, it is the row $(g - \mu_g)/\sigma_g$: a handbook value is treated as a reading of g taken before the commissioning test. The handbook conductance of 7.5 kW/K gives $\mu_g = \ln(7.5/6) = 0.2231$, and $\sigma_g = 0.3$ states that the true conductance is within a factor $\mathrm{e}^{0.3} = 1.35$ of it with probability of about 0.68. The prior is written $p(g)$ in general; its form is a modeling choice that must be stated and justified, and Section 1.4.4 tests how much this one weighs.

Three kinds of row. The model now holds three very different kinds of relation. Table 1.3 sets them side by side. Each row of the table is one kind of relation; the columns give its equation for the hall, its residual and its role.

Table 1.3: The three kinds of row in the equipment hall model, with $G = G_{\text{ref}}\,\mathrm{e}^{g}$.

Row	Equation	Residual	Role
Physics	$G\,(T - T_c) - Q = 0$	$G_{\text{ref}}\,\mathrm{e}^{g}(T - T_c) - Q$	constrains what is physically possible
Reading	$y_T = T + \varepsilon$	$(T - y_T)/\sigma_T$	provides imperfect information about the state
Prior	$g \sim \mathcal{N}(\mu_g, \sigma_g^2)$	$(g - \mu_g)/\sigma_g$	carries knowledge held before the readings

The distinction between the first two rows is foundational. The physical equation constrains the system: it says which combinations of temperature and conductance can occur at all.

The reading provides evidence about the unknown quantities: it says which of the possible combinations are plausible in light of what was measured. The prior row has the form of a reading row and the role of knowledge held in advance. A probabilistic inference method combines the three kinds of relation to obtain a distribution over the plausible states.

Rows and weights. Chapter 3 shows that, for normal errors, the negative logarithm of the posterior is, up to a constant, the objective $\mathcal{L}(\mathbf{z}) = \frac{1}{2}\sum_k w_k\, r_k(\mathbf{z})^2$, a weighted sum of the squared rows with weights $w_k = 1/\sigma_k^2$. A reading row may be written either normalized, as in equation (1.7), with unit weight, or in kelvin, as $T - y_T$ with weight $1/\sigma_T^2$; both put the same term $(T - y_T)^2/\sigma_T^2$ into the objective. Every row of this chapter is written normalized: the physics row in the form (1.6) with the tolerance σ_p, the reading row (1.7) and the prior row, so that every row counts in standard deviations and none dominates because of its unit.

The scenario node. The fourth question doubles the electrical load from 30 kW to 60 kW, an operating condition at which the hall has never run. A second operating condition is added to the graph, sharing the same unknown conductance. Figure 1.3 draws it. The upper row is the commissioning-test condition, at which the hall has been operated and observed: the free node at $Q_1 = Q = 30\,\mathrm{kW}$ with temperature T, its path to the pinned supply air, and its sensor. The lower row is the requested-load condition, a scenario node at $Q_2 = 60\,\mathrm{kW}$ with temperature T_2 and its own path to the same supply air. The vertical dashed line between the two paths marks that they carry one and the same conductance. The scenario node has no sensor: it is in the graph only because the question asks about it.

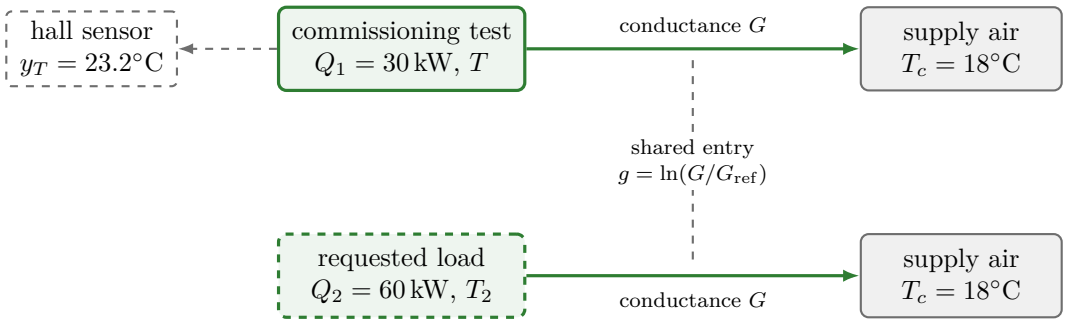


Figure 1.3: Two operating conditions of the equipment hall linked through the same physical conductance. Free: the commissioning-test temperature T and the shared log-conductance g, the two entries of the state; the requested-load temperature T_2 (dashed node) is a function of g. Pinned: the supply-air temperature and the two loads. Observed: T only, through the reading y_T; the scenario node is unobserved and enters the graph because the question asks about it.

The two operating conditions satisfy the energy balance of equation (1.1) each at its own load,

$$Q_1 = G\,(T - T_c), \tag{1.8}$$
$$Q_2 = G\,(T_2 - T_c), \tag{1.9}$$

with $Q_1 = Q$. The two temperatures are different, but the conductance is shared, and that connection is what makes the second condition predictable from the first. The scenario node adds no evidence and no free quantity: by equation (1.9) its temperature is fixed by the conductance, $T_2 = T_c + (Q_2/G_{\mathrm{ref}})\,\mathrm{e}^{-g}$, a function of the state $\mathbf{z} = (T, g)^{\mathsf{T}}$, which keeps its two entries and its three rows. Chapter 4 carries T_2 instead as a third entry with a second physics row, the same formulation; with the same rows it gives the same answers.

Which row acts on which entry. The three rows act on the two entries of the state in a pattern fixed by the graph, before any number is computed. The physics row ties the temperature to the conductance and involves both; the sensor sees the temperature and not the conductance; the handbook speaks about the conductance and not the temperature; the scenario node adds no row. The Jacobian of the rows, the matrix of their derivatives with respect to the entries, can therefore be nonzero only where marked $*$:

$$\mathbf{J} = \begin{bmatrix} * & * \\ * & 0 \\ 0 & * \end{bmatrix} \begin{matrix} \text{physics row} \\ \text{reading row,} \\ \text{prior row} \end{matrix} \qquad \text{columns } z_1 = T,\ z_2 = g. \tag{1.10}$$

The zeros are structural: they hold at every state. Only the physics row has entries in both columns, and it is the one row through which a reading of the temperature reaches the conductance.

The factor graph of the hall. The pattern (1.10) is what a factor graph (Section 1.1.3) draws. Figure 1.4 draws it for the hall with its scenario node. The two solid circles are the entries of the state, T and g. The black squares are the rows: the physics row of equation (1.6), linked to both entries; the reading row, linked to T alone; and the prior row, linked to g alone. The dashed circle is the full-load temperature T_2, joined to g by a dashed line: it is a function of g through equation (1.9), not an entry, and no row acts on it. The figure makes visible what the physical graph only implies: T_2 has no reading of its own and is connected to the evidence only through g, and g is connected to the reading only through the physics row. The physical drawing has two supply-air boxes and two paths; the factor graph has no node for the supply air at all, since T_c is pinned and appears only as a constant inside the physics row.

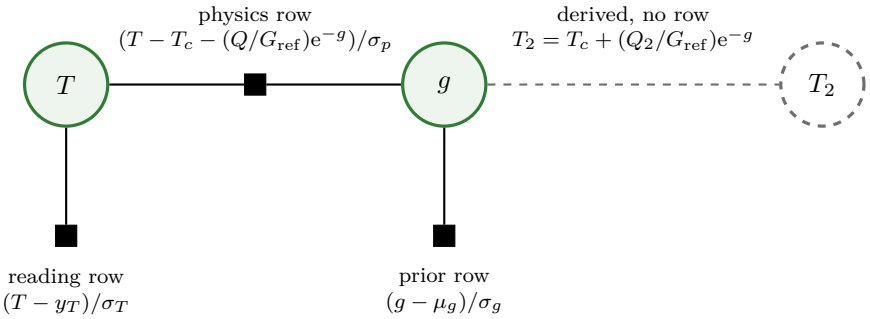


Figure 1.4: Factor graph of the equipment hall with its scenario node. Solid circles: the entries of the state, T and g. Squares: the rows of the model, each linked to the entries it involves, the pattern (1.10) of the Jacobian. Dashed: the full-load temperature T_2, a function of g with no row of its own, which reaches the evidence only through the shared g.

The residual vector. The state is the column vector of equation (1.4),

$$\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} T \\ g \end{bmatrix},$$

with T in °C and g dimensionless. Three rows act on it, the physics row (1.6), the reading row (1.7) and the prior row, each divided by its standard deviation, so that every row counts in standard deviations and none dominates because of its unit. Stacked into one residual vector, the three rows are

$$\mathbf{r}(\mathbf{z}) = \begin{bmatrix} r_1(\mathbf{z}) \\ r_2(\mathbf{z}) \\ r_3(\mathbf{z}) \end{bmatrix} = \begin{bmatrix} (T - T_c - (Q/G_{\text{ref}})\,\mathrm{e}^{-g})/\sigma_p \\ (T - y_T)/\sigma_T \\ (g - \mu_g)/\sigma_g \end{bmatrix}, \tag{1.11}$$

with $T_c = 18°\mathrm{C}$, $Q = 30\,\mathrm{kW}$, $G_{\mathrm{ref}} = 6\,\mathrm{kW/K}$, the reading $y_T = 23.2°\mathrm{C}$ with $\sigma_T = 0.5\,\mathrm{K}$, and the handbook prior $\mu_g = \ln(7.5/6) = 0.2231$ with $\sigma_g = 0.3$. Each entry of $\mathbf{r}$ is one row: a dimensionless number stating by how many standard deviations a candidate state violates one relation. The objective is half the sum of their squares,

$$\mathcal{L}(\mathbf{z}) = \tfrac{1}{2}\,\mathbf{r}(\mathbf{z})^{\mathsf{T}}\,\mathbf{r}(\mathbf{z}) = \tfrac{1}{2}\sum_{k=1}^{3} r_k(\mathbf{z})^2, \tag{1.12}$$

which Chapter 3 shows to be minus the logarithm of the posterior up to a constant, so that the most probable state $\hat{\mathbf{z}}$, the maximum a posteriori (MAP) state, is the minimizer of $\mathcal{L}$. Conservation of energy is exact in this model, and an exact constraint is the limit $\sigma_p \to 0$ of the physics row (Section 1.1.4). The computation takes $\sigma_p = 0.01\,\mathrm{K}$, fifty times tighter than the sensor, and Section 1.4.1 checks that the limit has been reached.

The Jacobian. The Jacobian $\mathbf{J}$ collects the derivative of every row with respect to every entry of the state: its entry in row k and column i is $\partial r_k/\partial z_i$, so it has one row per row of the model and one column per entry of the state, here three by two. Differentiating equation (1.11) entry by entry,

$$\mathbf{J}(\mathbf{z}) = \frac{\partial \mathbf{r}}{\partial \mathbf{z}} = \begin{bmatrix} 1/\sigma_p & (Q/G_{\mathrm{ref}})\,\mathrm{e}^{-g}/\sigma_p \\ 1/\sigma_T & 0 \\ 0 & 1/\sigma_g \end{bmatrix}, \tag{1.13}$$

the pattern (1.10) with its entries filled in. Each entry is a sensitivity in standard deviations: how far one row moves when one entry of the state moves by one unit. Only the entry in the first row and second column depends on the state; the others are constants. The first row is the physics row's own sensitivities in its physical unit, rescaled. Differentiating the kilowatt form of equation (1.5),

$$\frac{\partial r_{\mathrm{phys}}}{\partial T} = G_{\mathrm{ref}}\,\mathrm{e}^{g} = G, \qquad \frac{\partial r_{\mathrm{phys}}}{\partial g} = G_{\mathrm{ref}}\,\mathrm{e}^{g}\,(T - T_c) = G\,(T - T_c),$$

in kW/K and kW; on the zero set the second equals the load Q by equation (1.1). A one-kelvin error in the temperature unbalances the hall by G kilowatts, and a unit change in g, a factor e in the conductance, by about Q kilowatts to first order. Divided by G, to reach the temperature form, and by σ_p, the pair becomes $(1/\sigma_p,\ (Q/G)/\sigma_p)$, the first row of equation (1.13) on the zero set.

1.2 The posterior of the state (step 3)

All the questions of Table 1.2 are read from one posterior: the posterior of the state that the physics row, the reading row and the prior row of step 2 give, approximately the Laplace normal of (T, g) and exactly a density along the physics curve. This section derives it and computes it for the hall; Section 1.3 reads every answer from it.

Four objects of the fit recur in the specialized forms, and they are named here by their roles. The *MAP state* $\hat{\mathbf{z}} = (\hat{T}, \hat{g})^{\mathsf{T}}$, the maximum a posteriori state, is the minimum of the objective that the three rows put on the state, the most probable state; the *information matrix* $\mathbf{\Lambda} = \mathbf{J}^{\mathsf{T}}\mathbf{J}$ is the curvature that the rows give that objective at the minimum, with $\mathbf{J}$ the 3×2 Jacobian of the normalized rows; the *Laplace covariance* $\mathbf{\Sigma}$, a 2×2 matrix, is the inverse of that curvature after a term from the physics row's own curvature is added to it, and it makes the posterior of the state approximately normal, $\mathbf{z} \mid \mathbf{y} \approx \mathcal{N}(\hat{\mathbf{z}}, \mathbf{\Sigma})$; and a derived quantity f is carried by its *gradient* $\mathbf{g} = \partial f/\partial \mathbf{z}$ at $\hat{\mathbf{z}}$ (bold, to tell it from the entry g), with the variance $\mathbf{g}^{\mathsf{T}}\mathbf{\Sigma}\mathbf{g}$. Section 1.2.1 derives the four from the rows. Beside them stands the exact posterior that

this problem admits in the limit $\sigma_p \to 0$: on the physics curve the state has one free direction, and the posterior is the density of one entry, evaluated by quadrature. The Laplace route, a normal distribution centered on the MAP state, is what lets every answer be read from $\hat{\mathbf{z}}$ and $\boldsymbol{\Sigma}$ by operations that serve a plant of any size, and it misleads where the posterior is skewed along the curved ridge, which Section 1.4.2 measures against the exact posterior. The chapter reports the Laplace values, sets the exact value beside each answer with the gap, and collects the comparison in Section 1.4.

1.2.1 The posterior from the rows

The answers use the objects of one posterior: the MAP state $\hat{\mathbf{z}}$, the covariance $\boldsymbol{\Sigma}$ and the Jacobian $\mathbf{J}$ of the rows. The Jacobian is written out in Section 1.1.7; the other two are derived here for the hall. Chapter 3 derives these objects from the probability rules for Gaussian rows over any state, and Chapter 4 dissects the same objects in full on the hall with its scenario node carried as an entry, including the Laplace approximation and its curvature term. Here each is written out once for a reader meeting it for the first time.

The posterior. Four ingredients are in hand: the unknown state $\mathbf{z} = (T, g)^{\mathsf{T}}$; the observation $y_T = 23.2°\mathrm{C}$; the physical relation $Q = G\,(T - T_c)$; and the prior information $p(g)$. What is ultimately wanted is

$$p(\mathbf{z} \mid \mathbf{y}), \tag{1.14}$$

the *posterior distribution* of the state, conditional on the evidence $\mathbf{y}$, here the reading y_T and the handbook value, and on the assumed physical model. It states which values of the state remain plausible after the physics, the measurement and the prior information have been combined. Chapter 3 derives Bayes' theorem, by which the posterior is proportional to the likelihood times the prior, $p(\mathbf{z} \mid \mathbf{y}) \propto p(y_T \mid \mathbf{z})\, p(\mathbf{z})$; here the reading row plays the part of the likelihood, and the physics row and the prior row on g together play the part of the prior $p(\mathbf{z})$.

The normal equations and the Gauss–Newton step. Near a state $\mathbf{z}$, each row is replaced by its tangent, $\mathbf{r}(\mathbf{z} + \Delta\mathbf{z}) \approx \mathbf{r}(\mathbf{z}) + \mathbf{J}\,\Delta\mathbf{z}$, and the objective by $\frac{1}{2}\,\|\mathbf{r} + \mathbf{J}\,\Delta\mathbf{z}\|^2$, a quadratic in the step $\Delta\mathbf{z}$. Its gradient with respect to the step is $\mathbf{J}^{\mathsf{T}}(\mathbf{r} + \mathbf{J}\,\Delta\mathbf{z})$, and setting it to zero gives the *normal equations*

$$(\mathbf{J}^{\mathsf{T}}\mathbf{J})\,\Delta\mathbf{z} = -\mathbf{J}^{\mathsf{T}}\mathbf{r}, \tag{1.15}$$

a linear system of two equations in the two entries of the step. The Gauss–Newton method solves it, moves to $\mathbf{z} + \Delta\mathbf{z}$, evaluates $\mathbf{r}$ and $\mathbf{J}$ there, and repeats until the step no longer changes the state at the precision wanted. The vector $\mathbf{J}^{\mathsf{T}}\mathbf{r}$ is the gradient of the objective (1.12), so the iteration stops where that gradient vanishes, at the MAP state. The iteration starts here from the design state $\mathbf{z}^0 = (23°\mathrm{C},\ 0)^{\mathsf{T}}$ of the forward calculation, at which the physics row is zero.

The information matrix. The matrix $\boldsymbol{\Lambda} = \mathbf{J}^{\mathsf{T}}\mathbf{J}$ of the normal equations is the *information matrix* of the rows. It is a sum over the rows: writing $\mathbf{J}_k$ for the k-th row of $\mathbf{J}$,

$$\mathbf{J}^{\mathsf{T}}\mathbf{J} = \sum_{k=1}^{3} \mathbf{J}_k^{\mathsf{T}}\mathbf{J}_k,$$

so that each row contributes the outer product of its own sensitivities, and the precisions of independent rows add. The information matrix of normalized rows is $\mathbf{J}^{\mathsf{T}}\mathbf{J}$; when the rows are written in their physical units with weights $w_k = 1/\sigma_k^2$, as in Section 1.1.7, the same matrix is $\mathbf{J}^{\mathsf{T}}\mathbf{W}\mathbf{J}$, with $\mathbf{W} = \mathrm{diag}(w_k)$ and $\mathbf{J}$ the Jacobian of the unnormalized rows. For the hall the three outer products are of two kinds. The reading row and the prior row each have one

nonzero entry, so each contributes one diagonal entry, $1/\sigma_T^2$ in the T–T position and $1/\sigma_g^2$ in the g–g position. The physics row has two, and its outer product $\mathbf{J}_1^\mathsf{T}\mathbf{J}_1$ fills the whole matrix, off-diagonal entries included: it is the only source of the coupling between the two entries, and it has rank one, with all its information along its own row and none along the tangent of the curve. The eigenvectors of $\mathbf{J}^\mathsf{T}\mathbf{J}$ give the directions in which the rows fix the state, and its eigenvalues how sharply: a large one across the curve, set by $1/\sigma_p^2$, and a small one along it, set by the reading and the prior, as the ridge below makes exact.

The Laplace covariance and the curvature term. The inverse of the information matrix is the covariance of the posterior: exactly so for rows linear in the state (Chapter 3), and as the *Laplace approximation*, the normal distribution fitted at the MAP state, for curved rows (Chapter 4). The Gauss–Newton matrix $\mathbf{J}^\mathsf{T}\mathbf{J}$ drops the curvature of the rows themselves. The full Hessian of the objective (1.12), its matrix of second derivatives, is

$$\frac{\partial^2 \mathcal{L}}{\partial \mathbf{z}\,\partial \mathbf{z}^\mathsf{T}} = \mathbf{J}^\mathsf{T}\mathbf{J} + \sum_{k=1}^{3} r_k \,\frac{\partial^2 r_k}{\partial \mathbf{z}\,\partial \mathbf{z}^\mathsf{T}}, \tag{1.16}$$

and of the three rows only the physics row is curved, through e^{-g}, with the single nonzero second derivative

$$\frac{\partial^2 r_1}{\partial g^2} = -\frac{(Q/G_\mathrm{ref})\,\mathrm{e}^{-g}}{\sigma_p}.$$

The *curvature term* $r_1\,\partial^2 r_1/\partial g^2$ therefore adds to the g–g entry alone. It does not vanish in the exact-physics limit, because r_1 shrinks in proportion to σ_p while $\partial^2 r_1/\partial g^2$ grows as $1/\sigma_p$: at the MAP state the gradient $\mathbf{J}^\mathsf{T}\mathbf{r}$ is zero, its T entry reads $r_1/\sigma_p + r_2/\sigma_T = 0$, so $r_1 = -(\sigma_p/\sigma_T)\,r_2$, and

$$r_1\,\frac{\partial^2 r_1}{\partial g^2} = \frac{r_2}{\sigma_T}\,\frac{Q}{G_\mathrm{ref}}\,\mathrm{e}^{-\hat{g}}$$

at every σ_p. It is the curvature of the physics curve itself, and it acts in the one direction, along the curve, in which the information is small. The *Laplace covariance* is the inverse of the full Hessian at the MAP state,

$$\mathbf{\Sigma} = \Big(\mathbf{J}^\mathsf{T}\mathbf{J} + \sum_k r_k\,\frac{\partial^2 r_k}{\partial \mathbf{z}\,\partial \mathbf{z}^\mathsf{T}}\Big)^{-1}_{\hat{\mathbf{z}}}, \qquad \mathbf{z} \mid \mathbf{y} \approx \mathcal{N}(\hat{\mathbf{z}}, \mathbf{\Sigma}); \tag{1.17}$$

a Newton iteration, which uses the full Hessian in place of $\mathbf{J}^\mathsf{T}\mathbf{J}$ in equation (1.15), reaches the same $\hat{\mathbf{z}}$.

The ridge. The spreads of the two entries are not independent of each other, and the reason is read off the physics row. Its Jacobian row, the derivative of equation (1.6) with respect to (T, g), and the vector it annihilates are

$$\mathbf{J}_1 = \frac{1}{\sigma_p}\begin{bmatrix} 1 & (Q/G_\mathrm{ref})\,\mathrm{e}^{-g} \end{bmatrix}, \qquad \mathbf{n} = \begin{bmatrix} (Q/G_\mathrm{ref})\,\mathrm{e}^{-g} \\ -1 \end{bmatrix}, \qquad \mathbf{J}_1\,\mathbf{n} = 0. \tag{1.18}$$

A step along $\mathbf{n}$ leaves the physics row unchanged to first order, so $\mathbf{n}$ is the tangent to the zero set of the row, a change of $(Q/G_\mathrm{ref})\,\mathrm{e}^{-g}$ kelvin of temperature per unit of g with the opposite sign. With $\mathbf{J}_k$ the k-th row of $\mathbf{J}$, the information is a sum over rows, $\mathbf{J}^\mathsf{T}\mathbf{J} = \sum_k \mathbf{J}_k^\mathsf{T}\mathbf{J}_k$. The information the physics row contributes, $\mathbf{J}_1^\mathsf{T}\mathbf{J}_1$, is a matrix of rank one: it grows as $1/\sigma_p^2$ across the curve and is zero along $\mathbf{n}$. Along the curve only the reading row and the prior row inform the state, and by the pattern (1.10) each acts on one entry, with the contributions $\mathbf{J}_2^\mathsf{T}\mathbf{J}_2 = \mathrm{diag}(1/\sigma_T^2,\ 0)$ and $\mathbf{J}_3^\mathsf{T}\mathbf{J}_3 = \mathrm{diag}(0,\ 1/\sigma_g^2)$; the physics row's own curvature adds the

term $\mathbf{C} = \mathrm{diag}(0,\ n_1 r_2/\sigma_T)$, derived above, with r_2 the reading row at the MAP state. With $\mathbf{u} = \mathbf{n}/\|\mathbf{n}\|$ the unit tangent, the information along the curve is

$$\lambda_{\|} = \mathbf{u}^{\mathsf{T}}(\mathbf{J}_2^{\mathsf{T}}\mathbf{J}_2 + \mathbf{J}_3^{\mathsf{T}}\mathbf{J}_3 + \mathbf{C})\mathbf{u} = \frac{n_1^2/\sigma_T^2 + 1/\sigma_g^2 + n_1 r_2/\sigma_T}{n_1^2 + 1}, \qquad n_1 = (Q/G_{\mathrm{ref}})\,\mathrm{e}^{-\hat{g}}, \tag{1.19}$$

which does not depend on σ_p. The tangent $\mathbf{u}$ is orthogonal to $\mathbf{J}_1$, so in the orthonormal basis $(\mathbf{J}_1^{\mathsf{T}}/\|\mathbf{J}_1\|,\ \mathbf{u})$ the curvature of the objective has the entry $\lambda_{\|}$ along $\mathbf{u}$, an entry across the curve that grows as $1/\sigma_p^2$, and a coupling that stays bounded; its inverse, by the Schur complement of Chapter 3, leaves $1/\lambda_{\|}$ along $\mathbf{u}$ and nothing else as $\sigma_p \to 0$. The covariance therefore tends to the rank-one matrix $\boldsymbol{\Sigma} \to \mathbf{u}\,\mathbf{u}^{\mathsf{T}}/\lambda_{\|}$: the posterior collapses onto a segment of the tangent, a *ridge* of the objective. Then $\sqrt{\Sigma_{11}} \to |[\mathbf{u}]_1|/\sqrt{\lambda_{\|}}$ and $\sqrt{\Sigma_{22}} \to |[\mathbf{u}]_2|/\sqrt{\lambda_{\|}}$ are one spread along the curve seen through the two components of the tangent, and the correlation $\Sigma_{12}/\sqrt{\Sigma_{11}\Sigma_{22}}$ tends to the sign of $[\mathbf{u}]_1[\mathbf{u}]_2$, which is -1. The two entries are not known separately: a hall slightly warmer than $\hat{T}$ is a hall with a slightly lower conductance, and the posterior allows the one only together with the other. The reading row has a zero in the column of g, and it narrows the conductance all the same, through the first component of $\mathbf{u}$.

> **Principle 1.2.** *Physical rows constrain the system; reading rows provide evidence about its unknown quantities. The posterior is where the two agree, weighted by the prior, and neither kind of row alone determines it. An exact physical row leaves the posterior only along its own zero set, and its signature in $\boldsymbol{\Sigma}$ is a correlation near -1 or $+1$.*

The exact posterior along the curve. In the limit $\sigma_p \to 0$ the posterior lies on the physics curve, and along the curve it can be computed without approximation. On the curve the temperature is a function of the conductance entry, $T(g) = T_c + (Q/G_{\mathrm{ref}})\,\mathrm{e}^{-g}$. The rows define the posterior relative to a base flat in (T, g). Integrated across the curve, the physics row contributes a factor that does not depend on g, because the row is written in kelvin and its slope in T, $1/\sigma_p$, does not depend on g; the base is therefore flat in g along the curve, and what remains is the density of g along the curve, the reading row and the prior row evaluated on it:

$$p(g \mid \mathbf{y}) \propto \exp\left[-\frac{1}{2}\left(\frac{T(g) - y_T}{\sigma_T}\right)^2 - \frac{1}{2}\left(\frac{g - \mu_g}{\sigma_g}\right)^2\right]. \tag{1.20}$$

One-dimensional quadrature of this density over g gives every probability and moment of a quantity derived from the state exactly, for T through $T(g)$ and for G through $G_{\mathrm{ref}}\,\mathrm{e}^{g}$. The chapter computes every answer from the Laplace normal and sets the exact value from this density beside it, with the gap between the two.

The answer read off. An answer that is an entry of the state is read from the two objects directly: the entry $\hat{z}_i$ of $\hat{\mathbf{z}}$, with the spread $\sqrt{\Sigma_{ii}}$. The relation between two entries is the correlation $\Sigma_{ij}/\sqrt{\Sigma_{ii}\Sigma_{jj}}$, and the directions of the posterior are the eigenvectors of $\boldsymbol{\Sigma}$. A derived quantity f is carried by its gradient, with the spread $\sqrt{\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}}$, and a threshold on an entry is a normal tail of its marginal. Section 1.2.2 computes $\hat{\mathbf{z}}$, $\mathbf{J}$ and $\boldsymbol{\Sigma}$ for the hall.

1.2.2 The posterior of the hall

The hall's inputs are the pinned values $T_c = 18°\mathrm{C}$, $Q = Q_1 = 30\,\mathrm{kW}$ and $G_{\mathrm{ref}} = 6\,\mathrm{kW/K}$; the reading $y_T = 23.2°\mathrm{C}$ with $\sigma_T = 0.5\,\mathrm{K}$; the handbook prior $\mu_g = 0.2231$ with $\sigma_g = 0.3$; and the tolerance $\sigma_p = 0.01\,\mathrm{K}$ of the physics row.

The first step. At the design state $\mathbf{z}^0 = (23^\circ\mathrm{C},\ 0)^\mathsf{T}$ the Jacobian (1.13) is

$$\mathbf{J}(\mathbf{z}^0) = \begin{bmatrix} 100 & 500 \\ 2 & 0 \\ 0 & 3.333 \end{bmatrix} :$$

a one-kelvin change in T moves the physics row by 100 and the reading row by 2, and a unit change in g moves the physics row by 500 and the prior row by 3.33. The first row is the pair of kilowatt sensitivities $(G,\ Q) = (6,\ 30)$ divided by $G = 6\,\mathrm{kW/K}$ and by σ_p. The residual vector there is $\mathbf{r} = (0,\ -0.400,\ -0.744)$: the physics holds, the temperature lies 0.4 standard deviations below the reading, and g lies 0.74 standard deviations below the handbook value. The two sides of the normal equations (1.15) are

$$\mathbf{J}(\mathbf{z}^0)^\mathsf{T}\mathbf{J}(\mathbf{z}^0) = \begin{bmatrix} 10004 & 50000 \\ 50000 & 250011 \end{bmatrix}, \qquad -\mathbf{J}(\mathbf{z}^0)^\mathsf{T}\mathbf{r}(\mathbf{z}^0) = \begin{bmatrix} 0.800 \\ 2.479 \end{bmatrix},$$

and the first step is their solution,

$$\Delta\mathbf{z} = \begin{bmatrix} 0.0684 \\ -0.0137 \end{bmatrix}.$$

The iterations. Table 1.4 lists the iterations. Each row is one iterate $\mathbf{z}^k$; the columns give its two entries, its three rows and the objective of equation (1.12).

Table 1.4: Gauss–Newton iterations for the posterior of the equipment hall (equations (1.11) to (1.15)), with $\sigma_p = 0.01\,\mathrm{K}$, from the design state $(23^\circ\mathrm{C},\ 0)$.

k	T (°C)	g [–]	r_1 [–]	r_2 [–]	r_3 [–]	$\mathcal{L}$ [–]
0	23.0000	0.0000	0.000	−0.400	−0.744	0.3566
1	23.0684	−0.0137	−0.042	−0.263	−0.789	0.3471
2	23.0701	−0.0139	0.005	−0.260	−0.790	0.3460
3	23.0701	−0.0139	0.005	−0.260	−0.790	0.3460

The first step does almost all the work, because over a step this short the rows are nearly linear. It leaves the physics row slightly violated, $r_1 = -0.042$, through the curvature of e^{-g} that the tangent ignores; the second step restores it, and the third changes nothing at the printed precision. The MAP state is

$$\hat{\mathbf{z}} = \begin{bmatrix} \hat{T} \\ \hat{g} \end{bmatrix} = \begin{bmatrix} 23.070^\circ\mathrm{C} \\ -0.0139 \end{bmatrix}, \qquad \hat{G} = 6\,\mathrm{e}^{-0.0139} = 5.917\,\mathrm{kW/K}. \tag{1.21}$$

There the physics holds to $r_1\,\sigma_p = 5 \times 10^{-5}\,\mathrm{K}$, and the reading and the prior share the disagreement: the temperature lies 0.26 standard deviations below the reading and g lies 0.79 standard deviations below the handbook value. The second entry is the commissioned value of Section 1.1.5, $\hat{g}_0 = \hat{g} = -0.0139$.

The information matrix. At the MAP state the Jacobian (1.13) is

$$\mathbf{J}(\hat{\mathbf{z}}) = \begin{bmatrix} 100 & 507.0 \\ 2 & 0 \\ 0 & 3.333 \end{bmatrix}, \qquad \frac{(Q/G_{\mathrm{ref}})\,\mathrm{e}^{-\hat{g}}}{\sigma_p} = \frac{5\,\mathrm{e}^{0.0139}}{0.01} = 507.0,$$

and the information matrix is the sum of the three outer products,

$$\mathbf{J}^{\mathsf{T}}\mathbf{J} = \underbrace{\begin{bmatrix} 10000 & 50701 \\ 50701 & 257056 \end{bmatrix}}_{\text{physics}} + \underbrace{\begin{bmatrix} 4 & 0 \\ 0 & 0 \end{bmatrix}}_{\text{reading}} + \underbrace{\begin{bmatrix} 0 & 0 \\ 0 & 11.11 \end{bmatrix}}_{\text{prior}} = \begin{bmatrix} 10004 & 50701 \\ 50701 & 257067 \end{bmatrix}.$$

The physics term has rank one: all its information lies along its own row $(100,\ 507.0)$, across the physics curve, and none along the curve; only the reading's 4 and the prior's 11.11 inform along it. The eigenvalues of $\mathbf{J}^{\mathsf{T}}\mathbf{J}$ are 267067 across the curve and 4.27 along it, with the eigenvector $(0.9811,\ -0.1935)^{\mathsf{T}}$ of the small one pointing along the curve; the first two terms of equation (1.19) give it directly, $(5.070^2 \times 4 + 11.11)/(5.070^2 + 1) = 4.27$. The Gauss–Newton covariance is the inverse,

$$\mathbf{\Sigma}_{\text{GN}} = (\mathbf{J}^{\mathsf{T}}\mathbf{J})^{-1} = \begin{bmatrix} 0.2256 & -0.0445 \\ -0.0445 & 0.0088 \end{bmatrix},$$

with the standard deviations $\sqrt{0.2256} = 0.475\,\text{K}$ and $\sqrt{0.0088} = 0.094$.

The curvature term and the Laplace covariance. The second derivative of the physics row at the MAP state is $\partial^2 r_1/\partial g^2 = -507.0$; multiplied by $r_1 = 0.0052$, it adds -2.63 to the g–g entry, $257067 \to 257064$, the same value as $(r_2/\sigma_T)(Q/G_{\text{ref}})\,\mathrm{e}^{-\hat{g}} = (-0.2598/0.5) \times 5.0701$. Projected on the along-curve eigenvector $(0.9811,\ -0.1935)^{\mathsf{T}}$, the term is $-2.63 \times 0.1935^2 = -0.099$ against the along-curve information 4.27, a loss of 2.3 % in the one direction in which the information is small; it raises the variances by 2.4 % and the standard deviations by 1.2 %. The Laplace covariance (1.17) is

$$\mathbf{\Sigma} = \begin{bmatrix} 0.23096 & -0.045552 \\ -0.045552 & 0.0089881 \end{bmatrix}, \qquad \sqrt{\Sigma_{11}} = 0.4806\,\text{K}, \quad \sqrt{\Sigma_{22}} = 0.09481. \tag{1.22}$$

These are the Laplace values used in the rest of the chapter, quoted as 0.481 K and 0.0948; Gauss–Newton alone gives 0.475 K and 0.094. The correlation of the two entries, computed from the unrounded $\mathbf{\Sigma}$, is

$$\frac{\Sigma_{12}}{\sqrt{\Sigma_{11}\,\Sigma_{22}}} = \frac{-0.045552}{0.48058 \times 0.094806} = -0.9998.$$

The ridge. The physics row of $\mathbf{J}(\hat{\mathbf{z}})$, $(100,\ 507.0)$, is annihilated by $\mathbf{n} = (5.070,\ -1)^{\mathsf{T}}$, equation (1.18) at $\hat{g}$, and the long axis of $\mathbf{\Sigma}$ is $\mathbf{u} = (0.9811,\ -0.1935)^{\mathsf{T}}$, the same direction normalized. By equation (1.19),

$$\lambda_{\parallel} = \frac{5.070^2 \times 4 + 11.11 + 5.070 \times (-0.260)/0.5}{5.070^2 + 1} = 4.168,$$

$$\frac{0.9811}{\sqrt{4.168}} = 0.481\,\text{K}, \quad \frac{0.1935}{\sqrt{4.168}} = 0.0948,$$

the two diagonal entries of equation (1.22).

Figure 1.5 shows the geometry. The horizontal axis is the log-conductance g and the vertical axis the hall temperature T, the two entries of the state, so that every point of the plane is a candidate state. The green curve is the zero set of the physics row: the states the physics allows. The gray band is the set of temperatures within one standard deviation of the reading, $23.2 \pm 0.5°\text{C}$, and the dashed line inside it the reading itself. Points are labeled by their plane coordinates (g, T), horizontal first. The three black dots are the design state,

$(g, T) = (0, 23)$, a path built exactly to design; the state on the curve that matches the reading exactly, $(g, T) = (-0.039, 23.2)$; and, between them, the MAP state of equation (1.21), the fit of the commissioning test, $(g, T) = (-0.014, 23.07)$. The purple segment is the one-standard-deviation ellipse of the covariance of equation (1.22), centered on the MAP state; it is drawn as a thick line because its half-width across the curve, about 0.002 in g, is below the resolution of the figure.

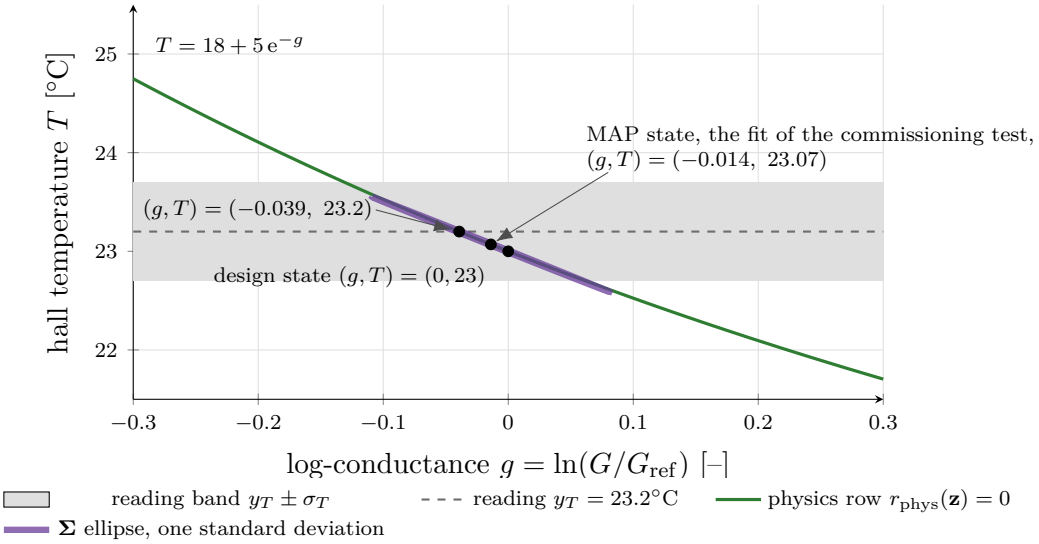


Figure 1.5: The state plane of the equipment hall with $T_c = 18$°C, $Q = 30$ kW and $G_{\text{ref}} = 6$ kW/K. Green: the zero set of the physics row, $T = 18 + 5\,e^{-g}$ (equation (1.5)). Gray: the reading $y_T = 23.2$°C with its band of one standard deviation, $\sigma_T = 0.5$ K (equation (1.7)). The plausible states lie on the arc of the curve inside the band. Purple: the one-standard-deviation ellipse of the Laplace covariance (equation (1.22)) around the MAP state, the fit of the commissioning test, $(g, T) = (-0.014,\ 23.07)$, collapsed onto the curve by the physics row. Points are labeled (g, T), horizontal coordinate first; the design state is a path built exactly to design.

The plausible states lie on the arc of the curve inside the band, from $g = -0.131$ to $g = 0.062$ ($G = 5.26$ to 6.38 kW/K, a health index h from 0.877 to 1.064), and the MAP state sits on it between the reading-match point and the design state, pulled toward the handbook value $\mu_g = 0.2231$ by the prior row.

The pair $(\hat{\mathbf{z}}, \mathbf{\Sigma})$ of equations (1.21) and (1.22) is what step 4 uses: every answer of the chapter is computed from it. Because the physics row is exact in the limit, the posterior also has the exact form (1.20) along the curve, and each answer has an exact value beside it, with the gap between the two reported.

1.3 The answers (step 4)

The first four questions of Table 1.2 take two general forms, written in the book's notation: $p(\mathbf{z} \mid \mathbf{y})$ is the posterior density of the state, $f(\mathbf{z})$ a quantity derived from the state and f^* a threshold on it. The fifth, a design question about a row not yet attached, is derived with the next measurement in Section 1.6.1. The questions name the limit of 24°C at the present load and the limit of 27.5°C at $Q_2 = 60$ kW.

1.3.1 The hall temperature and the health index

Asset manager's questions. What temperature is the hall running at now? How far is the as-commissioned conductance from its design value? The first asks for the temperature of the hall at the commissioning-test load, given its one reading and the handbook value; the second for the conductance of its air path at commissioning read against the design sheet, the health index of equation (1.3), and for the probability that the path was commissioned below design.

General form. An estimation question about a quantity $f(\mathbf{z})$ derived from the state asks for its posterior distribution, the probability that the quantity lies at or below each value f^* (the value at which it equals one half is the median),

$$\Pr(f(\mathbf{z}) \le f^* \mid \mathbf{y}) = \int \mathbf{1}\big[f(\mathbf{z}) \le f^*\big]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}, \tag{1.23}$$

where $\mathbf{1}[\cdot]$ is one when the event holds and zero otherwise, summarized by a location and a spread,

$$\mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}] = \int f(\mathbf{z})\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}, \qquad \mathrm{Var}(f(\mathbf{z}) \mid \mathbf{y}) = \mathbb{E}\big[(f(\mathbf{z}) - \mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}])^2 \mid \mathbf{y}\big]. \tag{1.24}$$

An entry of the state is the case $f(\mathbf{z}) = z_i$.

Specialized form. The two questions are equations (1.23) and (1.24) with the state $\mathbf{z} = (T, g)^{\mathsf{T}}$ of the hall, the derived quantity $f(\mathbf{z}) = z_1 = T$ for the temperature; $f(\mathbf{z}) = z_2 = g$, $f(\mathbf{z}) = G = G_{\mathrm{ref}}\, \mathrm{e}^{z_2}$ or $f(\mathbf{z}) = h = \mathrm{e}^{z_2}$ for the conductance and its health index, with $f^* = 1$ on h for the event below design; and $p(\mathbf{z} \mid \mathbf{y})$ the posterior of step 3, the Laplace normal $\mathbf{z} \mid \mathbf{y} \approx \mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ of equations (1.21) and (1.22). Each answer is read from $\hat{\mathbf{z}}$ and $\boldsymbol{\Sigma}$; the exact density (1.20) gives the value set beside it.

The temperature. Here $f(\mathbf{z}) = z_1 = T$, the first entry of the state. Under the Laplace posterior the marginal of an entry of a normal vector is normal, with the entry's mean and the diagonal entry of the covariance, so

$$T \mid \mathbf{y} \approx \mathcal{N}(\hat{z}_1, \Sigma_{11}),$$

and the answer is read off without any transformation: the first entry of $\hat{\mathbf{z}}$, with the spread $\sqrt{\Sigma_{11}}$.

The conductance and the health index. The second entry $z_2 = g$ is read off in the same way, $g \mid \mathbf{y} \approx \mathcal{N}(\hat{z}_2, \Sigma_{22})$. The conductance itself, $G = G_{\mathrm{ref}}\, \mathrm{e}^{z_2}$, is a derived quantity of the state. Linearized at the MAP state, its variance is $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$ with the gradient

$$\mathbf{g} = \frac{\partial G}{\partial \mathbf{z}} = \begin{bmatrix} \partial G/\partial T \\ \partial G/\partial g \end{bmatrix}_{\hat{\mathbf{z}}} = \begin{bmatrix} 0 \\ G_{\mathrm{ref}}\, \mathrm{e}^{\hat{g}} \end{bmatrix} = \begin{bmatrix} 0 \\ \hat{G} \end{bmatrix}, \qquad \sigma_G^2 = \mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g} = \hat{G}^2\, \Sigma_{22}, \tag{1.25}$$

the delta-method spread, written σ_G. The ratio $\sigma_G/\hat{G} = \sqrt{\Sigma_{22}}$: the spread of the log-conductance is the fractional spread of the conductance, as Section 1.1.4 anticipated. The health index $h = \mathrm{e}^{z_2}$ differs from G only by the constant G_{ref}, so its gradient is $\mathbf{g} = (0,\ \mathrm{e}^{\hat{g}})^{\mathsf{T}} = (0,\ \hat{h})^{\mathsf{T}}$ and its band is $\sigma_h = \hat{h}\sqrt{\Sigma_{22}} = \sigma_G/G_{\mathrm{ref}}$. The event below design, $h < 1$, is the event $g < 0$ on the second entry, since h increases with g. Equation (1.23) with $f = h$ and $f^* = 1$ is then the cumulative of the second entry's marginal at zero,

$$\Pr(h < 1 \mid \mathbf{y}) = \Pr(z_2 < 0 \mid \mathbf{y}) \approx \Phi\!\left(\frac{0 - \hat{z}_2}{\sqrt{\Sigma_{22}}}\right), \tag{1.26}$$

where Φ is the standard normal cumulative distribution function.

The value beside each comes from the exact posterior: the moments (1.24) under the density (1.20), evaluated by one-dimensional quadrature over g, with T through $T(g) = T_c + (Q/G_{\mathrm{ref}})\,\mathrm{e}^{-g}$ and G and h through $G_{\mathrm{ref}}\,\mathrm{e}^{g}$ and e^{g}, and the median of each, the value f^* at which equation (1.23) equals one half; the exact probability below design is the mass of equation (1.20) below $g = 0$. The Laplace location, the entry of $\hat{\mathbf{z}}$, is the median of the Laplace normal, and it is set beside the exact median, like with like; the exact mode of g is $\hat{g}$ itself, and the exact mean is given as well, labeled as a mean. They differ by the skew of the posterior along the curve.

Building blocks. The MAP state $\hat{\mathbf{z}} = (23.070\,^{\circ}\mathrm{C},\ -0.0139)^{\mathsf{T}}$ and the Laplace covariance $\boldsymbol{\Sigma}$, curvature term included, with $\sqrt{\Sigma_{11}} = 0.4806\,\mathrm{K}$ and $\sqrt{\Sigma_{22}} = 0.09481$ (equations (1.21) and (1.22)); the gradient (1.25) and its counterpart $(0,\ \hat{h})^{\mathsf{T}}$ for the health index; the normal cumulative Φ of equation (1.26); and, for the exact values beside them, the density (1.20) and a one-dimensional quadrature over g.

The hall temperature now. The hall temperature is the first entry of the state: $\hat{T} = 23.070\,^{\circ}\mathrm{C}$ with $\sqrt{\Sigma_{11}} = 0.481\,\mathrm{K}$ (Gauss–Newton alone, $0.475\,\mathrm{K}$). The value beside it comes from the exact posterior: the median of T along the curve is $23.028\,^{\circ}\mathrm{C}$, $0.042\,\mathrm{K}$ below $\hat{T}$ (the mean is $23.030\,^{\circ}\mathrm{C}$), with the standard deviation $0.4813\,\mathrm{K}$, $0.0007\,\mathrm{K}$ wider.

The commissioned conductance and its health index. The conductance entry is $\hat{g} = -0.0139$ with $\sqrt{\Sigma_{22}} = 0.0948$. By equation (1.25), with $\mathbf{g} = (0,\ 5.917)^{\mathsf{T}}$,

$$\hat{G} = 5.917\,\mathrm{kW/K}, \qquad \sigma_G = \sqrt{\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\,\mathbf{g}} = 5.917\,\sqrt{0.0089881} = 0.561\,\mathrm{kW/K},$$

a band of 9.5 % of its value, against a factor $\mathrm{e}^{0.3} = 1.35$ in the handbook before the reading. The values beside them come from the exact posterior: the conductance entry has the median -0.0056, 0.0083 above $\hat{g}$ (the mean is -0.0013), and the standard deviation 0.0967, 0.0019 wider; the conductance has the median $5.967\,\mathrm{kW/K}$, $0.049\,\mathrm{kW/K}$ above $\hat{G}$ (the mean is $6.020\,\mathrm{kW/K}$), and the standard deviation $0.591\,\mathrm{kW/K}$, $0.030\,\mathrm{kW/K}$ or 5 % wider. The exact posterior extends further toward large conductances than the normal at $\hat{\mathbf{z}}$ does, the skew that the full-load question below meets in its tail.

Read against the design sheet, the commissioned value $\hat{g}_0 = \hat{g} = -0.0139$, that is $\hat{G} = 5.92\,\mathrm{kW/K}$, is the health index

$$\hat{h} = \mathrm{e}^{\hat{g}} = \mathrm{e}^{-0.0139} = 0.986, \qquad \sigma_h = 0.986 \times 0.0948 = 0.0935,$$

1.4 % below design, with a band of 9.5 % of its value; exactly, its median is 0.994 and its mean 1.003, with the standard deviation 0.099. By equation (1.26), the probability that the path was commissioned below design is

$$\Pr(G < G_{\mathrm{ref}} \mid \mathbf{y}) = \Pr(h < 1 \mid \mathbf{y}) \approx \Phi\!\left(\frac{0 + 0.0139}{0.0948}\right) = \Phi(0.147) = 0.558.$$

The value beside it comes from the exact posterior: the mass of equation (1.20) below $g = 0$ is 0.523, 0.035 lower, because the exact posterior leans toward large conductances. On both routes the probability is close to one half: the commissioning test does not say whether the path met its design.

1.3.2 The limits at the present load and at 60 kilowatts

Asset manager's questions. Is the hall within its 24 °C limit now? Can the path carry 60 kW within 27.5 °C? The first asks about the hall as it runs at the commissioning-test load. The second is the owner's request, which for the asset manager reads

What is the probability that the air path, as commissioned, holds the hall below 27.5°C *when the load increases to* 60 kW*?*

Its unknown is the temperature T_2 of the scenario node at the full load of 60 kW, and over the state of the graph it is

$$\Pr(T_2 < 27.5^\circ\mathrm{C} \mid \mathbf{y}), \tag{1.27}$$

with the unknown, the full-load temperature, on the left of the bar and the evidence, the commissioning-test reading and the handbook value, on the right; the graph and its pinned inputs are conditions of the model throughout.

General form. A prediction under a threshold is a tail probability of a derived quantity under the posterior, the posterior mass of the states that satisfy the event,

$$\Pr(f(\mathbf{z}) < f^* \mid \mathbf{y}) = \int \mathbf{1}\big[f(\mathbf{z}) < f^*\big]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}. \tag{1.28}$$

Specialized form. The two questions are equation (1.28) with the state $\mathbf{z} = (T, g)^\mathsf{T}$ of the hall, $p(\mathbf{z} \mid \mathbf{y})$ the posterior of step 3, the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ of equations (1.21) and (1.22), and the derived quantity $f(\mathbf{z})$ and the threshold f^* given for each below: $f(\mathbf{z}) = T$ with $f^* = 24^\circ\mathrm{C}$, and $f(\mathbf{z}) = T_2$ with $f^* = T_{\max} = 27.5^\circ\mathrm{C}$. Two choices enter here. The full-load condition is taken to share the conductance of the commissioning-test condition, with the load and the supply air pinned at both: this is what makes a condition the hall has never run predictable from a reading at 30 kW, and it fails if the conductance changes with the load, for instance through an airflow that rises with it (Section 1.1.4), or if the unit saturates and can no longer hold T_c at 60 kW (Section 1.6.5); no check of this chapter can test either, because nothing has been observed at 60 kW. The full-load answer is the steady temperature at 60 kW: this is what lets the event be written on the conductance alone, and it says nothing about the period just after the load rises, while the hall's stored heat still matters (Section 1.6.5).

The present load. Here $f(\mathbf{z}) = z_1 = T$ and $f^* = 24^\circ\mathrm{C}$. Under the Laplace posterior the tail is a normal tail of the first entry,

$$\Pr(T < 24^\circ\mathrm{C} \mid \mathbf{y}) \approx \Phi\!\left(\frac{24 - \hat{z}_1}{\sqrt{\Sigma_{11}}}\right), \tag{1.29}$$

with Φ the normal cumulative of equation (1.26). The value beside it comes from the exact posterior: on the curve the event $T < 24^\circ\mathrm{C}$ is the event $T(g) < 24^\circ\mathrm{C}$, that is $g > -\ln\big((24 - T_c)\, G_{\mathrm{ref}}/Q\big) = -\ln 1.2$, and its probability is the integral of equation (1.20) above that value.

The full load. Here $f(\mathbf{z}) = T_2$, the temperature of the scenario node. By equation (1.9) it is a function of the second entry alone, $T_2 = T_c + (Q_2/G_{\mathrm{ref}})\, \mathrm{e}^{-z_2}$, and on the zero set of the physics row it is also a function of the first: dividing equation (1.9) by equation (1.8) removes the conductance,

$$\frac{T_2 - T_c}{T - T_c} = \frac{Q_2}{Q_1} \qquad \Longrightarrow \qquad T_2 = T_c + \frac{Q_2}{Q_1}\,(T - T_c). \tag{1.30}$$

At twice the load the excess over the supply air is twice as large, whatever the conductance: a commissioning-test temperature of 23°C predicts $T_2 = 18 + 2 \times 5 = 28^\circ\mathrm{C}$, and one of 23.2°C predicts 28.4°C. Measurements taken during the first operating condition therefore inform the second, although the hall has never been operated at full load; this is how a model of the asset links the evidence of its commissioning test to predictions under operating conditions that have not yet occurred. In the vector form the link is the tangent $\mathbf{n}$ of equation (1.18): a step along the curve moves T by $(Q_1/G_{\mathrm{ref}})\, \mathrm{e}^{-g}$ and T_2 by $(Q_2/G_{\mathrm{ref}})\, \mathrm{e}^{-g}$ per unit of $-g$, in the ratio Q_2/Q_1.

The event itself needs no posterior to be rewritten. The full-load temperature $T_2 = T_c + Q_2/G$ decreases in G, so for a limit $T_{\max} > T_c$ the event $T_2 < T_{\max}$ is, in turn, an event on the conductance, on the conductance entry and, by equation (1.30), on the commissioning-test temperature:

$$T_2 < T_{\max} \iff G > G^* \iff g > g^* = \ln \frac{G^*}{G_{\text{ref}}} \iff T < T^*, \tag{1.31}$$

with the thresholds

$$G^* = \frac{Q_2}{T_{\max} - T_c}, \tag{1.32}$$

$$T^* = T_c + \frac{Q_1}{Q_2}\,(T_{\max} - T_c). \tag{1.33}$$

For $T_{\max} \le T_c$ the probability is zero, since $T_2 > T_c$ for every positive conductance. The inequalities of equation (1.31) describe the same set of states on the physics curve, so their probabilities are equal under any distribution on the curve; the equivalence holds under the assumed steady-state model, with known load and supply-air temperature. The question about a condition that has never been operated has become a question about the unknown conductance, an entry of the state that the commissioning-test condition informs, before any data are processed.

Figure 1.6 draws the equivalence (1.31) on the forward relation (1.2). The horizontal axis is the conductance G in kW/K and the vertical axis the hall temperature in °C. The solid blue curve is the hall at the load of 30 kW, the dashed green curve the same hall at 60 kW; the dot on the blue curve is the forward calculation of Section 1.1.4, $T = 23$°C at $G = 6$ kW/K, a path built to design, and the dot above it the same path at 60 kW, 28°C. The red line is the limit of 27.5°C. The two dotted vertical lines mark the design value $G_{\text{ref}} = 6$ kW/K, through the two dots, and the threshold G^* of equation (1.32), where the dashed curve crosses the limit. The dashed green curve lies below the red line exactly to the right of G^*: because the curve is strictly decreasing, the event "below the limit" on the vertical axis is the event "above G^*" on the horizontal axis.

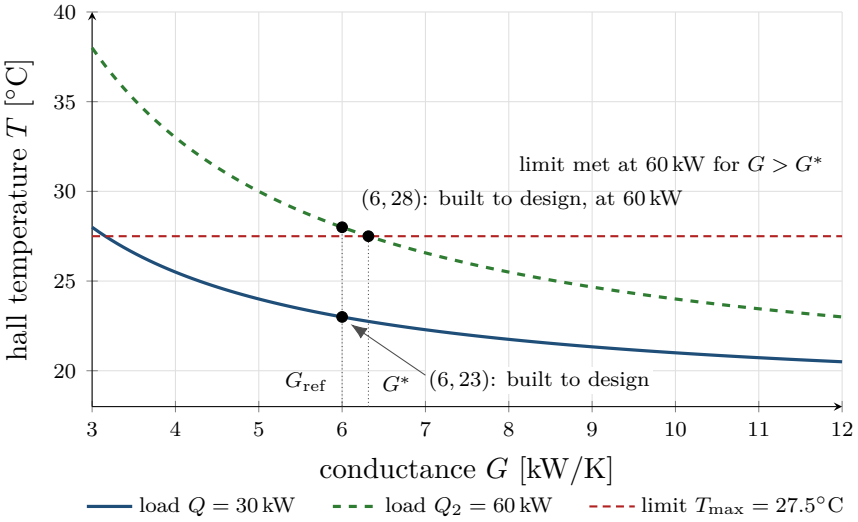


Figure 1.6: The forward relation $T = T_c + Q/G$ (equation (1.2)) with $T_c = 18$°C, for the loads $Q = 30$ kW and $Q_2 = 60$ kW. Dots: the forward calculation $T = 23$°C for a path built to design, $G = G_{\text{ref}} = 6$ kW/K; the same path at 60 kW, 28°C; and the threshold $G^* = Q_2/(T_{\max} - T_c)$ of equation (1.32) at which the 60 kW hall reaches the limit of 27.5°C. Dotted: the design value G_{ref} and the threshold G^*.

Under the Laplace posterior the probability is a normal tail of the second entry, the entry onto which the physics maps the event,

$$\Pr(T_2 < T_{\max} \mid \mathbf{y}) = \Pr(g > g^* \mid \mathbf{y}) \approx \Phi\left(\frac{\hat{z}_2 - g^*}{\sqrt{\Sigma_{22}}}\right). \tag{1.34}$$

The same event written on the first entry, $T < T^*$, has the same exact probability, but under the Laplace normal, whose ridge is the straight tangent rather than the curve, its tail differs slightly; the chapter reports the tail on g, the entry on which the full-load temperature depends alone. The value beside it comes from the exact posterior: the probability is the integral of equation (1.20) above g^*.

Principle 1.3. *When a quantity in a question is a monotone function of one entry of the state, a threshold on the quantity is a threshold on that entry, and the two events have the same probability under every posterior. The transformation uses the physics only, and it can be made before any data are processed.*

Building blocks. The entries $\hat{z}_1 = 23.070\,^\circ\mathrm{C}$ and $\hat{z}_2 = -0.0139$ of the MAP state and the standard deviations $\sqrt{\Sigma_{11}} = 0.481\,\mathrm{K}$ and $\sqrt{\Sigma_{22}} = 0.0948$ of the covariance (equations (1.21) and (1.22)); the threshold of equation (1.32) for the request's $T_{\max} = 27.5\,^\circ\mathrm{C}$ and $Q_2 = 60\,\mathrm{kW}$, and for other limits and loads, with its equivalent on the commissioning-test temperature, equation (1.33), and in the health index, $h^* = G^*/G_{\mathrm{ref}}$; the normal cumulative distribution Φ; and, for the exact values beside them, the density (1.20) integrated over g above each threshold.

The hall below 24 degrees. By equation (1.29),

$$\Pr(T < 24\,^\circ\mathrm{C} \mid \mathbf{y}) \approx \Phi\left(\frac{24 - 23.070}{0.481}\right) = \Phi(1.93) = 0.973.$$

The value beside it comes from the exact posterior: the mass of equation (1.20) above $g = -\ln 1.2 = -0.1823$ is 0.977, 0.004 higher.

The hall at 60 kilowatts. For the request's limit $T_{\max} = 27.5\,^\circ\mathrm{C}$ at $Q_2 = 60\,\mathrm{kW}$, equations (1.32) and (1.33) give

$$G^* = \frac{60}{27.5 - 18} = \frac{60}{9.5} = 6.316\,\mathrm{kW/K}, \qquad g^* = \ln\frac{6.316}{6} = 0.0513, \qquad h^* = \mathrm{e}^{g^*} = 1.053,$$

$$T^* = 18 + \tfrac{1}{2} \times 9.5 = 22.75\,^\circ\mathrm{C}:$$

the temperature at 60 kW stays below 27.5°C if and only if the temperature at 30 kW is below 22.75°C; on the physics curve, the states with $g > g^* = 0.0513$ are those below 22.75°C. In the health index, the request needs a path 5.3 % above its design value. By equation (1.34),

$$\Pr(T_2 < 27.5\,^\circ\mathrm{C} \mid \mathbf{y}) \approx \Phi\left(\frac{-0.0139 - 0.0513}{0.0948}\right) = \Phi(-0.688) = 0.246,$$

with the MAP conductance 5.917 kW/K to the left of G^* in Figure 1.6. The value beside it comes from the exact posterior: the mass of equation (1.20) above $g^* = 0.0513$ is 0.281, 0.036 higher. The event lies in the tail of g toward large conductances, the side on which the exact posterior extends further than the normal at $\hat{\mathbf{z}}$, so the state route understates it. The gap does not change the decision: on both routes the hall meets the limit at 60 kW with probability of about one in four. A path built exactly to design, $h = 1$, would itself run at 28.0°C at 60 kW by the forward calculation of Section 1.1.4, 0.50 K over the limit.

Table 1.5 evaluates the same blocks for other limits and loads. Each row is a limit $T_{\max}$ and a future load Q_2; the columns give the threshold conductance G^*, the threshold log-conductance g^*, the threshold health index h^*, the threshold on the commissioning-test temperature T^*, the probability of equation (1.34), which the chapter reports, and beside it the exact value, the mass of equation (1.20) above g^*. One row of the arithmetic, for $T_{\max} = 27°\text{C}$ and $Q_2 = 60\,\text{kW}$: $G^* = 60/9 = 6.667\,\text{kW/K}$, $g^* = \ln(6.667/6) = 0.1054$, $h^* = 6.667/6 = 1.111$, $T^* = 18 + 0.5 \times 9 = 22.5°\text{C}$, and $\Phi((-0.0139 - 0.1054)/0.0948) = \Phi(-1.258) = 0.104$.

Table 1.5: Equivalent thresholds and the full-load probability, with $T_c = 18°\text{C}$, $Q_1 = 30\,\text{kW}$ and $G_{\text{ref}} = 6\,\text{kW/K}$ (equations (1.32) and (1.33)); the probability by the Laplace tail (1.34), with $\hat{g} = -0.0139$ and $\sqrt{\Sigma_{22}} = 0.0948$, and exactly, by quadrature of equation (1.20). The threshold health index is $h^* = \mathrm{e}^{g^*}$. The row 27.5°C at 60 kW is the owner's request; the others are added to show the trend.

Limit $T_{\max}$ [°C]	Load Q_2 [kW]	G^* [kW/K]	g^* [–]	h^* [–]	T^* [°C]	Laplace [–]	Exact [–]
26	60	7.500	0.2231	1.250	22.00	0.0062	0.0155
27	60	6.667	0.1054	1.111	22.50	0.1042	0.1354
27.5	60	6.316	0.0513	1.053	22.75	0.2458	0.2815
29	60	5.455	−0.0953	0.909	23.50	0.8047	0.8355
27.5	45	4.737	−0.2364	0.789	24.33	0.9905	0.9964
27.5	75	7.895	0.2744	1.316	21.80	0.0012	0.0049

A stricter limit, or a larger future load, demands a larger conductance and a cooler commissioning-test temperature; the first six columns follow from the physics alone. Only the request at 29°C and the lighter load of 45 kW need less than the design conductance. The Laplace tail is below the exact one in every row, and the gap grows in ratio toward the far tail: a factor of 2.5 to 4 in the two smallest rows.

1.4 The checks (step 5)

The fit rests on two approximations and one assumption: a physics row with a finite tolerance in place of an exact law, a normal distribution in place of the posterior along a curved ridge, and a reading and a handbook value that are compatible with each other. Each is set against a full calculation, and the weight of the handbook value is measured by removing it.

1.4.1 The exact-physics limit

Table 1.6 repeats the computation at three model tolerances. Each row is one σ_p; the columns are the two entries of $\hat{\mathbf{z}}$, their Laplace standard deviations and their correlation. From 0.1 to 0.01 K the conductance entry moves by 0.001, its spread by 0.002 and the correlation from -0.979 to -0.9998; from 0.01 to 0.001 K nothing moves at the printed precision but the last digit of the correlation. The limit is reached at $\sigma_p = 0.01\,\text{K}$, and the correlation tends to -1 as equation (1.19) requires.

1.4.2 The Laplace posterior against the exact posterior

In the limit $\sigma_p \to 0$ the exact posterior is the density (1.20) along the physics curve, and one-dimensional quadrature of it gives every quantity of the questions; the event $T < 24°\text{C}$ is the event $g > -\ln 1.2 = -0.1823$ on the curve, the event below design is $g < 0$, and the full-load event is $g > g^* = 0.0513$. Table 1.7 collects the pairs that step 4 computed, the Laplace value it reports beside the exact one. Each row is one quantity of the posterior, with its unit; the

Table 1.6: The posterior of the equipment hall at three model tolerances σ_p of the physics row: MAP state, Laplace standard deviations (with the curvature term) and correlation. The exact-physics limit $\sigma_p \to 0$ is reached at 0.01 K.

σ_p [K]	$\hat{T}$ [°C]	$\hat{g}$ [–]	$\sqrt{\Sigma_{11}}$ [K]	$\sqrt{\Sigma_{22}}$ [–]	correlation [–]
0.1	23.0705	−0.0130	0.4806	0.0966	−0.9790
0.01	23.0701	−0.0139	0.4806	0.0948	−0.9998
0.001	23.0701	−0.0139	0.4806	0.0948	−1.0000

third column is its Laplace value, read from $\hat{\mathbf{z}}$ and $\boldsymbol{\Sigma}$ of equations (1.21) and (1.22), and the fourth its exact value from the quadrature. The location of each quantity is its median in both columns, the Laplace median being the value at $\hat{\mathbf{z}}$; the exact mean has a row of its own.

Table 1.7: The Laplace approximation of the posterior of the equipment hall against the exact posterior along the physics curve, in the limit $\sigma_p \to 0$.

Quantity	Unit	Laplace, at $\hat{\mathbf{z}}$	Exact, along the curve
Median of T	°C	23.070	23.028
Mean of T	°C	–	23.030
Standard deviation of T	K	0.4806	0.4813
$\Pr(T < 24°\text{C} \mid \mathbf{y})$	–	0.973	0.977
Median of g	–	−0.0139	−0.0056
Mean of g	–	–	−0.0013
Standard deviation of g	–	0.0948	0.0967
Median of G	kW/K	5.917	5.967
Mean of G	kW/K	–	6.020
Standard deviation of G	kW/K	0.561	0.591
Median of h	–	0.986	0.994
Mean of h	–	–	1.003
Standard deviation of h	–	0.0935	0.0986
$\Pr(h < 1 \mid \mathbf{y})$, below design	–	0.558	0.523
$\Pr(T_2 < 27.5°\text{C} \mid \mathbf{y})$ at 60 kW	–	0.246	0.281

For the temperature the approximation is adequate: the tail probability differs by 0.004 and the standard deviation by 0.0007 K. The posterior is slightly skewed along the curve, so its mode, median and mean differ: the mode of g is $\hat{g}$, the median lies 0.0083 and the mean 0.013 above it, the median conductance lies 0.049 kW/K and the mean 0.10 kW/K above $\hat{G}$, and the exact standard deviation of G is 0.591 kW/K, 5 % wider than the Laplace 0.561 kW/K. The same skew lowers the probability that the path was commissioned below design, 0.523 exactly against the Laplace 0.558, without moving it away from one half. The skew matters most for the full-load question, whose event lies in the tail of g toward large conductances, the side on which the exact posterior extends further: the Laplace tail 0.246 is low by 0.036 against the exact 0.281. Further into the tail the understatement grows: in Table 1.5 the Laplace tail on g is smaller than the exact one by a factor of 2.5 to 4 in the two smallest rows, 0.0062 against 0.0155 and 0.0012 against 0.0049. Written on the first entry instead, as $T < T^* = 22.75°\text{C}$, the full-load event has the Laplace tail 0.253, because the normal follows the straight tangent rather than the curve; the exact value is 0.281 on either entry. The exact values rest on the base of equation (1.20), flat in g along the curve; with a base flat in T along the curve instead, the full-load value is 0.249, a reminder that an exact tail depends on the base measure as well as on the rows. The conductance band and the full-load probability are therefore slightly narrow and

slightly low, and the report says so.

A scalar check. Eliminating T through the physics row and differentiating the remaining one-entry objective in g twice gives the same $\sqrt{\Sigma_{22}} = 0.0948$ and, through the slope of the curve, the same $\sqrt{\Sigma_{11}} = 0.481\,\mathrm{K}$. The vector form reaches them without the substitution, and it is the form that carries over unchanged to the cooling network of Section 1.1.3.

1.4.3 The residual check

Three rows act on two entries, and with the physics exact the reading and the handbook both inform the one direction along the curve: one row is redundant, and the rows at the MAP state measure how well the two agree. They are $r_1 = 0.005$, $r_2 = -0.260$ and $r_3 = -0.790$ (Table 1.4), and their squares sum to $2\mathcal{L}(\hat{\mathbf{z}}) = 0.692$. When the model and the standard deviations of the rows are right, that sum follows a chi-square distribution with as many degrees of freedom as there are rows less entries, here one, the chi-square test of bad-data detection that a later chapter derives. A sum of 0.692 or more then has probability 0.41, while a sum above 3.84 would have probability below 0.05 and would say that the reading or the handbook value is wrong. The two agree, which is a first, coarse look at the question of a drifting sensor.

1.4.4 The weight of the handbook prior

The handbook value is the one input of the fit that the hall did not supply, and removing its row shows how much the answers lean on it. Without the prior row the residual vector holds the rows r_1 and r_2 of equation (1.11): two rows, two entries, a square Jacobian, both rows zero at the MAP state ($23.2^\circ\mathrm{C},\ -0.0392$), and $\Sigma_{11} = \sigma_T^2$ whatever σ_p, and $\sqrt{\Sigma_{22}} = 0.0962$ (the last exercise of the chapter derives both). On the conductance entry, the route of equation (1.34), the full-load probability is then

$$\Pr(T_2 < 27.5^\circ\mathrm{C} \mid y_T) \approx \Phi\!\left(\frac{-0.0392 - 0.0513}{0.0962}\right) = \Phi(-0.941) = 0.173,$$

against 0.246 with the handbook row, a rise of 0.073. Exactly, with the same base flat in g along the curve and the prior term removed from equation (1.20), it is 0.212 against 0.281, a rise of 0.069. Figure 1.7 draws the two Laplace marginals of the conductance entry. The horizontal axis is g and the vertical axis the probability density; the dashed curve is the marginal under the reading alone, $\mathcal{N}(-0.0392,\ 0.0962^2)$, and the solid curve the marginal with the handbook row, $\mathcal{N}(-0.0139,\ 0.0948^2)$. The dotted red line is the threshold $g^* = 0.0513$, and the shaded areas to its right are the two probabilities. The handbook value, which places the conductance above what the reading alone implies, moves the marginal toward the threshold with nearly the same width.

The same holds in every row of Table 1.5, on both routes: without the handbook row each probability is lower than with it. The full-load answer therefore depends on the handbook value, and the handbook value is stated with the report. The probability that the path was commissioned below design leans on it too: without the handbook row it is $\Phi((0 + 0.0392)/0.0962) = \Phi(0.408) = 0.658$ by the Laplace route (0.618 exactly), against 0.558 with it.

1.5 The report (step 6)

The dashboard's "3.8 % below design" is one commissioning-test reading taken as exact. Held against the handbook value through the energy balance, the reading gives a commissioned health index of 0.986 ± 0.0935 against the design sheet and a probability of 0.56 (0.52 exactly) that the path was commissioned below design: the test does not show whether the path met

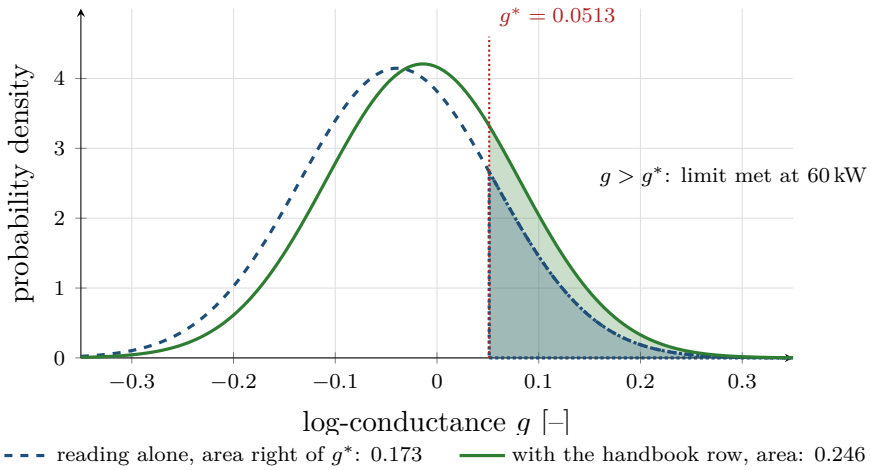


Figure 1.7: The weight of the handbook prior in the full-load answer: Laplace marginals of the conductance entry, $\mathcal{N}(-0.0392,\ 0.0962^2)$ under the reading alone (dashed) and $\mathcal{N}(-0.0139,\ 0.0948^2)$ with the handbook row (solid), densities $\varphi((g-\hat{g})/s)/s$ with φ the standard normal density. The shaded areas right of $g^* = 0.0513$ are the full-load probabilities, 0.173 and 0.246.

its design. Nor would a path at design carry the request. The limit of 27.5°C at 60 kW needs 1.053 times the design conductance, and the path as commissioned meets it with probability of about one in four, 0.246 by the Laplace route and 0.281 exactly: neither the comfortable pass of the handbook tile nor the certain failure of the reading taken alone. A duty test of the path, the performance test that seems to measure the very heat in question, would not change that answer (Section 1.6.2). The answers to the first four questions of Table 1.2 are given below in the units in which they were asked; the decision on the owner's request follows from the fourth, and the entry in the maintenance log records them.

1.5.1 The answers and their bands

Table 1.8 gives the answers. Each row is one question; the columns give the answer computed from $\hat{\mathbf{z}}$ and $\boldsymbol{\Sigma}$, which the chapter reports, the exact value beside it from the density along the curve, and the band (one standard deviation for the temperature and the conductance, Laplace and exact, with the band of the health index and the probability below design; for the full-load probability, its value without the handbook row, the check of Section 1.4.4).

The reading has placed the hall temperature to within about half a kelvin, slightly better than the sensor's own 0.5 K because the handbook value adds a little through the physics row. The conductance is known to within 9.5 % of its value, against a factor $e^{0.3} = 1.35$ in the handbook before the reading; the exact posterior widens that band to 0.59 kW/K. The hall is below 24°C with probability of about 0.97. The fifth question, whether another sensor would narrow the health index, is answered in Section 1.6.

1.5.2 The full-load decision

The owner's request has the answer of the fourth row: at 60 kW the hall stays below 27.5°C with probability of about one in four, 0.246 by the Laplace route and 0.281 exactly. The air path as commissioned does not support the doubled load. The answer assumes that the unit still holds the supply air at 18°C at 60 kW; if it cannot, T_c rises with the load and the hall is warmer than this answer says (Section 1.6.5). The reason is read from the equivalent threshold: the limit is met only if the conductance exceeds $G^* = 6.316$ kW/K, a health index of 1.053,

Table 1.8: The report on the equipment hall from the commissioning-test reading $y_T = 23.2°\text{C}$ and the handbook conductance of 7.5 kW/K, with the physics exact; the health index is read against the design conductance of 6 kW/K.

Question	Answer (Laplace route)	Exact	Band
What temperature is the hall running at now?	$\hat{T} = 23.07°\text{C}$	median 23.03°C	±0.48 K (0.481 K exact)
How far is the as-commissioned conductance from its design value?	$\hat{G} = 5.92$ kW/K, commissioned index 0.986	median 5.97 kW/K (0.994); mean 6.02 kW/K	±0.56 kW/K (±0.0935 in the index); Pr(below design) = 0.56 (0.52 exact)
Is the hall within its 24°C limit now?	$\Pr(T < 24°\text{C} \mid \mathbf{y}) =$ 0.973, about 0.97	0.977; gap 0.004	–
Can the path carry 60 kW within 27.5°C?	$\Pr(T_2 < 27.5°\text{C} \mid \mathbf{y}) =$ 0.246, about 0.25	0.281; gap 0.036	without the handbook value 0.173 (0.212 exact)

and the posterior places the conductance at 5.92 ± 0.56 kW/K; in the conductance entry the threshold $g^* = 0.0513$ lies 0.69 standard deviations above $\hat{g} = -0.0139$. The answer leans on the handbook value: without it the probability would be 0.173 by the Laplace route and 0.212 exactly. A measurement cannot raise the conductance; it can only show where it lies, and with the estimate below the threshold, further evidence is more likely to confirm the shortfall than to remove it. The decision is therefore an upgrade of the path beyond its design, not a restoration: a path built exactly to design would run at 28.0°C at 60 kW, 0.50 K over the limit, so bringing the path up to its design sheet, the remedy the dashboard's reading suggests, would not carry the request. How to upgrade the path, or replace it, is the repair-or-replace decision of Chapter 17, in Volume II.

1.5.3 The entry in the maintenance log

The commissioning test enters the path's maintenance log as its baseline, window 0, with the values of Table 1.8:

- commissioning baseline, window 0: commissioned value $\hat{g}_0 = -0.0139$, that is $\hat{G} = 5.92 \pm 0.56$ kW/K and a health index of 0.986 ± 0.0935 against the design sheet, below design with probability 0.56 (0.52 exactly); no shortfall against design is demonstrated, and the baseline is recorded for the condition tests of later windows;
- request for 60 kW within 27.5°C: met with probability 0.246 (0.281 exactly), and the limit needs 1.053 times the design conductance; referred as an upgrade beyond design;
- next measurement: an independent second hall sensor (Section 1.6.3).

The entry uses nothing later than window 0: no rate, no remaining life and no date can be read from one window, and the baseline waits for the condition tests that follow it.

1.6 The next measurement (step 7)

Whether the bands of Table 1.8 are narrow enough is the asset manager's decision. If they are not, the fifth question of Table 1.2 asks which measurement would narrow them most, and a measurement not yet taken is a row not yet attached.

1.6.1 The preposterior covariance

The value a candidate measurement will return is unknown, but its Jacobian row at $\hat{\mathbf{z}}$, its sensitivity to each entry divided by its standard deviation, is known before the measurement is made, and that is enough. The information matrix is a sum of one outer product per row (Section 1.2.1), so a candidate row with the Jacobian row $\mathbf{J}_a$ adds $\mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a$ to the inverse of the present covariance, and the covariance expected after the measurement is

$$\boldsymbol{\Sigma}_a = \left(\boldsymbol{\Sigma}^{-1} + \mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a\right)^{-1}, \tag{1.35}$$

with $\boldsymbol{\Sigma}$ of equation (1.22), curvature term included. The value the measurement will return does not appear in it. Chapter 2 derives $\boldsymbol{\Sigma}_a$ as the preposterior covariance, and Chapter 23, in Volume II, on sensor selection for prognostics, ranks candidate rows by it for quantities derived from the state. The band on the temperature after the measurement is $\sqrt{\Sigma_{a,11}}$, the band on the conductance, the quantity on which the full-load answer rests, is $\sigma_{G,a} = \sqrt{\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}_a\mathbf{g}}$ with the gradient of equation (1.25), and the band on the health index is $\sigma_{h,a} = \sigma_{G,a}/G_{\text{ref}}$.

1.6.2 The candidate measurements

Four candidates are considered, each written as a row over $\mathbf{z} = (T, g)^{\mathsf{T}}$:

- a second hall sensor of the same accuracy, with an error independent of the first: the row $(T - y'_T)/\sigma_T$, with the Jacobian row $\mathbf{J}_a = (2,\ 0)$, the same as the present reading row;
- a reference sensor of 0.2 K in the hall: $\mathbf{J}_a = (5,\ 0)$;
- a reading of the supply-air temperature, or a load meter on the servers: T_c and Q are pinned, not entries of the state, so the row has no entry in either column, $\mathbf{J}_a = (0,\ 0)$;
- a duty (heat-flow) test of the path, the asset manager's performance test, with a standard deviation $\sigma_{\text{duty}} = 1\,\text{kW}$ on the heat removed $G\,(T - T_c)$: $\mathbf{J}_a = (\hat{G},\ \hat{G}\,(\hat{T} - T_c))/\sigma_{\text{duty}} = (5.917,\ 30.00)$.

Table 1.9 evaluates equation (1.35) for each. Each row is one candidate; the columns give its Jacobian row, the standard deviation of T after it, $\sqrt{\Sigma_{a,11}}$, and the width it removes from the present 0.481 K, then the conductance band after it, $\sigma_{G,a}$ with the gradient $\mathbf{g} = (0,\ 5.917)^{\mathsf{T}}$, and the width it removes from the present 0.561 kW/K, and last the band on the health index after it, $\sigma_{h,a}$, against the present 0.0935. One row of the arithmetic, for the second sensor: $\mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a$ has the single entry 4, in its T–T position, and

$$\boldsymbol{\Sigma}_a = \begin{bmatrix} 0.12005 & -0.023678 \\ -0.023678 & 0.0046738 \end{bmatrix}, \qquad \sqrt{0.12005} = 0.346\,\text{K}, \quad 5.917\,\sqrt{0.0046738} = 0.405\,\text{kW/K},$$

which is $\sigma_{h,a} = 0.0674$ in the health index.

Only the two sensors in the hall remove anything, and the reason is visible in their Jacobian rows. The posterior is wide in one direction only, along the physics curve, $(0.9811,\ -0.1935)^{\mathsf{T}}$, where the present information is 4.17: the 4.27 of $\mathbf{J}^{\mathsf{T}}\mathbf{J}$ less the 0.099 of the curvature term. The second sensor's row projects on that direction as $2 \times 0.9811 = 1.962$ and adds $1.962^2 = 3.85$, nearly doubling the information, so that the variances along the curve fall by the factor $4.17/8.02 = 0.52$ and both standard deviations by the factor 0.72. The supply-air reading and the load meter read pinned quantities: in this model they are known already, their rows are zero, and they would inform only a graph in which T_c or Q were free. The duty test is the subtler case. Its row $(5.917,\ 30.00)$ is proportional to the physics row $(100,\ 507.0)$, both in the ratio 5.070, so it points across the curve, where the physics already fixes the state to within σ_p; it measures again what the energy balance asserts, and it adds nothing along the curve. A measurement is worth what it adds in the direction in which the posterior is wide.

Table 1.9: Candidate measurements for the equipment hall ranked by the preposterior covariance of equation (1.35), evaluated at $\hat{\mathbf{z}}$. Present bands: 0.481 K on T, 0.561 kW/K on G and 0.0935 on the health index h.

Candidate	$\mathbf{J}_a$ [K^{-1}, –]	$\sqrt{\Sigma_{a,11}}$ [K]	removed [K]	$\sigma_{G,a}$ [kW/K]	removed [kW/K]	$\sigma_{h,a}$ [–]
Second hall sensor, 0.5 K	(2, 0)	0.346	0.134	0.405	0.156	0.0674
Reference hall sensor, 0.2 K	(5, 0)	0.185	0.296	0.216	0.345	0.0360
Supply-air reading or load meter	(0, 0)	0.481	0	0.561	0	0.0935
Duty (heat-flow) test of the path, 1 kW	(5.917, 30.00)	0.481	$< 10^{-6}$	0.561	$< 10^{-6}$	0.0935

1.6.3 The recommendation

The recommendation, and the answer to the fifth question, is a second sensor in the hall: it removes 0.134 K of the 0.481 K on the temperature and 0.156 kW/K of the 0.561 kW/K on the conductance, 28 % of each, and narrows the health index from 0.0935 to 0.0674; a 0.2 K reference sensor, if one is available, removes 62 %, to 0.0360 in the index. A reading of the supply air, a load meter and a duty test of the path are not worth installing for these answers. Two conditions apply. The second sensor's error must be independent of the first: a second reading from the same sensor at the same condition would share any bias of the first, and a shared error is a discrepancy that repeats (Section 1.1.4), which a repeated reading does not reduce. And the gain is computed at $\hat{\mathbf{z}}$ under the Laplace approximation, which the checks found adequate for T and slightly narrow for G. Once the measurement is taken, its row joins the three of equation (1.11), and the method returns to step 3 with four rows. For the request the narrower conductance band sharpens the answer but does not move the threshold: the decision of Section 1.5.2 turns on an upgrade of the path, not on the sensor.

1.6.4 What the analysis showed

The dashboard, taking the commissioning test's reading as exact, reported the path 3.8 % below design and the doubled load over its limit at 28.4°C; its handbook tile put the path at 1.25 times design and the doubled load at a comfortable 26°C. The analysis, which holds the two against each other through the energy balance, finds neither: a commissioned health index of 0.986 with a band of 9.5 %, a probability close to one half that the path was commissioned below design at all, and a probability of about one in four that the path holds the hall below 27.5°C at 60 kW, 0.246 from the posterior of the state (0.281 exact). It also overturns the obvious instrument: a duty test of the path, although it measures the very heat the question is about, removes nothing, because it repeats what the energy balance already asserts. One reading and a handbook value pin the hall temperature to about half a degree but the conductance of its air path only to about ten percent, and the two are not known separately: the posterior lives on the physics curve, a ridge along which a warmer hall is a hall with a weaker path, which is why the correlation of the two entries is -0.9998. The owner's request is a question about that conductance. The hall meets 27.5°C at 60 kW only if its path conducts more than 6.316 kW/K, 1.053 times its design value, so a path restored to its design sheet would still fall short, and the answer leans on the handbook value. What would narrow the answers is a second sensor in the hall, which removes 28 % of both bands. What would change the answer to the request is an upgrade of the path beyond its design, not more measurement.

1.6.5 Limits of the model

The posterior is exact only under the model, the steady-state graph of one hall with one path to a supply air that the discharge controller holds, and that model can be wrong before any inference is run. The supply air is pinned only while the controller holds it: a cooling unit driven to full output can no longer hold T_c, which then rises with the load and would have to become a free quantity of the graph, including at the 60 kW of the scenario node, which the full-load answer assumes the unit can still hold at 18°C. The hall and its one reading are constructed for the example: the design sheet, the 23.2°C reading and the handbook value are illustrative values with no stated truth, so no answer of this chapter is compared with a true state. The condition record of the two-coil hall from Part II on is generated from a stated truth, against which those chapters check their answers.

Modeling mistakes. Several errors can be made before any inference algorithm is run, in the construction of the graph and its rows. Table 1.10 lists five. Each row names the mistake, gives its form in the equipment hall, and states the correction.

Table 1.10: Five modeling mistakes made before any inference is run.

Mistake	In the equipment hall	Correction
Treating a measurement as the true state	the dashboard takes the commissioning test's 23.2°C as exact and reports the path 3.8 % below design	the reading is evidence with error σ_T; T stays a free entry of the state
Confusing parameters and observations	the conductance and the reading are treated alike	G is a physical property, an entry of the state; y_T is evidence about T, connected to G through the physics row
Assuming the physical model is complete	the steady-state relation is used just after the path loses half its conductance	stored heat matters in a transient, $C_{\mathrm{th}}\,\mathrm{d}T/\mathrm{d}\vartheta = Q - G\,(T - T_c)$
Confusing structural with parameter uncertainty	doubt whether the hall has one or two effective air paths is folded into the spread of G	the value of G is a parameter question within one model; the number of paths is a question between two models
Believing the graph alone determines the posterior	a correct graph is expected to give a narrow answer	useful inference also needs suitable evidence, assumptions and constraints; some entries may remain poorly identified

The first two mistakes confuse the roles of Table 1.3. A sensor reporting 23.2°C does not establish that the true temperature is exactly 23.2°C: the observation has an uncertainty, and the physics row of Section 1.1.4 showed that 23.2°C and 23°C imply conductances of 5.77 and 6 kW/K, the difference of 3.8 % that the dashboard reports as a shortfall against design. The conductance is a physical property of the path; the reading is evidence about the temperature, which is connected to the conductance through the model. The fourth mistake mixes two kinds of uncertainty that require different treatments. Uncertainty about the numerical value of G within the model of Figure 1.2 is represented by the posterior of g. Uncertainty about whether the hall has one or two effective air paths is uncertainty about the model M itself, and it is represented by comparing the evidence of the two drawings. The fifth mistake is the converse of the first. A graph provides relations; the posterior also requires evidence, and some entries may remain poorly identified even when the model is perfectly specified. In the two-path drawing, for example, a reading of the hall temperature determines only the total conductance of the

two paths and leaves their split to the prior: the Jacobian of the physics and reading rows has a null vector along the split, and only a prior row adds information in that direction.

The steady-state assumption. The third mistake concerns the physics itself. The model of equation (1.1) assumes steady conditions and a single effective conductance. During transient operation the hall stores and releases heat, and the energy balance gains a storage term,

$$C_{\mathrm{th}}\,\frac{\mathrm{d}T}{\mathrm{d}\vartheta} = Q - G\,(T - T_c),$$

where C_{th} is the thermal capacitance of the hall, its air, structure and servers, in kJ/K, and ϑ is the elapsed time in seconds, within one window. For constant Q, G and T_c, write the departure from the steady temperature as $\Delta T = T - T_c - Q/G$. Then $\mathrm{d}\Delta T/\mathrm{d}\vartheta = \mathrm{d}T/\mathrm{d}\vartheta$ and the balance becomes

$$C_{\mathrm{th}}\,\frac{\mathrm{d}\Delta T}{\mathrm{d}\vartheta} = -G\,\Delta T,$$

whose solution is $\Delta T(\vartheta) = \Delta T(0)\,\mathrm{e}^{-\vartheta/\vartheta_{\mathrm{th}}}$ with the time constant $\vartheta_{\mathrm{th}} = C_{\mathrm{th}}/G$. Therefore

$$T(\vartheta) = T_c + \frac{Q}{G} + \Big(T(0) - T_c - \frac{Q}{G}\Big)\,\mathrm{e}^{-\vartheta/\vartheta_{\mathrm{th}}}, \qquad \vartheta_{\mathrm{th}} = \frac{C_{\mathrm{th}}}{G}. \tag{1.36}$$

The steady-state relation of equation (1.2) is the limit $\vartheta \to \infty$. Suppose (illustratively) that at $\vartheta = 0$ the path loses half its conductance at once, as when one of the cooling unit's two fans trips, so that the conductance falls from 6 to 3 kW/K while the hall is at 23°C under 30 kW. The new steady temperature is $18 + 30/3 = 28$°C, and equation (1.36) gives $T(\vartheta) = 28 - 5\,\mathrm{e}^{-\vartheta/\vartheta_{\mathrm{th}}}$: after one time constant the hall is at $28 - 5\,\mathrm{e}^{-1} = 26.16$°C. A steady-state model would place it at 28°C at once. It is therefore inadequate for predicting temperatures immediately after such a loss, where the time constant decides whether the operators have minutes or seconds; the graph for the asset manager's sixth question, the sudden loss of half the conductance, needs the storage term and C_{th} as a further entry or a pinned input.

The chapter has built the representation on which every question will be asked. The next chapter, Chapter 2, turns to the questions themselves: how an asset manager's question stated in the maintenance office's language is recognized as estimation, prediction, comparison, diagnosis or design, and written as a precise probability or expectation over the state of such a graph. The commissioning test then becomes window 0 of a trend in Chapter 4, and in Part II the hall grows to two coils fed from chilled water, each with its own health index read against its commissioning fit (Chapter 5).

1.7 What to retain

Definition 1.1 (Plant graph)**.** A plant graph represents a physical system as free and pinned nodes, paths and sources, with the physical relations that must hold among them written as residual rows; its free quantities, physical states and parameters alike, form the state vector $\mathbf{z}$; readings, records and prior knowledge are attached as further rows; and every asset manager's question is then a probability or an expectation over $\mathbf{z}$, conditional on the evidence $\mathbf{y}$, computed from the posterior $p(\mathbf{z} \mid \mathbf{y})$ that the rows define.

The forward calculation, the posterior and the question are three different things computed over the same graph. The forward calculation uses no evidence and gives a number; the posterior combines the physics, the readings and the prior; the question picks what is read from the posterior. The posterior is computed in state-vector form: the rows, each divided by its standard deviation, form one residual vector over $\mathbf{z}$; the normal equations $(\mathbf{J}^{\mathsf{T}}\mathbf{J})\,\Delta\mathbf{z} = -\mathbf{J}^{\mathsf{T}}\mathbf{r}$ give the Gauss–Newton steps to the MAP state $\hat{\mathbf{z}}$; the information matrix $\mathbf{J}^{\mathsf{T}}\mathbf{J}$ is a sum of one outer

product per row; the covariance $\boldsymbol{\Sigma}$ is the inverse of $\mathbf{J}^{\mathsf{T}}\mathbf{J}$ plus the rows' own curvature, here the curvature of the physics curve; and an answer is an entry of $\hat{\mathbf{z}}$ with its $\sqrt{\Sigma_{ii}}$, or a derived quantity with $\sqrt{\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\,\mathbf{g}}$. An exact physical law is the limit $\sigma_p \to 0$ of its row, and its signature in $\boldsymbol{\Sigma}$ is a correlation near -1: a ridge along the zero set of the row, in the direction of the null vector of the row's Jacobian. A question about a quantity that is a monotone function of one entry of the state is transformed into a threshold on that entry by the physics alone, as the full-load question became $G > 6.316\,\mathrm{kW/K}$, and its probability is a normal tail of that entry's marginal. The answer is computed from $\hat{\mathbf{z}}$ and $\boldsymbol{\Sigma}$, with the exact value along the curve set beside it and the gap stated, checked against the full calculation (the limit $\sigma_p \to 0$, the exact posterior along the curve, the agreement of the redundant rows, the weight of the prior), reported with its band in the units of the question, and followed by the next measurement, ranked by the preposterior covariance $\boldsymbol{\Sigma}_a = (\boldsymbol{\Sigma}^{-1} + \mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a)^{-1}$ of a row not yet taken; a reading of a pinned quantity, or a row parallel to a physics row, removes nothing. A health index is a parameter of the physics read against a reference, here $h = G/G_{\mathrm{ref}} = \mathrm{e}^{g}$ against the design sheet; the fit of the commissioning test gives the commissioned value $\hat{g}_0$, the baseline of every later window, and it is its band, not its most probable value, that says whether the asset met its design. Table 1.11 lists the chapter's terms with their expressions, units and the subsection that defines each.

Table 1.11: The terms of the plant graph and where each is defined. Units for the equipment hall.

Term	Expression	Units	Meaning	Defined in
Plant graph	nodes, paths, sources, rows	–	the system before any question	1.1.4
Free, pinned	T, g free; T_c, Q pinned	–	inferred against known	1.1.4
Energy balance	$Q = G\,(T - T_c)$	kW	conservation at the hall node	1.1.4
Forward calculation	$T = T_c + Q/G$	°C	temperature from known inputs	1.1.4
State vector	$\mathbf{z} = (T, g)^{\mathsf{T}}$	°C, –	all free quantities at once	1.1.4
Log-conductance	$g = \ln(G/G_{\mathrm{ref}})$	–	keeps G positive	1.1.4
Health index	$h = G/G_{\mathrm{ref}} = \mathrm{e}^{g}$	–	conductance against its design value; the commissioning test's fit of g is the commissioned value $\hat{g}_0$	1.1.4
Physics row	$(T - T_c - (Q/G_{\mathrm{ref}})\mathrm{e}^{-g})/\sigma_p$	–	a relation that must hold, tolerance σ_p	1.1.4
Model discrepancy	$r_{\mathrm{phys}}(\mathbf{z}) = \Delta_{\mathrm{m}}$	kW	model error, not sensor error	1.1.4
Reading row	$(T - y_T)/\sigma_T$	–	evidence about the state	1.1.7
Prior row	$(g - \mu_g)/\sigma_g$	–	knowledge before the readings	1.1.7
Jacobian pattern	which row acts on which entry	–	a property of the graph	1.1.7
Scenario node	$T_2 = T_c + (Q_2/G_{\mathrm{ref}})\mathrm{e}^{-g}$	°C	an unobserved condition sharing g	1.1.7
Factor graph	entries and rows	–	the graph of the inference	1.1.3
Delta-method band	$\sigma_G^2 = \mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$, $\mathbf{g} = (0, \hat{G})^{\mathsf{T}}$	kW/K	spread of a derived quantity	1.3.1
Ridge	$\mathbf{J}_1\mathbf{n} = 0$, $\boldsymbol{\Sigma} \to \mathbf{u}\mathbf{u}^{\mathsf{T}}/\lambda_{\parallel}$, $\lambda_{\parallel} = (n_1^2/\sigma_T^2 + 1/\sigma_g^2 + n_1 r_2/\sigma_T)/(n_1^2 + 1)$	–	posterior along the physics curve	1.2.1
Exact posterior on the curve	$p(g \mid \mathbf{y})$, equation (1.20)	–	the limit $\sigma_p \to 0$ by quadrature	1.2.1
Threshold probability	$\Pr(T_2 < T_{\max} \mid \mathbf{y})$	–	prediction under a threshold	1.3.2
Equivalent threshold	$G^{*} = Q_2/(T_{\max} - T_c)$	kW/K	the same event on an entry	1.3.2
Posterior	$p(\mathbf{z} \mid \mathbf{y})$	–	plausible states after the evidence	1.2.1
Residual vector	$\mathbf{r}(\mathbf{z})$, $\mathcal{L} = \frac{1}{2}\mathbf{r}^{\mathsf{T}}\mathbf{r}$	–	all rows in standard deviations	1.1.7
Jacobian	$\mathbf{J} = \partial\mathbf{r}/\partial\mathbf{z}$	varies	sensitivity of each row to each entry	1.1.7
Normal equations	$(\mathbf{J}^{\mathsf{T}}\mathbf{J})\,\Delta\mathbf{z} = -\mathbf{J}^{\mathsf{T}}\mathbf{r}$	varies	the Gauss–Newton step to $\hat{\mathbf{z}}$	1.2.1
MAP state	$\hat{\mathbf{z}} = \arg\min_{\mathbf{z}} \mathcal{L}(\mathbf{z})$	°C, –	the most probable state	1.2.1
Information matrix	$\boldsymbol{\Lambda} = \mathbf{J}^{\mathsf{T}}\mathbf{J} = \sum_k \mathbf{J}_k^{\mathsf{T}}\mathbf{J}_k$	varies	precisions of the rows, added	1.2.1
Laplace covariance	$\boldsymbol{\Sigma} = \left(\mathbf{J}^{\mathsf{T}}\mathbf{J} + \sum_k r_k\,\partial^2 r_k/\partial\mathbf{z}\,\partial\mathbf{z}^{\mathsf{T}}\right)^{-1}$	varies	spreads and correlation of the entries	1.2.1
Preposterior covariance	$\boldsymbol{\Sigma}_a = (\boldsymbol{\Sigma}^{-1} + \mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a)^{-1}$	varies	spreads after a row not yet taken	1.6.1
Time constant	$\vartheta_{\mathrm{th}} = C_{\mathrm{th}}/G$	s	transient omitted by steady state	1.6.5

1.8 Exercises

Exercise 1.1. (a) Which statement best describes the role of the hall sensor, the condition-monitoring sensor of the path: it defines the physical law; it reports the exact true state; or it supplies evidence about an uncertain state? (The third: the sensor provides noisy evidence, while the physical model relates the state to the inputs and the conductance.) (b) A hall loses heat to the same supply air through two cooling units' air paths, with conductances G_1 and G_2. Name the free and pinned quantities, write the physics row, and say what a temperature reading alone can determine. (Free: T, g_1, g_2; pinned: T_c, Q; the physics row, normalized as in equation (1.6), is $(T - T_c - Q/(G_{\text{ref}}(\mathrm{e}^{g_1} + \mathrm{e}^{g_2})))/\sigma_p$; on its zero set its Jacobian row over $\mathbf{z} = (T, g_1, g_2)^\mathsf{T}$ is $(1,\ (T - T_c)G_1/(G_1 + G_2),\ (T - T_c)G_2/(G_1 + G_2))/\sigma_p$ and the reading row's is $(1/\sigma_T,\ 0,\ 0)$, so the two-by-three Jacobian of the two rows has the null vector $(0,\ G_2,\ -G_1)^\mathsf{T}$, the step $\Delta G_1 = -\Delta G_2$ that leaves the total unchanged; $\mathbf{J}^\mathsf{T}\mathbf{J}$ is singular along it, the reading determines only $G_1 + G_2$, and the split rests on the priors.) (c) Write "will the hall stay below 27°C if the load rises to 45 kW?" as a probability, and say what the graph needs. ($\Pr(T_3 < 27°\text{C} \mid \mathbf{y})$, with a scenario node $T_3 = T_c + (45/G_{\text{ref}})\,\mathrm{e}^{-g}$ at 45 kW sharing g, unobserved.)

Exercise 1.2. (a) A room is cooled by supply air at 20°C; it has a heat load of 24 kW and a conductance of 4 kW/K. Compute its steady-state temperature. ($T = 20 + 24/4 = 26°$C.) (b) A room is cooled by supply air at 19°C; it has a load of 40 kW and a strict limit of 27°C. What conductance keeps it below the limit? ($19 + 40/G < 27$ gives $40/G < 8$, so $G > 5$ kW/K.) (c) For the hall of the chapter, with $G_{\text{ref}} = 6$ kW/K, evaluate r_{phys} and r_{obs} at $\mathbf{z} = (23.2°\text{C},\ 0)$ and at $\mathbf{z} = (23°\text{C},\ 0)$. (1.2 kW and 0; 0 and -0.4: the first state matches the reading and violates the physics, the second satisfies the physics and lies 0.4 standard deviations from the reading.)

Exercise 1.3. Suppose both the supply-air temperature T_c and the conductance G are uncertain. Starting from $T = T_c + Q/G$, write the probability that the temperature exceeds a limit $T_{\max}$, and state what joint distribution is needed to evaluate it. ($\Pr(T > T_{\max} \mid \mathbf{y}) = \iint \mathbf{1}\{T_c + Q/G > T_{\max}\}\, p(T_c, G \mid \mathbf{y})\, \mathrm{d}T_c\, \mathrm{d}G$, which needs the joint posterior of T_c and G; if the two are independent given $\mathbf{y}$, it reduces to $\int p(T_c \mid \mathbf{y})\, \Pr\big(G < Q/(T_{\max} - T_c) \mid \mathbf{y}\big)\, \mathrm{d}T_c$ over $T_c < T_{\max}$, plus $\Pr(T_c \geq T_{\max} \mid \mathbf{y})$. A later chapter evaluates this integral.)

Exercise 1.4. For the two-condition hall of Figure 1.3, with a proposed load of $Q_2 = 50$ kW and a limit of 28°C: (a) derive the threshold conductance and the equivalent threshold on the commissioning-test temperature. ($T_2 = 18 + 50/G < 28$ gives $G > 5$ kW/K; by equation (1.33), $T < 18 + (30/50) \times 10 = 24°$C.) (b) Show that the commissioning-test temperature determines the full-load temperature without knowledge of G. (On the physics curve both temperatures are functions of g, $T = T_c + (Q_1/G_{\text{ref}})\,\mathrm{e}^{-g}$ and $T_2 = T_c + (Q_2/G_{\text{ref}})\,\mathrm{e}^{-g}$, so a step along the null vector of the physics row moves T_2 by $Q_2/Q_1 = 5/3$ times the move of T; in closed form, dividing the two balances gives $T_2 - T_c = (Q_2/Q_1)(T - T_c)$, equation (1.30).) (c) If the reading $y_T = 23.2°$C with $\sigma_T = 0.5$ K were the only information about T, what would the probability of meeting the limit be? ($\Phi((24 - 23.2)/0.5) = \Phi(1.6) = 0.945$.)

Exercise 1.5. The hall is at 23°C under 30 kW with $T_c = 18°$C when its conductance falls at once from 6 to 3 kW/K; its thermal capacitance is $C_{\text{th}} = 3600$ kJ/K (illustrative). (a) Compute the time constant and the new steady temperature. ($\vartheta_{\text{th}} = 3600/3 = 1200$ s $= 20$ min; 28°C.) (b) Compute the temperature after 10 and 60 minutes, and the time at which it reaches 27.5°C. ($28 - 5\,\mathrm{e}^{-0.5} = 24.97°$C; $28 - 5\,\mathrm{e}^{-3} = 27.75°$C; $5\,\mathrm{e}^{-\vartheta/\vartheta_{\text{th}}} = 0.5$ gives $\vartheta = \vartheta_{\text{th}} \ln 10 = 46.1$ min.) (c) State what a steady-state model would predict and when it is adequate. (It predicts 28°C immediately; it is adequate for conditions held for several time constants, not for the first minutes after the loss.)

Exercise 1.6. Remove the prior row from the hall of Section 1.2.1, so that the residual vector holds the rows r_1 and r_2 of equation (1.11) over $\mathbf{z} = (T, g)^{\mathsf{T}}$, with $\sigma_p = 0.01\,\mathrm{K}$ and $\sigma_T = 0.5\,\mathrm{K}$. (a) Find the MAP state and write the two-by-two Jacobian there. (Both rows vanish at $T = y_T = 23.2°\mathrm{C}$ and $g = \ln(5/5.2) = -0.0392$; there $\mathbf{J} = \left[\begin{smallmatrix} 100 & 520 \\ 2 & 0 \end{smallmatrix}\right]$, since $5\,\mathrm{e}^{0.0392}/0.01 = 520$.) (b) Show that $\mathbf{\Sigma} = (\mathbf{J}^{\mathsf{T}}\mathbf{J})^{-1}$ has $\Sigma_{11} = \sigma_T^2$ for every σ_p, and compute $\sqrt{\Sigma_{22}}$. (The Jacobian is square and invertible, so $\mathbf{\Sigma} = \mathbf{J}^{-1}\mathbf{J}^{-\mathsf{T}}$. Writing $\mathbf{J} = \left[\begin{smallmatrix} a & b \\ c & 0 \end{smallmatrix}\right]$, $\mathbf{J}^{-1} = \left[\begin{smallmatrix} 0 & 1/c \\ 1/b & -a/(bc) \end{smallmatrix}\right]$; its first row gives $\Sigma_{11} = 1/c^2 = \sigma_T^2 = 0.25\,\mathrm{K}^2$, and its second $\Sigma_{22} = (\sigma_p^2 + \sigma_T^2)/(5\,\mathrm{e}^{-\hat{g}})^2$, so $\sqrt{\Sigma_{22}} = \sqrt{0.2501}/5.2 = 0.0962$.) (c) Explain why no curvature term is needed here, and how the result gives the reading-alone probability 0.173 of Section 1.4.4. (Both rows are zero at $\hat{\mathbf{z}}$, so $\sum_k r_k\,\partial^2 r_k/\partial\mathbf{z}\,\partial\mathbf{z}^{\mathsf{T}}$ vanishes; the entry g then has the Laplace marginal $\mathcal{N}(-0.0392,\ 0.0962^2)$, and its tail above $g^* = 0.0513$ is $\Phi((-0.0392 - 0.0513)/0.0962) = \Phi(-0.941) = 0.173$.) (d) Without the handbook row, what is the probability that the path was commissioned below design? (By equation (1.26) on the same marginal, $\Pr(h < 1 \mid y_T) \approx \Phi(0.0392/0.0962) = \Phi(0.408) = 0.658$.)

Chapter 2

The Language of Asset-Management Questions

2.1 Asset-management questions and their mathematical form

The equipment hall of Chapter 1 was commissioned against a design sheet that gives its air path a conductance of 6 kW/K. Its commissioning test, window 0, read 23.2°C at 30 kW; the condition test of window 24, December of year 2, read 23.5°C at the same load. The asset dashboard reads the approach $T - T_c$, the hall air above the 18°C supply air, as the health index of the air path: it rose from 5.2 to 5.5 K, up 5.77 %, and the tile says that the air path has lost 5.8 % in two years. The straight line through the two approaches gains 0.150 K a year and reaches 6 K, where the hall can no longer hold 24°C at 30 kW, at year 5.33, window 64: 3.33 years after the later test, about three years left.

The asset manager has more questions than the tile has numbers. How far has the conductance really fallen against its design sheet? Can the hall take the requested 60 kW and stay below 27.5°C? Is the rise between the two tests fouling of the air path, or the hall sensor reading high? When is the limit reached, and with what band? Is another condition test before the next shutdown worth scheduling, and should a filter be changed now? The dashboard answers each with the same approach and the same line, and an inference algorithm can return a mathematically correct result to a question nobody asked; the step that prevents both is the translation of the asset manager's question into the mathematical quantity that answers it. The question of this chapter is

how an asset manager's question about an asset is translated into a mathematical expression that probabilistic inference can evaluate, and how that expression determines what must be computed.

The answer is short. Once a probability statement names its event, its evidence and its model, every question about the asset is a probability or an expectation over the state of its plant graph, with the unknown on the left of the conditioning bar and the evidence on the right. Asset-management questions then fall into five categories, estimation, prediction under a threshold, hypothesis comparison, diagnosis and design, each with its general forms, integrals or ratios written without reference to any particular asset. Five computations supply every quantity those forms contain: the posterior of the state; the trend, which reads that posterior along time; the evidence of a model; the preposterior of a measurement not yet taken; and the expected consequence of an action. Two of them, the trend and the expected consequence, have general forms of their own, which the categories call on. What differs from one asset to the next is the specialization: which entry or function of the state the form is applied to, which event or family of models it names, and which objects of the fit carry the answer.

Jargon: event, conditional probability, indicator function, health index, condition test, corrective work order, MTBF, posterior predictive, trend, crossing time, functional limit, evidence, Bayes factor, expected consequence, risk, value of information

An *event* is a statement about the asset that is either true or false, such as "the hall temperature exceeds 28°C"; a probability assigns it a number between zero and one. A *conditional probability* $\Pr(A \mid B)$ is the probability of event A when the information B is treated as known. The *indicator function* $\mathbf{1}_A(\mathbf{z})$ of an event equals one for the states at which the event is true and zero elsewhere, so that the probability of the event is the expectation of its indicator. The *health index* of the hall is the conductance of its air path read against the design sheet, $h = G/G_{\text{ref}}$, a parameter of the physics and not a score; a *condition test* repeats the commissioning test at the same inputs. A *corrective work order* is a repair raised because a unit failed or stopped meeting its duty, and the *MTBF* (mean time between failures) of a constant rate is its reciprocal. A *posterior predictive* probability is the probability of a future event that stays random even when the state is known, averaged over the posterior of the state. A *trend* is the change per unit time of a quantity the state carries, and the *crossing time* is the time at which a quantity that moves reaches a stated level; the *functional limit* is the level of the health index at which the asset stops doing its job, the F of the P–F interval. The *evidence*, or marginal likelihood, of a model is the probability of the observed evidence averaged over the prior of that model's unknowns; the *Bayes factor* is the ratio of the evidences of two models. The *expected consequence* of an action, also called its expected loss, is the probability-weighted average of the consequences that may follow it; summed over the events that may follow, it is each event's probability times its consequence, the *risk* of the action. The *value of information*, here the expected value of sample information (EVSI), is the reduction in expected loss that a measurement is expected to bring because the decision can be taken after its result is known.

Key idea

Every asset-management question is answered by a specific mathematical object computed from the same plant graph and its rows: a marginal, a tail probability, a predictive average, the posterior read along time, a ratio of evidences, a preposterior or an expected consequence minimized over actions. Five computations produce these objects, the posterior, the trend, the evidence, the preposterior and the expected consequence. The question, not the dashboard tile or the inference tool, decides which object, and so which computation, is required.

The questions are put to one asset, the air path of the equipment hall of Chapter 1, and to the work-order log of its coil design, so that the differences between them lie in the questions alone.

2.1.1 The plant graph of the equipment hall

Figure 2.1 draws the plant graph of Chapter 1, with the scenario node that the hall example of Section 2.4.1 adds to it. The upper green box is the free node of the hall at the test load, whose temperature T_1 is an entry of the state; the lower green box is the same hall at the requested full load, a scenario node whose temperature T_2 is also an entry and which has no sensor; the gray box on the right is the pinned node, the supply air at $T_c = 18°\text{C}$, held there by the discharge controller of the cooling unit; the two solid green arrows are the heat-removal path at the two loads, which carry one and the same effective thermal conductance G, held in the state as the log-conductance $g = \ln(G/G_{\text{ref}})$ with $G_{\text{ref}} = 6\,\text{kW/K}$, the design conductance on the design sheet, so that $g = 0$ is the design and $h = \mathrm{e}^g$ the health index; the white boxes are the sources, the heat loads Q_1 and Q_2; and the dashed box on the left, joined to the hall by

a dashed arrow, is the observation, the reading y_T of the test-load temperature. Each hall node carries one physical row, its steady-state energy balance, written for a node of temperature T and load Q as

$$Q = G\,(T - T_c), \tag{2.1}$$

where Q is the heat load in kW, G the conductance in kW/K, T the hall air temperature and T_c the supply-air temperature, both in °C. The supply air is pinned because the unit's controller holds it; the pin fails when the unit saturates, at a load its coil cannot carry, which the questions of this chapter do not reach. Solved for the hall temperature, it gives the relation used throughout the chapter,

$$T = T_c + \frac{Q}{G}. \tag{2.2}$$

The relation is deterministic: once G is known, T is known. What makes the hall a probabilistic problem is that G is not known, and that a reading y_T differs from T by a measurement error.

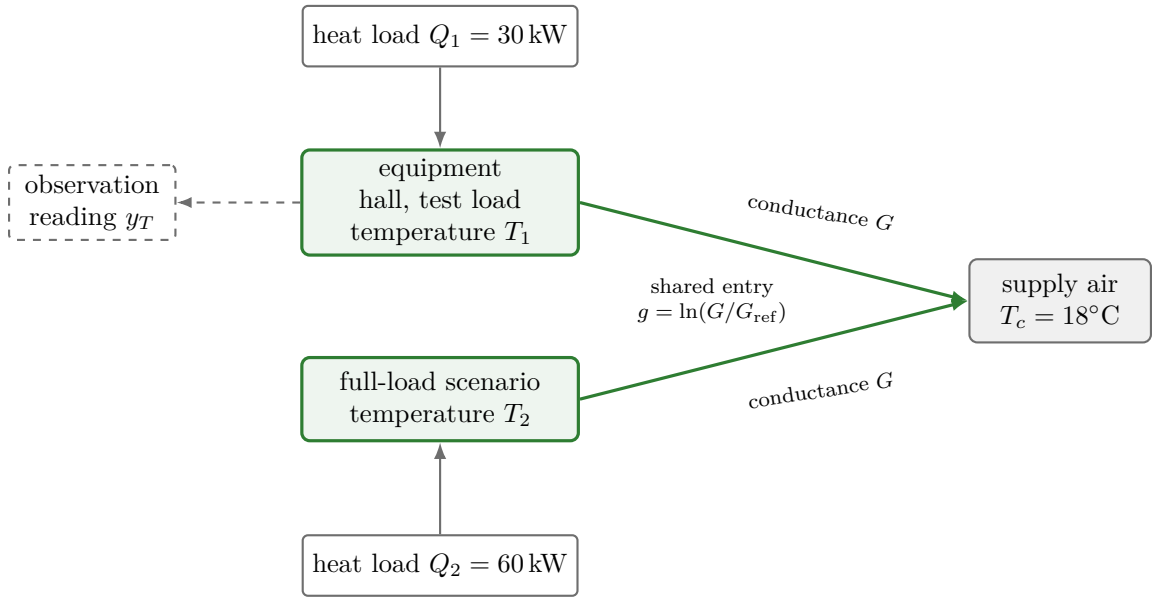


Figure 2.1: The plant graph of the equipment hall with the full-load scenario node. Free: the test-load temperature T_1 and the full-load temperature T_2 (green nodes) and the log-conductance g shared by the two heat-removal paths (solid green edges), with $G_{\mathrm{ref}} = 6\,\mathrm{kW/K}$ the design conductance. Pinned: the supply-air temperature $T_c = 18°\mathrm{C}$, held by the cooling unit's discharge controller (gray node). Sources: the heat loads $Q_1 = 30\,\mathrm{kW}$ and $Q_2 = 60\,\mathrm{kW}$. Observed: the reading y_T of the test-load temperature in a condition test (dashed box and arrow); the scenario node has no sensor.

2.1.2 Four questions about one asset

Suppose the asset manager asks the four questions of Table 2.1. Each row is a question in the asset manager's words; the second column names the kind of calculation it requires and the third column the quantity that answers it, in the notation that Section 2.2 defines.

Question A estimates an unknown physical quantity and reads it against its design value. Question B asks for the probability of an event at a load not yet applied. Question C distinguishes between two explanations of the same readings, one in which the air path has lost conductance and one in which the sensor carries a bias. Question D predicts how a measurement that does not yet exist would change a maintenance decision. All four concern the same graph, and

Table 2.1: Four questions about the air path of the equipment hall and the calculation each requires.

Asset manager's question	Calculation	Answering quantity
A. What is the air path's conductance, and where does it stand against its design sheet?	estimation of an unknown physical quantity	$p(G \mid \mathbf{y})$, and the health index $h = G/G_{\text{ref}}$
B. What is the probability that the hall carries the requested 60 kW below 27.5°C?	probability of an event	$\Pr(T_2 < 27.5^\circ\text{C} \mid \mathbf{y})$
C. Is the fall between the hall's two tests fouling, or is the hall sensor reading high?	comparison of competing explanations	$\Pr(M_{\text{path}} \mid \mathbf{y})$ against $\Pr(M_{\text{bias}} \mid \mathbf{y})$
D. Is another condition test before the next shutdown worth scheduling?	prediction of what a future measurement would change	the preposterior, and the expected loss after the test

all four are computed from the same physical model; none of them is answered by the same calculation as another.

Principle 2.1. *The same physical graph does not imply the same probabilistic calculation. The formulation of the question determines which calculation is needed.*

2.2 The grammar of probability statements

A question can be classified only when the statement that expresses it is complete: an event or an unknown quantity, the evidence it is conditioned on, and the model under which its probability is computed.

2.2.1 Events

An event is a statement that can be true or false. Three events about the asset are

$$T > 28^\circ\text{C}, \qquad G < 5\,\text{kW/K}, \qquad N_w \geq 2,$$

where N_w is the number of corrective work orders raised on the coil in a year. A probability assigns to an event a number between zero and one. The statement

$$\Pr(T > 28^\circ\text{C}) = 0.10$$

means that the probability assigned to the temperature-exceedance event is 10 %. The statement is incomplete in one respect: it does not say how the probability was calculated or what information it incorporates. A probability computed before the commissioning test and one computed after thirty condition tests are different numbers for the same event, and the notation must say which is meant. That is the role of conditioning.

2.2.2 Conditional probability and the evidence

The probability of event A given the information B is written

$$\Pr(A \mid B), \tag{2.3}$$

and read "the probability of A, given B". In asset management the information on the right of the bar is the evidence: the commissioning test and the condition tests, the trend log of the building management system, the work-order and repair records, inspection results, the design sheet, manufacturer specifications and prior engineering knowledge. The book collects all of it in the bold symbol $\mathbf{y}$, and keeps a plain y with a subscript, such as y_T, for a single reading. The statement

$$\Pr(T > 28^{\circ}\mathrm{C} \mid \mathbf{y})$$

asks for the probability that the hall temperature exceeds 28°C, conditional on the available evidence. It differs from $\Pr(T > 28^{\circ}\mathrm{C})$, which does not condition on that evidence explicitly. The unknown, the event about the asset, stands on the left of the bar; everything treated as known stands on the right. Keeping that order fixed is what prevents questions from becoming entangled when several are asked of one asset.

2.2.3 The role of the model

A complete probability statement also names the model under which it is computed:

$$\Pr(A \mid \mathbf{y}, M), \tag{2.4}$$

where A is the event of interest, $\mathbf{y}$ the evidence, and M the assumed mathematical model, which comprises the physical relations of the plant graph and the probability assumptions about errors and unknowns. The thermal model of equation (2.2) assumes steady state; a model that adds the thermal storage of the hall would assign a different probability to the same temperature event under the same evidence, because it predicts a different temperature after a change of load. The book sometimes suppresses M to keep the notation readable. The dependence on the model remains in every probability, and a probability is a statement about the asset only to the extent that the model is.

2.2.4 Unknown quantities that are not events

An unknown conductance G is a numerical quantity, not an event; it has a value, not a truth value. What can be asked of it is its probability distribution, described for a continuous quantity by a probability density,

$$p(G \mid \mathbf{y}, M),$$

where the lowercase p denotes a density. Alternatively, a question can ask for the probability that the conductance falls within a range, for example

$$\Pr(5 < G < 7 \mid \mathbf{y}, M),$$

and that is an event probability, because "$5 < G < 7$" is a statement that is true or false. The two are connected by integration of the density over the range that defines the event:

$$\Pr(5 < G < 7 \mid \mathbf{y}, M) = \int_5^7 p(G \mid \mathbf{y}, M)\,\mathrm{d}G, \tag{2.5}$$

with the limits in kW/K. A distribution describes uncertainty across all possible values of a quantity; an event probability describes how much of that probability falls within a specified set of values. A density is not a probability: it has units, $\mathrm{kW}^{-1}\,\mathrm{K}$ here, and it can exceed one. Only its integral over a set is a probability.

2.2.5 A reusable notation

The questions of this chapter, and of the book, are written with the symbols of Table 2.2. The first column is the symbol, the second its meaning. The state $\mathbf{z}$ holds every free quantity of the plant graph at once, the physical quantities such as the log-conductance as well as the temperatures; there is no separate parameter vector. The action a stands for a maintenance decision such as inspecting, cleaning a coil, changing a filter, repairing, replacing, or scheduling a condition test; the symbol $\tilde{y}_a$ stands for the outcome of a candidate measurement a that has not yet been made, and y_{new} for a future outcome of the asset itself, such as the time of its next corrective work order.

Table 2.2: The reusable notation of probabilistic questions.

Symbol	Meaning
$\mathbf{z}$	state vector of the plant graph: every unknown quantity, physical parameters included
$\mathbf{y}$	observed evidence: readings, counts, records, priors
M, M_j	assumed model or system structure; one of several competing models
$p(\mathbf{y} \mid M)$	evidence of model M: the probability of the evidence with the model's state integrated out
B_{AB}	Bayes factor of model M_A against M_B, $p(\mathbf{y} \mid M_A)/p(\mathbf{y} \mid M_B)$; B_{fc} is M_{fouled} against M_{clean}
A	an event
$f(\mathbf{z})$, f^*	a quantity derived from the state; a requirement or limit on it
κ	assurance level: the posterior probability a requirement must reach
y_{new}	a future observation that stays random when the state is known
a	a maintenance action or decision, or a candidate measurement
$\tilde{y}_a$	the outcome of candidate measurement a, not yet observed
U_{p}, $U(a)$	a quantity of the posterior after $\tilde{y}_a$; its expectation over the outcomes, which ranks the candidates
$\boldsymbol{\Sigma}_a$	covariance of the state after candidate measurement a (the preposterior)
$L(a, \mathbf{z})$	loss or cost of action a when the state is $\mathbf{z}$: its consequence
$\mathcal{R}(a)$	expected loss of action a over the posterior, $\mathbb{E}[L(a, \mathbf{z}) \mid \mathbf{y}]$
c_a, C_e	the fixed cost of action a; the consequence of event e
$\mathbf{g}_a$, $\mathbf{H}_a$	gradient and curvature of $L(a, \mathbf{z})$ in the state at $\hat{\mathbf{z}}$
$f_t(\mathbf{z})$, $\mathbf{g}(t)$	a quantity at time t carried by a state indexed by time; its gradient
α, γ	the level and the trend entries of such a state: the quantity at the record's center $\bar{t}$ and its change per unit time
t^*	crossing time: the time at which f_t reaches a level f^*

With these symbols, questions as different as the conductance of an air path, the work-order rate of a coil design and the choice of a condition test are written in one mathematical language, and the differences between them become visible as differences in the expression, not in the words.

2.3 The probability over the state, the general forms and the five computations

Every question of Section 2.1 is an integral over the posterior of the state, a ratio of such integrals, or the minimum of one over a set of actions. The categories of question differ in the

integrand, and the integrands draw on five computations: the posterior, the trend, the evidence, the preposterior and the expected consequence.

2.3.1 The probability of an event as an integral over the state

Suppose that inference has produced the posterior distribution of the state,

$$p(\mathbf{z} \mid \mathbf{y}, M),$$

and that a question asks whether the state lies in some region $\mathcal{A}$ of the state space. For the hall, $\mathcal{A}$ might be the set of all combinations of temperature and log-conductance that satisfy an operating limit. The probability of the region is the posterior integrated over it,

$$\Pr(\mathbf{z} \in \mathcal{A} \mid \mathbf{y}, M) = \int_{\mathcal{A}} p(\mathbf{z} \mid \mathbf{y}, M)\, \mathrm{d}\mathbf{z}, \tag{2.6}$$

which is equation (2.5) with a vector in place of one quantity. The same expression is written with the indicator function of the event,

$$\mathbf{1}_A(\mathbf{z}) = \begin{cases} 1 & \text{if event } A \text{ holds at } \mathbf{z}, \\ 0 & \text{otherwise,} \end{cases}$$

which extends the integral to the whole state space and lets the indicator select the region:

$$\Pr(A \mid \mathbf{y}, M) = \int \mathbf{1}_A(\mathbf{z})\, p(\mathbf{z} \mid \mathbf{y}, M)\, \mathrm{d}\mathbf{z} = \mathbb{E}\big[\mathbf{1}_A(\mathbf{z}) \mid \mathbf{y}, M\big]. \tag{2.7}$$

The last form reads the probability of an event as the posterior expectation of its indicator. The equation states the central rule of the chapter: *the probability of an event is the posterior integrated over the states at which the event is true.* Many asset-management calculations are variations of it, differing only in the event and in how the integral is evaluated.

2.3.2 Future events and the posterior predictive

Some events remain random even when the state is known exactly. A coil design whose work-order rate is known still needs its next corrective work order at a random time; knowing the rate fixes only the probability of passing a season without one. For such a future event A_{future}, the model supplies the conditional probability $\Pr(A_{\text{future}} \mid \mathbf{z}, M)$, and the question asks for its probability given the evidence. The derivation uses three rules. By the law of total probability, the probability of the event is the integral of its joint probability with the state:

$$\Pr(A_{\text{future}} \mid \mathbf{y}, M) = \int \Pr(A_{\text{future}}, \mathbf{z} \mid \mathbf{y}, M)\, \mathrm{d}\mathbf{z}.$$

By the product rule, the joint factors into a conditional and a marginal:

$$\Pr(A_{\text{future}} \mid \mathbf{y}, M) = \int \Pr(A_{\text{future}} \mid \mathbf{z}, \mathbf{y}, M)\, p(\mathbf{z} \mid \mathbf{y}, M)\, \mathrm{d}\mathbf{z}.$$

Finally, the model assumes that the future event depends on the past evidence only through the state: once $\mathbf{z}$ is known, the readings and records of the past add nothing about it. Under that conditional independence, $\Pr(A_{\text{future}} \mid \mathbf{z}, \mathbf{y}, M) = \Pr(A_{\text{future}} \mid \mathbf{z}, M)$, and

$$\Pr(A_{\text{future}} \mid \mathbf{y}, M) = \int \Pr(A_{\text{future}} \mid \mathbf{z}, M)\, p(\mathbf{z} \mid \mathbf{y}, M)\, \mathrm{d}\mathbf{z}. \tag{2.8}$$

This is the posterior predictive probability, the foundation of every prediction of a future failure or work order. Equation (2.7) is the special case in which the conditional probability is an indicator, zero or one at each state. The difference between the two matters as soon as the event is a future failure rather than a property of the present asset, which Section 2.4 shows with a work-order log.

2.3.3 The general forms of the five categories

Asset-management questions are many, but the mathematical objects that answer them are few, and the book classifies questions by the object rather than by the domain word in the question. Five categories cover every question the book asks, each with its general forms, written here in the notation of Table 2.2: $p(\mathbf{z} \mid \mathbf{y})$ is the posterior density of the state; $f(\mathbf{z})$ is a quantity derived from the state, an entry being the case $f(\mathbf{z}) = z_i$, and f^* a value of it, a requirement in a threshold question; y_{new} is a future observation; M_j is a complete model of the asset, a drawing with its rows; and $\mathbf{\Sigma}_a$ is the covariance of the state after a candidate measurement a. Nothing of any particular asset appears in the forms. Two further forms, of a trend (Section 2.3.4) and of an expected consequence (Section 2.3.5), belong to computations that the categories call on rather than to a category of their own. Every later chapter opens its derivation with one of these forms and then specializes it to its asset.

Estimation. An estimation question asks what a quantity is and how well it is known. Its general form is the posterior distribution of the quantity, the probability that it lies at or below each value f^*,

$$\Pr(f(\mathbf{z}) \le f^* \mid \mathbf{y}) = \int \mathbf{1}\big[f(\mathbf{z}) \le f^*\big]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}, \tag{2.9}$$

which is equation (2.7) for the event $f(\mathbf{z}) \le f^*$, read as a function of f^*. It is summarized by a named location, the mean, the median or the mode, and a spread; the mean and the variance are

$$\begin{aligned} \mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}] &= \int f(\mathbf{z})\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}, \\ \mathrm{Var}(f(\mathbf{z}) \mid \mathbf{y}) &= \int \big(f(\mathbf{z}) - \mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}]\big)^2\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}, \end{aligned} \tag{2.10}$$

the median is the value f^* at which equation (2.9) equals $\frac{1}{2}$, and the mode is the value at which the density of $f(\mathbf{z})$ is largest. The uncertainty is also stated as the central 95 % credible interval $[f_L, f_U]$, whose endpoints leave 2.5 % of posterior probability in each tail:

$$\Pr(f(\mathbf{z}) \le f_L \mid \mathbf{y}) = 0.025, \qquad \Pr(f(\mathbf{z}) \le f_U \mid \mathbf{y}) = 0.975. \tag{2.11}$$

The quantiles of a monotone function of a quantity are the images of that quantity's quantiles, in the same order for an increasing function and in reverse order for a decreasing one; the median is a quantile and maps the same way, while the mean and the mode in general do not. Under the Laplace posterior $\mathbf{z} \mid \mathbf{y} \approx \mathcal{N}(\hat{\mathbf{z}}, \mathbf{\Sigma})$ an entry is read off directly, with location $\hat{z}_i$ and spread $\sqrt{\Sigma_{ii}}$, and a derived quantity through its gradient $\mathbf{g} = \partial f / \partial \mathbf{z}$ at $\hat{\mathbf{z}}$, with location $f(\hat{\mathbf{z}})$ and spread $\sqrt{\mathbf{g}^\top \mathbf{\Sigma} \mathbf{g}}$ by the delta method. The form is the posterior seen through one quantity: a distribution, of which any single number reported is a named summary. A health index is such a quantity, $h = \mathrm{e}^g$ read through the entry g that carries it.

Prediction under a threshold. A prediction under a threshold asks whether the asset meets a requirement. The requirement is written as the event $f(\mathbf{z}) \ge f^*$, "the requirement holds", with f^* the required value, and its general form is the posterior probability of that event, compared with an assurance level κ, the posterior probability the requirement must reach:

$$\Pr(f(\mathbf{z}) \ge f^* \mid \mathbf{y}) = \int \mathbf{1}\big[f(\mathbf{z}) \ge f^*\big]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}, \qquad \text{accept if } \Pr(f(\mathbf{z}) \ge f^* \mid \mathbf{y}) \ge \kappa. \tag{2.12}$$

A requirement that a quantity stay below an upper limit, such as a temperature $T \le T_{\max}$, is the same form with $f(\mathbf{z}) = -T$ and $f^* = -T_{\max}$; equivalently, the probability of the requirement is one minus the probability of the complementary event, the exceedance $T > T_{\max}$, and the rule is applied to the requirement, never to the exceedance. A joint event is the same integral with

the indicator of an intersection. The required value is the asset manager's and the assurance level is the owner's, and the form keeps the two apart. When the event lies in the future and stays random even when the state is known, the tail is taken of the predictive distribution of the future observation,

$$p(y_{\text{new}} \mid \mathbf{y}) = \int p(y_{\text{new}} \mid \mathbf{z})\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}, \tag{2.13}$$

of which equation (2.8) is the case of an event, with y_{new} the indicator of A_{future}. The predictive is the randomness the model gives the future outcome, averaged over what is not known about the state. When the quantity moves in time and the requirement concerns a date, the event is a crossing, and its form is the trend form of Section 2.3.4.

Hypothesis comparison. A hypothesis comparison asks which of two descriptions of the asset is true. Its general form is the ratio of the evidences of two complete models fitted to the same evidence, the Bayes factor,

$$B_{AB} = \frac{p(\mathbf{y} \mid M_A)}{p(\mathbf{y} \mid M_B)}, \tag{2.14}$$

where the evidence $p(\mathbf{y} \mid M)$ of a model is the probability of the observed evidence with the model's own state integrated out over its prior. The ratio multiplies the prior odds of the two models into their posterior odds, and because each evidence integrates its own state out, it compares the drawings, not the values of their entries. The Bayes factor B_{fc} of the notation table is the case $M_A = M_{\text{fouled}}$ against $M_B = M_{\text{clean}}$.

Diagnosis. A diagnosis asks where and what a fault is. Its general form is hypothesis comparison run over a family M_j, $j = 1, \ldots, J$, of candidate faults generated from the plant graph, one for each component or reading that could be at fault, with the answer returned as the ranking of

$$\Pr(M_j \mid \mathbf{y}) \propto p(\mathbf{y} \mid M_j)\, \Pr(M_j), \qquad j = 1, \ldots, J. \tag{2.15}$$

The answer is the order of the candidates with the posterior probability of each, not a single verdict.

Design. A design question asks what to measure or do next, before the measurement exists. Its general form is the preposterior expectation: for a candidate measurement a whose outcome $\tilde{y}_a$ is not yet known, the expected value, over the outcomes weighted by their predictive distribution, of a quantity U_{p} of the posterior that the outcome would produce,

$$U(a) = \mathbb{E}_{\tilde{y}_a \mid \mathbf{y}}\Big[\, U_{\text{p}}\big(p(\mathbf{z} \mid \mathbf{y}, \tilde{y}_a)\big)\Big] = \int U_{\text{p}}\big(p(\mathbf{z} \mid \mathbf{y}, \tilde{y}_a)\big)\, p(\tilde{y}_a \mid \mathbf{y})\, \mathrm{d}\tilde{y}_a, \tag{2.16}$$

where $p(\tilde{y}_a \mid \mathbf{y})$ is the predictive (2.13) of the outcome; for an outcome that takes discrete values, such as a number of work orders in a season, the integral is a sum over those values weighted by their predictive probabilities. The quantity U_{p} is one of four: the spread of a derived quantity after the outcome; the information gained; the indicator that an acceptance rule (2.12) is met after the outcome, which makes $U(a)$ the probability that the campaign ends with the requirement accepted; or the expected loss of the best action once the outcome is known. The candidates are ranked by $U(a)$, the largest first when U_{p} is an information gain or an acceptance indicator, the smallest first when U_{p} is a spread or a loss, and a candidate's cost enters as a loss of its own. When the posterior is Gaussian and the candidate is a linear Gaussian row, $r_a(\mathbf{z}) = (\mathbf{h}_a\mathbf{z} - \tilde{y}_a)/\sigma_a$ with $\mathbf{h}_a$ the row vector of its sensitivities and σ_a the standard deviation of its noise, the covariance of the updated posterior does not depend on

the outcome, and for a spread or an information gain the form needs only two objects: the preposterior covariance,

$$\boldsymbol{\Sigma}_a = \left(\boldsymbol{\Sigma}^{-1} + \boldsymbol{\Lambda}_a\right)^{-1}, \qquad \boldsymbol{\Lambda}_a = \mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a = \frac{\mathbf{h}_a^{\mathsf{T}}\mathbf{h}_a}{\sigma_a^2}, \tag{2.17}$$

with $\mathbf{J}_a = \mathbf{h}_a/\sigma_a$ the Jacobian of the candidate row, and the expected information gain, in nats,

$$\mathrm{EIG}(a) = \tfrac{1}{2}\ln\frac{\det\boldsymbol{\Sigma}}{\det\boldsymbol{\Sigma}_a}. \tag{2.18}$$

The spread of a derived quantity f after the measurement is then $\sqrt{\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}_a\mathbf{g}}$, with the same gradient $\mathbf{g}$ at $\hat{\mathbf{z}}$. For a count, demand or lifetime row the candidate's information is the expected information: $\boldsymbol{\Lambda}_a = \mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a$ evaluated at $\hat{\mathbf{z}}$ with the expected counts of the campaign in place of observed ones, a count row entering $\mathbf{J}_a$ through the line $\sqrt{\mu_\ell}\,\partial \ln\mu_\ell/\partial\mathbf{z}$ of its working form (Section 4.2.7); equation (2.17) is then the Laplace preposterior, exact only to the order of the Laplace approximation, and it still does not depend on the outcome. A design question values a posterior that does not exist yet, through what the measurement would do to it. When the candidate is an action that brings no new evidence, a repair or a cleaning, there is no outcome to average over, the posterior after the action is the present one, and the form reduces to the expected consequence of Section 2.3.5.

2.3.4 The general form of a trend

Some questions concern a quantity that moves in time: whether a conductance is falling, when it will cross the level at which the hall can no longer be held, what it will be next year. By its integrand each is an estimation or a prediction under a threshold; what sets it apart is that the quantity is indexed by time, so that the state must carry how the quantity changes as well as where it stands. Write the quantity, on the scale on which the state carries it (a log-conductance, a log-rate, a temperature), as $f_t(\mathbf{z})$, its value at time t. In the simplest drawing it is a straight line in time about the center $\bar{t}$ of the record, with a level entry α, the quantity at $\bar{t}$, and a *trend entry* γ, its change per unit time (the degradation rate, when the quantity is a health index):

$$\mathbf{z} = \begin{bmatrix}\alpha\\ \gamma\end{bmatrix}, \qquad f_t(\mathbf{z}) = \alpha + \gamma\,(t-\bar{t}) = \mathbf{g}(t)^{\mathsf{T}}\mathbf{z}, \qquad \mathbf{g}(t) = \frac{\partial f_t}{\partial \mathbf{z}} = \begin{bmatrix}1\\ t-\bar{t}\end{bmatrix}. \tag{2.19}$$

The evidence arrives in windows, one row per window. Window ℓ, centered at the time t_ℓ, gives the row $r_\ell(\mathbf{z})$, which sees the state only through the quantity at its own time, $f_{t_\ell}(\mathbf{z})$, so its Jacobian row is the gradient $\mathbf{g}(t_\ell)^{\mathsf{T}}$ times the row's sensitivity $d_\ell = \partial r_\ell/\partial f_{t_\ell}$: $1/\sigma_\ell$ for a reading row of standard deviation σ_ℓ, and $\sqrt{\mu_\ell}$, the square root of the window's expected count, for a count row in the working form of Section 4.2.7. The rows of the record stack to

$$\begin{gathered}\mathbf{J} = \mathbf{D}\,\mathbf{X}, \qquad \mathbf{X} = \begin{bmatrix}\mathbf{g}(t_1)^{\mathsf{T}}\\ \mathbf{g}(t_2)^{\mathsf{T}}\\ \vdots\end{bmatrix}, \qquad \mathbf{D} = \mathrm{diag}(d_1, d_2, \ldots),\\ \boldsymbol{\Lambda} = \mathbf{J}^{\mathsf{T}}\mathbf{J} = \begin{bmatrix}\sum_\ell d_\ell^2 & \sum_\ell d_\ell^2\,(t_\ell-\bar{t})\\ \sum_\ell d_\ell^2\,(t_\ell-\bar{t}) & \sum_\ell d_\ell^2\,(t_\ell-\bar{t})^2\end{bmatrix}, \qquad \boldsymbol{\Sigma} = \boldsymbol{\Lambda}^{-1}.\end{gathered} \tag{2.20}$$

The trend entry of the information, $\sum_\ell d_\ell^2(t_\ell-\bar{t})^2$, weighs each window's distance from the center squared, while the windows and their precisions d_ℓ^2 add linearly: widening the span of a record raises the information on the trend faster than adding windows inside it. Chapter 4 writes the same line as the limit of trend rows that tie each window's entry to the next; the straight line here is that limit.

The posterior of this state is read along time in three ways, and the three make up the general form of a trend:

$$\begin{aligned}
\Pr(\gamma > 0 \mid \mathbf{y}) &= \int \mathbf{1}[\gamma > 0]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}, \\
\Pr(t^* \le s \mid \mathbf{y}) &= \int \mathbf{1}[\gamma > 0]\, \mathbf{1}\big[f_s(\mathbf{z}) \ge f^*\big]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}, \\
\Pr(f_t(\mathbf{z}) \le f^* \mid \mathbf{y}) &= \int \mathbf{1}\big[f_t(\mathbf{z}) \le f^*\big]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}, \qquad t \text{ beyond the record.}
\end{aligned} \tag{2.21}$$

The first line is the sign of the trend, the probability that the quantity is rising: the threshold form on the trend entry. The second is the distribution function of the crossing time $t^*(\mathbf{z})$, the time at which the quantity reaches the level f^* from below. A rising path, $\gamma > 0$, has crossed by a time s if and only if it is at or above f^* at s, whether s lies before or after $\bar{t}$, and a path with $\gamma \le 0$ never crosses, a mass $\Pr(\gamma \le 0 \mid \mathbf{y})$ at $t^* = \infty$: the joint indicator is the crossing, the first passage of a rising path. The threshold probability $\Pr(f_s(\mathbf{z}) \ge f^* \mid \mathbf{y})$ alone is not the crossing, because it also counts the states with $\gamma \le 0$ that stand at or above the level at s, paths that started above it and have never crossed it; the two agree only when the posterior puts no such state there. The third is the forecast, the estimation form at a time beyond the last window of the record. Nothing of a particular asset appears in the three lines; an asset enters through the scale of f_t, the rows of its windows and the level f^*. A quantity that falls to its level, as a health index does, is the same form with $f_t = -g_t$ and $f^* = -g^*$: the sign of the trend is then $\Pr(\gamma < 0 \mid \mathbf{y})$ on the trend of g, the crossing by s is the joint event $\gamma < 0$ and $g_s \le g^*$, and the mass that never crosses is $\Pr(\gamma \ge 0 \mid \mathbf{y})$.

Under the Laplace posterior $\mathbf{z} \mid \mathbf{y} \approx \mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ the quantity at any time is linear in the state, so the sign and the forecast are exact normal tails in two numbers, the center $\mathbf{g}(t)^{\mathsf{T}}\hat{\mathbf{z}}$ and the variance $\mathbf{g}(t)^{\mathsf{T}}\boldsymbol{\Sigma}\,\mathbf{g}(t)$, and the crossing is an exact bivariate normal probability of the pair (γ, f_s), whose correlation $\varrho(s)$ is read from the same $\boldsymbol{\Sigma}$:

$$\begin{gathered}
\Pr(\gamma > 0 \mid \mathbf{y}) \approx \Phi\!\left(\frac{\hat{\gamma}}{\sqrt{\Sigma_{\gamma\gamma}}}\right), \qquad \Pr(t^* \le s \mid \mathbf{y}) \approx \Phi_2\!\left(\frac{\hat{\gamma}}{\sqrt{\Sigma_{\gamma\gamma}}},\ \frac{\mathbf{g}(s)^{\mathsf{T}}\hat{\mathbf{z}} - f^*}{\sqrt{\mathbf{g}(s)^{\mathsf{T}}\boldsymbol{\Sigma}\,\mathbf{g}(s)}};\ \varrho(s)\right), \\
\mathbf{g}(t)^{\mathsf{T}}\boldsymbol{\Sigma}\,\mathbf{g}(t) = \Sigma_{\alpha\alpha} + 2\,(t - \bar{t})\,\Sigma_{\alpha\gamma} + (t - \bar{t})^2\,\Sigma_{\gamma\gamma}, \qquad \varrho(s) = \frac{\Sigma_{\alpha\gamma} + (s - \bar{t})\,\Sigma_{\gamma\gamma}}{\sqrt{\Sigma_{\gamma\gamma}\,\mathbf{g}(s)^{\mathsf{T}}\boldsymbol{\Sigma}\,\mathbf{g}(s)}},
\end{gathered} \tag{2.22}$$

where $\Phi_2(u_1, u_2; \varrho)$ is the distribution function of two standard normals of correlation ϱ, the probability that the first lies below u_1 and the second below u_2. It is one integral, $\Phi_2(u_1, u_2; \varrho) = \int_{-\infty}^{u_1} \varphi(x)\, \Phi\big((u_2 - \varrho x)/\sqrt{1 - \varrho^2}\big)\, \mathrm{d}x$, and it lies below the single tail $\Phi(u_2)$ by the mass of the states that stand beyond the level without a rising trend. The sign is the single tail with the gradient $(0,\ 1)^{\mathsf{T}}$, which reads the trend entry alone, so that its variance is $\Sigma_{\gamma\gamma}$; the forecast is the same with the time asked. The variance of the quantity grows with the square of the distance $t - \bar{t}$, the lever arm: a forecast pays for the uncertainty of the trend with that square. The crossing time itself, $t^* = \bar{t} + (f^* - \alpha)/\gamma$ for $\gamma > 0$, is a ratio with the trend entry in its denominator. When the posterior of γ has a positive density at zero, as a Gaussian posterior always has, t^* has no posterior mean and no variance, so it is reported by its median, the date at which the second line is one half, which lies near the plug-in date $t^*(\hat{\mathbf{z}})$ when little of the posterior stands beyond the level without a rising trend, by the band whose ends solve the second line at 0.025 and 0.975, and by the mass that never crosses; the delta band $t^*(\hat{\mathbf{z}}) \pm 1.96\sqrt{\mathbf{g}_*^{\mathsf{T}}\boldsymbol{\Sigma}\,\mathbf{g}_*}$, with $\mathbf{g}_* = \partial t^*/\partial \mathbf{z}$ at $\hat{\mathbf{z}}$, is a finite number standing in for an infinite one.

The trend on the hall's two tests. The hall's own record shows each of these readings. Its two tests, the commissioning test of window 0 at year 0 and the condition test of window 24 at

year 2, both at 30 kW, are two reading rows of standard deviation 0.5 K, and the handbook prior $\mathcal{N}(\ln(7.5/6), 0.3^2)$ is a row on the commissioned log-conductance; there is no prior row on the trend. Time t is in years in this chapter, window $12t$, and the crossing time t^* is a year. Chapter 4 fits this record, and its straight line, the last row of its table in the tolerance σ_q of the trend rows, is the limit $\sigma_q \to 0$ used here. Centered at the middle of the record, $\bar{t} = 1$ year (window 12), its state and Laplace covariance, the curvature term of the nonlinear reading rows included, are

$$\hat{\mathbf{z}} = \begin{bmatrix} \hat{\alpha} \\ \hat{\gamma} \end{bmatrix} = \begin{bmatrix} -0.054618 \\ -0.040692\,\mathrm{yr}^{-1} \end{bmatrix}, \qquad \boldsymbol{\Sigma} = \begin{bmatrix} 4.312 \times 10^{-3} & -1.80 \times 10^{-4} \\ -1.80 \times 10^{-4} & 4.312 \times 10^{-3} \end{bmatrix},$$

a degradation rate of $-3.99\,\%$ of conductance a year with $\sqrt{\Sigma_{\gamma\gamma}} = 0.065668\,\mathrm{yr}^{-1}$, larger than the rate itself. The health index falls, so the sign of the trend is the first line of equation (2.22) for $-\gamma$, $\Pr(\gamma < 0 \mid \mathbf{y}) = \Phi(0.6197) = 0.732$; the quadrature of Chapter 4 over the exact posterior of the two tests gives 0.7292 beside it. The level is the functional limit: the hall can no longer hold 24°C at today's 30 kW when G falls to $G_F = 30/(24 - 18) = 5\,\mathrm{kW/K}$, the health index $h_F = 0.833$ and $g_F = -0.182322$. The plug-in date $t^*(\hat{\mathbf{z}})$ is year 4.14, 2.14 years after the later test. The crossing by s is the first passage of a falling path, $\Pr(t^* \le s \mid \mathbf{y}) = \Pr(\gamma < 0,\ g_s \le g_F \mid \mathbf{y})$, the second line of equation (2.22) for $-g$; at the plug-in date it is 0.4990, because a mass of 0.0010 stands below the limit there with $\gamma \ge 0$ and has never crossed, so the median crossing is year 4.15 (window 49.8). The exact posterior of the two tests, by quadrature, puts the median at year 4.32 (window 51.8).

Figure 2.2 draws the record and the crossing. In panel (a) the horizontal axis is the year and the vertical axis the log-conductance g; the two dots are the tests as the dashboard reads them, each reading turned into g as if it were exact; the green line is $\mathbf{g}(t)^{\mathsf{T}}\hat{\mathbf{z}}$ and the shaded band its 95 % band $\pm 1.96\sqrt{\mathbf{g}(t)^{\mathsf{T}}\boldsymbol{\Sigma}\,\mathbf{g}(t)}$; the dashed red line is the functional limit g_F. In panel (b) the horizontal axis is the date s in years and the vertical axis the probability that the hall has crossed the functional limit by s, the first passage of the second line of equation (2.22); the dashed gray line is its ceiling, $\Pr(\gamma < 0 \mid \mathbf{y})$; the dotted lines mark the median and the year at which the dashboard's straight line on the approach reaches 6 K; and the bar above the curve is the delta band. The line of panel (a) passes through the window-24 dot but not the window-0 dot, because the handbook prior acts on window 0 and pulls the commissioned value up, while the later reading, with nothing else on it, is fitted exactly. The band of panel (a) opens with the square of the lever arm: its lower edge stands at -0.1833 against the line's -0.0546 at the center of the record and at -1.5810 against -0.4208 at year 10, while its upper edge stays above the commissioned value at every year drawn. In panel (b) the distribution function does not rise to one. It stands at 0.0181 at the center of the record and at 0.1665 at the later test, and its 0.025 point, year 1.13 (window 13.6), lies inside the record; it passes 0.5 at the median and then flattens toward the ceiling, 0.7156 at year 40, never above 0.7323, because a mass $\Pr(\gamma \ge 0 \mid \mathbf{y}) = 0.268$ never crosses. *The band by inversion has no upper end.* The delta band, with $\mathbf{g}_* = (24.575,\ 77.122)^{\mathsf{T}}$ and a spread of 5.251 years, runs from -6.15 to 14.43 years: finite where the upper end is infinite and beginning before the commissioning test, it is wrong at both ends. The dashboard's line lands at year 5.33, 1.20 years after the plug-in date, and carries no band at all. Exercise 2.5 works the crossing by year 5 from these numbers.

2.3.5 The general form of an expected consequence

Other questions ask what to do: whether to change a filter before the summer peak, whether to clean a coil now or at the shutdown, whether to repair or replace a unit. Each compares actions a from a set of candidates by what they are expected to cost. Let $L(a, \mathbf{z})$ be the loss of taking action a when the state is $\mathbf{z}$, its consequence in whatever unit the decision is taken in: reais, hours of shed IT load, lost service. It is a derived quantity of the state, one for each

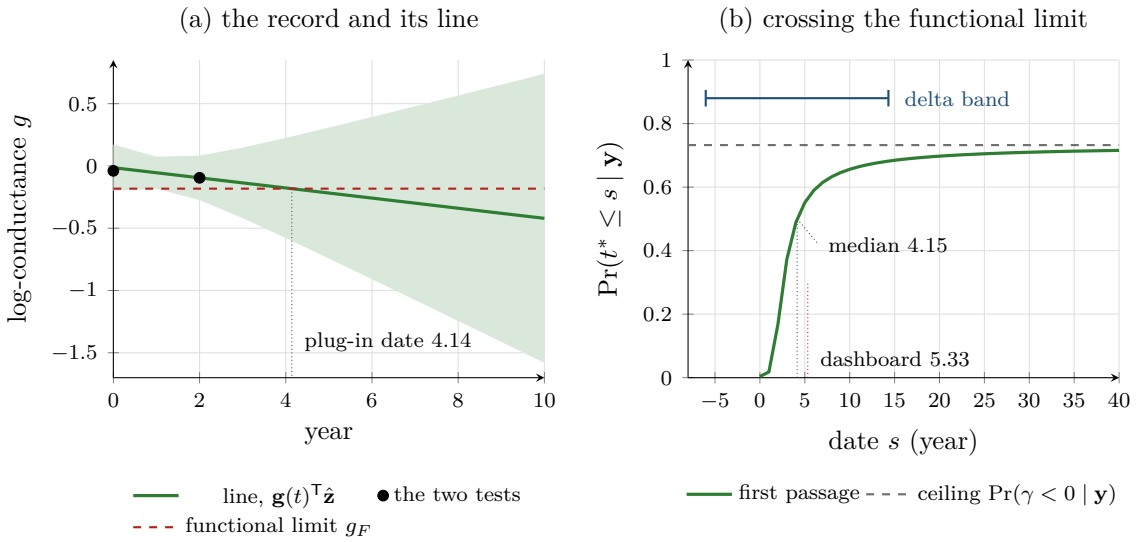


Figure 2.2: The trend form on the hall's two tests (straight line, t in years, $\bar{t} = 1$). (a) The tests as the dashboard reads them, $g = \ln(30/(y_T - 18)/6)$, at -0.0392 and -0.0953; the line $\hat{\alpha} + \hat{\gamma}(t - 1)$ with $\hat{\alpha} = -0.054618$, $\hat{\gamma} = -0.040692\,\mathrm{yr}^{-1}$ and its 95 % band; the functional limit $g_F = \ln(5/6) = -0.1823$ and the plug-in date 4.14. (b) The first passage $\Pr(t^* \le s \mid \mathbf{y}) = \Pr(\gamma < 0,\ g_s \le g_F \mid \mathbf{y})$, a bivariate normal probability under the Laplace posterior, its ceiling 0.7323, the median 4.15, the year 5.33 of the dashboard's line, and the delta band from -6.15 to 14.43 years.

action. The general form of an expected consequence is its posterior expectation, the expected loss $\mathcal{R}(a)$ of the action, and the action that minimizes it. When the action costs c_a for certain and leaves events e that occur with probability $\Pr(e \mid \mathbf{z}, a)$ and have the consequence C_e each, the loss given the state is the cost of the action plus the consequence of each event weighted by its probability, and by the predictive (2.8) the posterior expectation turns each conditional probability into its predictive one:

$$\begin{aligned} \mathcal{R}(a) &= \mathbb{E}[L(a,\mathbf{z}) \mid \mathbf{y}] = \int L(a,\mathbf{z})\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}, \qquad a^* = \arg\min_a \mathcal{R}(a), \\ L(a,\mathbf{z}) &= c_a + \sum_e C_e \Pr(e \mid \mathbf{z}, a) \quad \Longrightarrow \quad \mathcal{R}(a) = c_a + \sum_e C_e \Pr(e \mid \mathbf{y}, a). \end{aligned} \tag{2.23}$$

The second line is the rule that risk is probability times consequence, summed over the events and added to the certain cost of the action. Nothing of a particular asset appears in it either; an asset enters through the events, the probabilities its posterior gives them and the consequence schedule c_a, C_e that the asset owner supplies.

Under the Laplace posterior the expected consequence is read from the same objects as every other answer. With $\mathbf{g}_a = \partial L(a,\mathbf{z})/\partial \mathbf{z}$ and $\mathbf{H}_a = \partial^2 L(a,\mathbf{z})/\partial \mathbf{z}\,\partial \mathbf{z}^\mathsf{T}$ at $\hat{\mathbf{z}}$, the expansion of the loss to second order about $\hat{\mathbf{z}}$, averaged over $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$, keeps the constant and the curvature term, since $\mathbb{E}[\mathbf{z} - \hat{\mathbf{z}} \mid \mathbf{y}] = \mathbf{0}$ and $\mathbb{E}[(\mathbf{z} - \hat{\mathbf{z}})^\mathsf{T} \mathbf{H}_a (\mathbf{z} - \hat{\mathbf{z}}) \mid \mathbf{y}] = \mathrm{tr}(\mathbf{H}_a \boldsymbol{\Sigma})$; the spread follows from the gradient, and the comparison of two actions from the gradient of their difference:

$$\begin{aligned} \mathcal{R}(a) &\approx L(a,\hat{\mathbf{z}}) + \tfrac{1}{2}\,\mathrm{tr}(\mathbf{H}_a \boldsymbol{\Sigma}), \qquad \mathrm{sd}(L(a,\mathbf{z}) \mid \mathbf{y}) \approx \sqrt{\mathbf{g}_a^\mathsf{T} \boldsymbol{\Sigma}\, \mathbf{g}_a}, \\ \Pr(L(a,\mathbf{z}) < L(a',\mathbf{z}) \mid \mathbf{y}) &\approx \Phi\left(\frac{\mathcal{R}(a') - \mathcal{R}(a)}{\sqrt{(\mathbf{g}_a - \mathbf{g}_{a'})^\mathsf{T} \boldsymbol{\Sigma}\, (\mathbf{g}_a - \mathbf{g}_{a'})}} \right). \end{aligned} \tag{2.24}$$

The loss at the most probable state, $L(a,\hat{\mathbf{z}})$, misses the curvature term. The curvature is taken in the coordinates of the state, and its sign there, not in a physical quantity, sets the direction

of the second-order term. The work-order log of Section 2.4.4 shows the difference over one year. The probability of a year without a corrective work order, $\mathrm{e}^{-\lambda\tau}$, is convex in the rate λ, so by Jensen's inequality the predictive 0.7725 lies above the plug-in 0.7659; but the state carries the log-rate ω, in which the probability is concave, with curvature -0.150 at $\hat{\omega}$, and the second-order reading with $\sigma_\omega^2 = 0.25$ moves the plug-in down to 0.747, close to the average 0.748 over the Laplace normal of the log-rate and away from the exact 0.7725. The term is right about the Laplace posterior and wrong about the coil design, because the error lies in the Laplace posterior itself; where the second-order term and an exact route move the plug-in in opposite directions, the value read from the state is stated with the exact one beside it, and the gap is what the Laplace posterior costs. The second line of equation (2.24) asks how firm a ranking is: the probability that action a costs less than a' at the true state. The variance of the difference is the sum of the two variances minus $2\,\mathbf{g}_a^{\mathsf{T}}\boldsymbol{\Sigma}\,\mathbf{g}_{a'}$, so the band of the comparison is narrower than the two bands taken apart exactly when $\mathbf{g}_a^{\mathsf{T}}\boldsymbol{\Sigma}\,\mathbf{g}_{a'} > 0$, when the two losses move together across the posterior. Section 2.5.4 applies the form to a filter change before the summer peak, and the repair-or-replace, inspection, trigger, priority and warranty questions of Chapters 17 to 21, in Volume II, are this form over the posterior that the earlier chapters produce.

2.3.6 The five computations beneath the forms

Five computations sit underneath the forms, each defined by the object it returns: the posterior of the state, the trend, the evidence, the preposterior and the expected consequence.

The posterior. The posterior of the state follows from Bayes' rule, the likelihood of the evidence times the prior divided by their integral over the state. For Gaussian rows divided by their standard deviations it is the exponential of minus the objective of the rows,

$$p(\mathbf{z} \mid \mathbf{y}) = \frac{p(\mathbf{y} \mid \mathbf{z})\,p(\mathbf{z})}{p(\mathbf{y})} \propto \exp\bigl[-\mathcal{L}(\mathbf{z})\bigr], \qquad \mathcal{L}(\mathbf{z}) = \tfrac{1}{2}\,r(\mathbf{z})^{\mathsf{T}} r(\mathbf{z}), \tag{2.25}$$

and a count, demand or lifetime row puts its negative log-probability term into $\mathcal{L}$ instead. Under the Laplace approximation the posterior is the normal $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ centered at the MAP state $\hat{\mathbf{z}}$, the minimum of $\mathcal{L}$, with the covariance $\boldsymbol{\Sigma} = (\mathbf{J}^{\mathsf{T}}\mathbf{J})^{-1}$ built from the Jacobian $\mathbf{J}$ of the rows; Chapter 3 derives the normal equations and the covariance, and Chapter 4 the Laplace approximation. The estimation, threshold and predictive forms are integrals against this posterior, read for an entry $\hat{z}_i$ with $\sqrt{\Sigma_{ii}}$ or for a derived quantity through $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$.

The trend. The trend is the posterior over a state indexed by time or by window, with a trend entry among its entries, read along time: as the probability that the trend entry is positive, as a crossing time with its band, or as a forecast, the three lines of equation (2.21). Its building blocks are those of the posterior, $\hat{\mathbf{z}}$ and $\boldsymbol{\Sigma}$ from the rows of the windows (2.20), with the gradient $\mathbf{g}(t)$ of the quantity at the time asked. It is a reading of the posterior, and it is named as a computation of its own because its object, its check and its failures are its own: the object is a function of the time asked rather than one number; the check orders the record's residuals in time and inverts the threshold line for the band of the crossing instead of linearizing the crossing time; and it fails where no other reading does, in a forecast that pays for the uncertainty of the trend with the square of its lever arm, in a crossing time without a mean when the posterior of the trend reaches zero, and in the shape of the line, which the record tests only inside its own span.

The evidence. The denominator of equation (2.25), written with its model explicit, is the evidence of the model, the probability of the observed evidence with the model's own unknowns

integrated out over their prior:

$$p(\mathbf{y} \mid M_j) = \int p(\mathbf{y} \mid \mathbf{z}, M_j)\, p(\mathbf{z} \mid M_j)\, \mathrm{d}\mathbf{z}. \tag{2.26}$$

It is the normalizing constant of the posterior within model M_j, and it is the only place where the comparison between models enters: the posterior inside either model, however sharp, does not say which model is true. Under the Laplace approximation it is read from the model's own fit, as the volume of its posterior times the height of the integrand at its MAP state, $p(\mathbf{y} \mid M_j) \approx (2\pi)^{d_j/2} \det(\mathbf{\Sigma}_j)^{1/2}\, p(\mathbf{y} \mid \hat{\mathbf{z}}_j, M_j)\, p(\hat{\mathbf{z}}_j \mid M_j)$ with d_j the number of entries of its state, exact when the objective is quadratic (Section 4.2.2). The comparison and diagnosis forms are a ratio and a ranking of it.

The preposterior. The preposterior is the expected posterior under a measurement not yet taken: the posterior that each outcome $\tilde{y}_a$ would produce, averaged over the predictive distribution of the outcomes, the inner object of the design form (2.16). For a linear Gaussian candidate row its covariance does not depend on the outcome, and equation (2.17) follows from the rows. Let the current posterior of the state be approximately Gaussian with mean $\hat{\mathbf{z}}$ and covariance $\mathbf{\Sigma} = (\mathbf{J}^\mathsf{T}\mathbf{J})^{-1}$, and let the proposed measurement depend on the state through its sensitivities, $\tilde{y}_a = \mathbf{h}_a\mathbf{z} + \varepsilon$ to first order, with independent Gaussian noise ε of variance σ_a^2. Divided by its standard deviation, the candidate row is $r_a(\mathbf{z}) = (\mathbf{h}_a\mathbf{z} - \tilde{y}_a)/\sigma_a$, and its Jacobian is the row vector $\mathbf{J}_a = \partial r_a/\partial \mathbf{z} = \mathbf{h}_a/\sigma_a$. The objective after the measurement is the current one plus that row,

$$\mathcal{L}_a(\mathbf{z}) = \tfrac{1}{2}\, r(\mathbf{z})^\mathsf{T} r(\mathbf{z}) + \tfrac{1}{2}\, r_a(\mathbf{z})^2,$$

and its curvature, the information, is the curvature of the stacked Jacobian $[\mathbf{J};\ \mathbf{J}_a]$, the sum $\mathbf{J}^\mathsf{T}\mathbf{J} + \mathbf{J}_a^\mathsf{T}\mathbf{J}_a$: the information of independent rows adds. With $\mathbf{J}^\mathsf{T}\mathbf{J} = \mathbf{\Sigma}^{-1}$, the covariance after the measurement under the linear-Gaussian approximation is the inverse of that sum, which is equation (2.17). When the current rows include exact constraints, $\mathbf{\Sigma}$ is singular in the limit $\sigma_p \to 0$ and $\mathbf{\Sigma}^{-1}$ does not exist there. The limit is then taken through the equivalent update form, which uses $\mathbf{\Sigma}$ alone,

$$\mathbf{\Sigma}_a = \mathbf{\Sigma} - \mathbf{\Sigma}\mathbf{J}_a^\mathsf{T}\big(\mathbf{I} + \mathbf{J}_a\mathbf{\Sigma}\mathbf{J}_a^\mathsf{T}\big)^{-1}\mathbf{J}_a\mathbf{\Sigma}, \tag{2.27}$$

where $\mathbf{I}$ is the identity of the size of the number of candidate rows. The value $\tilde{y}_a$ of the measurement does not appear in either form, only its Jacobian, its sensitivity divided by its noise. That is why equation (2.17) predicts how the uncertainty would shrink before the measurement is taken, and why it is one foundation of sensor placement and of the planning of condition tests. The information gain (2.18) is the standard result for a Gaussian prior and a linear Gaussian measurement; where both determinants vanish in the limit of exact rows, the determinant identity $\det\mathbf{\Sigma}/\det\mathbf{\Sigma}_a = \det(\mathbf{I} + \mathbf{J}_a\mathbf{\Sigma}\mathbf{J}_a^\mathsf{T})$, which also uses $\mathbf{\Sigma}$ alone, evaluates it.

The expected consequence. This computation returns the expected loss of each candidate action over the posterior, $\mathcal{R}(a) = \mathbb{E}[L(a, \mathbf{z}) \mid \mathbf{y}]$, and the choice of the action that minimizes it, equation (2.23); summed over events, it is probability times consequence. Its building blocks are $\hat{\mathbf{z}}$ and $\mathbf{\Sigma}$ with the gradient $\mathbf{g}_a$ and the curvature $\mathbf{H}_a$ of each action's loss, equation (2.24). It is the expectation of one more derived quantity, and it is named as a computation of its own because its object is a choice among actions rather than a description of the state, because its check is the curvature term that separates it from the loss at the most probable state together with the probability that the ranking holds, and because its failures are its own: a second-order reading whose sign follows the coordinates of the state, a ranking that is clear in expectation and uncertain at the true state, and a consequence schedule that the posterior does not contain, whose own uncertainty no band of the state carries.

Table 2.3 sets the five categories against the five computations: its first column is the category, its second the computation with the kind of question in the category that calls on it, and its third the object the computation returns. A category calls on more than one computation when its questions differ in kind.

Table 2.3: The five categories of question, the computations that answer them, and the object each computation returns.

Category	Computation	Object returned
estimation	posterior	an entry, $\hat{z}_i$ with $\sqrt{\Sigma_{ii}}$; a derived quantity, $f(\hat{\mathbf{z}})$ with $\sqrt{\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}}$
	trend, for the quantity at a stated time and for the distribution of a crossing time	$\mathbf{g}(t)^{\mathsf{T}}\hat{\mathbf{z}}$ with $\mathbf{g}(t)^{\mathsf{T}}\boldsymbol{\Sigma}\,\mathbf{g}(t)$; the median crossing time $t^*(\hat{\mathbf{z}})$ and its band by inversion
	expected consequence, for the expected loss of one course of action	$\mathcal{R}(a) = \mathbb{E}[L(a, \mathbf{z}) \mid \mathbf{y}]$
prediction under a threshold	posterior	a tail of an entry or of a derived quantity; a predictive
	trend, for the sign of a trend and for a crossing by a date	$\Pr(\gamma > 0 \mid \mathbf{y})$; $\Pr(t^* \le s \mid \mathbf{y})$
hypothesis comparison	evidence	a Bayes factor $B_{AB} = p(\mathbf{y} \mid M_A)/p(\mathbf{y} \mid M_B)$
diagnosis	evidence	the ranking of $\Pr(M_j \mid \mathbf{y})$ over a family generated from the graph
design	preposterior, for the next measurement	$\boldsymbol{\Sigma}_a$, the information gain, the value of information
	expected consequence, for the choice of an action	$a^* = \arg\min_a \mathcal{R}(a)$; the probability that the ranking holds

In this book the posterior is used in every chapter; the trend in Chapter 4, which derives the trend rows, and in Chapters 10, 11 and 13, and in Chapters 14, 16, 17, 18 and 20 to 24 of Volume II, where the remaining-life band feeds the decisions and the monitoring questions; the evidence in Chapters 4, 7, 11 and 12, and in Chapter 21 of Volume II; the preposterior at step 7 of every chapter that solves a problem, and as the lead computation of Chapters 22 to 24 of Volume II; and the expected consequence in Chapters 17 to 21 of Volume II, for the decisions of the asset's life.

2.3.7 The finer question types inside the five categories

A finer list of question types is useful in practice, because each type has a characteristic wording. Table 2.4 lists twelve of them. The first column is the type, the second an example in the asset manager's words, the third the mathematical target, and the last two the category of the book it belongs to and the computation beneath it. The list does not limit the applications; it collects the mathematical patterns that recur across them.

Most rows fall into a category directly; five need a sentence. A *joint event* such as "both need replacing" is an event like any other, a region of the state space, so its probability is equation (2.7) with the indicator of an intersection; the wording "both" changes the region, not the category. A *trend* question asks for threshold events, on the trend entry and on the quantity at a date, so its category is prediction under a threshold; its computation is the trend, because the state is indexed by time and the crossing time is a ratio that the threshold line

Table 2.4: Twelve question types, their mathematical targets, and the category and computation each belongs to.

Question type	Example	Mathematical target	Category	Computation
Estimation	How far has the coil's conductance fallen against its design sheet?	$p(h \mid \mathbf{y})$, $h = \mathrm{e}^{g}$	estimation	posterior
Threshold probability	Will the hall stay within its limit at the requested load?	$\Pr(T \le T_{\max} \mid \mathbf{y})$	prediction under a threshold	posterior
Joint and conditional events	Will the fan motor and the valve actuator both need replacing this year?	$\Pr(F_A \cap F_B \mid \mathbf{y})$	prediction under a threshold	posterior
Prediction	Will the summer season pass without a corrective work order?	$\Pr(T > 90 \text{ days} \mid \mathbf{y})$	prediction under a threshold	posterior
Trend and crossing	Is the conductance falling, and when will it reach the functional limit?	$\Pr(\gamma < 0 \mid \mathbf{y})$, $\Pr(t^{*} \le s \mid \mathbf{y})$	prediction under a threshold	trend
Derived quantities	How much heat can the air path carry within its limit?	$p(G(T_{\max} - T_c) \mid \mathbf{y})$	estimation	posterior
Model comparison	Is the conductance falling, or is the sensor drifting?	$\Pr(M_j \mid \mathbf{y})$	hypothesis comparison	evidence
Diagnosis	Which element explains the warmer hall?	$\Pr(F_j \mid \mathbf{y})$	diagnosis	evidence
Risk and expectation	What does waiting until the shutdown cost on average?	$\mathbb{E}[L \mid \mathbf{y}]$	estimation	expected consequence
Importance	Which entry moves the hall temperature most?	$\partial T / \partial z_i$, an entry of the gradient $\mathbf{g}$	estimation	posterior
Information and experiment design	Which condition test should be scheduled?	expected information gain or value of information	design	preposterior
Decision optimization	Clean the coil now or at the shutdown?	$\arg\min_a \mathbb{E}[L(a, \mathbf{z}) \mid \mathbf{y}]$	design	expected consequence, with the preposterior when a test comes first

must be inverted to bound. An *expectation* of a consequence, $\mathbb{E}[L \mid \mathbf{y}] = \int L(\mathbf{z})\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}$, is the mean of the marginal of the derived quantity $L(\mathbf{z})$: its category is estimation, applied to a loss, and its computation is the expected consequence, whose curvature term separates it from the loss at the most probable state; its tail $\Pr(L > l \mid \mathbf{y})$ would be a prediction under a threshold. An *importance* measure is a derivative of a derived quantity with respect to an entry of the state, the gradient that the delta method already uses, and it belongs with estimation. A *decision* among actions is a design question answered by the expected consequence, with the preposterior as well when a test comes before the action.

2.3.8 The twenty-one questions of the asset's life

The book is organized around twenty-one questions that come up in the life of an asset, grouped by the horizon they look over: what condition it is in today, how fast it is changing, when it will need attention, what should be done and when, and what the monitoring itself is worth. Table 2.5 lists them in the asset manager's words. Its first column is the number of the question; the second the question as the asset manager puts it; the third the general form it takes, by equation number; the fourth its category; the fifth the computation beneath it, every computation the question uses; and the last the chapter that leads with it.

Table 2.5: The twenty-one asset-management questions of the book, with the general form, the category and every computation each uses, and the chapter that leads with it. Chapters 14 to 24 are in Volume II.

No.	Asset manager's question	General form	Category	Computation	Chapter
What condition is it in today					
1	**Health index from physics.** How far has this coil's conductance fallen from its commissioned value?	(2.9), (2.10), the ratio $\mathrm{e}^{g_t - g_0}$ to the commissioned value	estimation	posterior	5
2	**Hidden degradation.** The unit still meets its set-point; is it working harder to do so?	(2.9), the duty as an entry against its commissioned value at the same inputs	estimation	posterior	6
3	**Fouling versus sensor.** The approach temperature has risen 1.5 K; is the exchanger fouled or is the sensor drifting?	(2.14)	hypothesis comparison	evidence	7
4	**Battery state of health.** What is the usable capacity now, and the internal resistance?	(2.9), (2.10), two entries and their correlation	estimation	posterior	8
5	**Fleet position.** Where does this unit sit among its fifty siblings?	(2.9), with a shared prior	estimation	posterior	9
How fast is it changing					
6	**Degradation rate.** How many percent of conductance per year is it losing?	(2.9) on the trend entry γ, with its band	estimation	trend	10
7	**Is the rate itself changing.** Is the loss accelerating?	(2.21), the sign of a curvature entry; (2.14)	prediction under a threshold; hypothesis comparison	trend; evidence	11
8	**Step change.** Did something happen in March?	(2.15) over candidate windows; (2.9) at the winner	diagnosis; estimation	evidence; posterior	12
9	**Seasonal confound.** Is the summer drop in performance degradation or just the ambient?	(2.21), the ambient a known input of the physics rows	prediction under a threshold	trend	13

No.	Asset manager's question	General form	Category	Computation	Chapter
When will it need attention					
10	**Remaining useful life.** When does the conductance cross the level at which the hall can no longer be held?	(2.21), second line: median and band by inversion	estimation	trend	14
11	**Probability of surviving to the next outage window.** Will it hold until the planned shutdown in October?	(2.21), second line at $s = t_{\mathrm{sd}}$	prediction under a threshold	trend	14
12	**Failure probability in a mission window.** Given its rate today, what is the probability it fails in the next 90 days?	(2.13)	prediction under a threshold	posterior	15
13	**Age-dependent rate.** Does the failure rate rise with age, and where is this unit on that curve?	(2.21) over age, the sign of θ_a; (2.9) at the unit's age	prediction under a threshold; estimation	trend	16
What should I do and when					
14	**Repair or replace.** Is restoring this unit's conductance worth more than its replacement, given what the remaining life band says?	(2.23), (2.24)	design	trend; expected consequence	17
15	**Inspection interval.** How often should the exchanger be inspected so that the probability of an undetected crossing stays under 5 %?	(2.21), second line against the interval; (2.23)	design	trend; expected consequence	18
16	**Condition-based trigger.** At what measured value should the work order fire?	(2.23) over decision levels, with (2.13)	design	expected consequence	19
17	**Which unit first.** Fifty units, one crew; which one does the crew visit next?	(2.23), ranked over the units	design	trend; expected consequence	20
18	**Warranty claim.** Has this unit degraded faster than its specification allows, with enough confidence to make the claim?	(2.14); (2.23) for the claim	hypothesis comparison; design	trend; evidence; expected consequence	21
What is the monitoring itself worth					
19	**Does the trend log see the degradation.** Can the existing sensors resolve a 5 % per year loss in the time I have?	(2.16), U_{p} the acceptance indicator of the trend test	design	trend; preposterior	22
20	**Which sensor to add.** Which added measurement most tightens the remaining-life band?	(2.16), (2.17), U_{p} the width of the crossing band	design	trend; preposterior	23
21	**Sampling frequency.** Hourly, daily, or monthly?	(2.16), (2.17), autocorrelated reading rows	design	trend; preposterior	24

Three readings of the table carry through the book. Questions 1, 2, 4 and 9 are where the plant graph separates the parameter from the load and the weather: a trend on raw performance mixes aging with load and weather, and a trend on the parameter does not, because the physics rows have already accounted for them. Questions 6 to 13 are where time enters, and it enters as rows that tie one window's parameter to the next, so that the trend is more rows of the same fit and not a second tool. Questions 14 to 18 are decisions: the asset manager wants a date and a decision rather than a health index, and each of these questions is an expected consequence over the posterior that the earlier questions produce. Counted over the table, the trend is called on by thirteen questions and the expected consequence by five; together they serve fourteen of the twenty-one, every one of them among Questions 6 to 21. The posterior serves six questions, the evidence four and the preposterior three.

2.4 Worked examples of estimation, prediction and joint events

Three pairs of questions are often confused in asset management: the estimate of a quantity and the probability of a threshold event, the estimate of a work-order rate and the prediction of a season without a work order, two replacements in one year and the product of two marginal probabilities. In each pair the two questions are put to the same model and differ in the form put to it.

2.4.1 Estimation of the conductance of the hall

The asset manager holds the commissioning test of the hall, 23.2°C at 30 kW, and wants to know where the conductance of the air path stands against its design sheet (question A) and what will happen at the requested load of $Q_2 = 60\,\mathrm{kW}$ (question B). Step 2 of the workflow has extended the graph with the scenario node of Figure 2.1, the hall at full load, whose temperature T_2 obeys the balance of equation (2.2),

$$T_2 = T_c + \frac{Q_2}{G}, \qquad T_c = 18^\circ\mathrm{C}, \quad Q_2 = 60\,\mathrm{kW}. \tag{2.28}$$

The conductance enters the state as the log-conductance $g = \ln(G/G_{\mathrm{ref}})$ with $G_{\mathrm{ref}} = 6\,\mathrm{kW/K}$, the design sheet's value. Chapter 1 fitted the commissioning test, the reading row and the handbook prior on the hall's two-entry state, and that fit is summarized here as one row on the log-conductance, its commissioned value $\hat{g}_0$ with the standard deviation $\sigma_{\hat{g}_0}$ of its posterior,

$$r_0(\mathbf{z}) = \frac{g - \hat{g}_0}{\sigma_{\hat{g}_0}}, \qquad \hat{g}_0 = -0.013926, \quad \sigma_{\hat{g}_0} = 0.094787. \tag{2.29}$$

The two numbers are the fit of Chapter 1 in the limit of an exact balance, where the hall temperature is a function of g and the objective has the one entry g: $\hat{g}_0$ is its Newton minimum, and $\sigma_{\hat{g}_0}$ the inverse square root of its curvature there, 111.3013, which is the Gauss–Newton part 113.9354 plus the curvature of the reading row, −2.6341. (Chapter 1 prints its fit at $\sigma_p = 0.01\,\mathrm{K}$, $\hat{g} = -0.013916$ with $\sqrt{\Sigma_{gg}} = 0.09481$, and the full-load tail $\Phi(-0.6878) = 0.2458$; this chapter takes the exact-physics limit, and the two agree to the fourth decimal.) Summarizing the commissioning record as one Gaussian row is a modeling choice. It is what lets the extended graph be fitted without carrying the reading again, and it drops the one thing the Gaussian cannot hold, the skew that the nonlinear reading row gives the exact posterior of Chapter 1; Section 2.4.2 states what that costs at full load. The row is the only one that acts on g alone, and the fit below returns the posterior of the third entry with center $\hat{g}_0$ and spread $\sqrt{\Sigma_{33}} = \sigma_{\hat{g}_0}$.

Question A is an estimation question, of the general form (2.9). The quantity asked for is the conductance, a derived quantity of the log-conductance, $f(\mathbf{z}) = G = G_{\mathrm{ref}}\,\mathrm{e}^{g}$, with g the third entry of the state, read against the design sheet as the health index $h = G/G_{\mathrm{ref}} = \mathrm{e}^{g}$; both are increasing in that one entry. The answer is therefore read from three objects of the fit of the hall: the MAP state $\hat{\mathbf{z}}$, the Laplace covariance $\boldsymbol{\Sigma}$ and the gradient of f. The gradient has a single nonzero entry, $\mathbf{g} = (0,\ 0,\ G)^{\mathsf{T}}$, so the band $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g} = G^2\,\Sigma_{33}$ uses only the third diagonal entry of $\boldsymbol{\Sigma}$; and because f is increasing in one entry, its quantiles are the images of that entry's quantiles under the exponential. When the third entry is exactly normal, as it is under the row (2.29), G is exactly lognormal, with median $G_{\mathrm{ref}}\,\mathrm{e}^{\hat{z}_3}$, mean $G_{\mathrm{ref}}\,\mathrm{e}^{\hat{z}_3+\Sigma_{33}/2}$ and central 95 % interval $G_{\mathrm{ref}}\,\mathrm{e}^{\hat{z}_3 \pm 1.960\sqrt{\Sigma_{33}}}$, and these exact summaries are reported beside the Laplace band. The fit of the state of the extended graph is the building block.

The posterior of the whole state of the extended graph follows from the row (2.29) and the two balance rows as a fit. The state holds the hall temperature T_1 at the test load $Q_1 = 30\,\mathrm{kW}$,

the full-load temperature T_2 and the log-conductance,

$$\mathbf{z} = \begin{bmatrix} T_1 \\ T_2 \\ g \end{bmatrix}. \tag{2.30}$$

The rows are the energy balances of the two hall nodes, written in temperature form as in Chapter 1 and divided by a model tolerance σ_p in kelvin, and the commissioning fit, which enters as a row on g of the same form as a prior row:

$$r(\mathbf{z}) = \begin{bmatrix} (T_1 - T_c - (Q_1/G_{\mathrm{ref}})\,\mathrm{e}^{-g})/\sigma_p \\ (T_2 - T_c - (Q_2/G_{\mathrm{ref}})\,\mathrm{e}^{-g})/\sigma_p \\ (g - \hat{g}_0)/\sigma_{\hat{g}_0} \end{bmatrix}, \qquad \mathcal{L}(\mathbf{z}) = \tfrac{1}{2}\, r(\mathbf{z})^{\mathsf{T}} r(\mathbf{z}). \tag{2.31}$$

The exact balance of equation (2.28) is the limit $\sigma_p \to 0$. The Jacobian has one row per row of r and one column per entry of $\mathbf{z}$, so it is 3×3; each entry is the sensitivity of a row to an entry of the state, in standard deviations of the row per unit of the entry:

$$\mathbf{J} = \frac{\partial r}{\partial \mathbf{z}} = \begin{bmatrix} 1/\sigma_p & 0 & (Q_1/G_{\mathrm{ref}})\,\mathrm{e}^{-g}/\sigma_p \\ 0 & 1/\sigma_p & (Q_2/G_{\mathrm{ref}})\,\mathrm{e}^{-g}/\sigma_p \\ 0 & 0 & 1/\sigma_{\hat{g}_0} \end{bmatrix} = \begin{bmatrix} 100 & 0 & 507.01 \\ 0 & 100 & 1014.02 \\ 0 & 0 & 10.55 \end{bmatrix}$$

at the most probable state with $\sigma_p = 0.01\,\mathrm{K}$, in K^{-1} in the first two columns. The third column says that a unit of g moves the full-load balance twice as far as the test-load balance, because the full load is twice the test load. The normal equations $(\mathbf{J}^{\mathsf{T}}\mathbf{J})\,\Delta\mathbf{z} = -\mathbf{J}^{\mathsf{T}} r$ give the Gauss–Newton step. The table below lists the iterations from the starting state (23°C, 28°C, 0), the forward solution of both balances at the design conductance $G = G_{\mathrm{ref}}$: the columns are the iteration, the three entries and the objective.

iteration [–]	T_1 [°C]	T_2 [°C]	g [–]	$\mathcal{L}$ [–]
0	23.0000	28.0000	0	1.079×10^{-2}
1	23.0696	28.1393	−0.013926	5.931×10^{-3}
2	23.0701	28.1402	−0.013926	6×10^{-21}
3	23.0701	28.1402	−0.013926	2×10^{-26}

At the start both balances hold, so the objective 0.01079 is the commissioning row's alone, $\frac{1}{2}(0.013926/0.094787)^2$: a path built to its design sheet is 0.15 standard deviations from the commissioned value. The row on g is linear, so the first step puts g at $\hat{g}_0$; the objective falls at that step only to 0.005931, because the step is taken on the linearization of e^{-g}, and the balance residuals it leaves, -4.9×10^{-4} and $-9.7 \times 10^{-4}\,\mathrm{K}$, are magnified a hundredfold by the division by σ_p. The second step closes both balances. At the most probable state

$$\hat{\mathbf{z}} = \begin{bmatrix} 23.070\,^{\circ}\mathrm{C} \\ 28.140\,^{\circ}\mathrm{C} \\ -0.013926 \end{bmatrix}$$

every row is zero, $T_1 = 18 + 30/5.9170$ and $T_2 = 18 + 60/5.9170$, with $\hat{G}_0 = 6\,\mathrm{e}^{\hat{g}_0} = 5.9170\,\mathrm{kW/K}$.

For rows divided by their standard deviations, as here, the information matrix is $\mathbf{\Lambda} = \mathbf{J}^{\mathsf{T}}\mathbf{J}$. When the rows are left in their physical units, with Jacobian $\mathbf{J}_{\mathrm{ph}}$ and weights $\mathbf{W} = \mathrm{diag}(1/\sigma_\ell^2)$, it is $\mathbf{J}_{\mathrm{ph}}^{\mathsf{T}}\mathbf{W}\mathbf{J}_{\mathrm{ph}}$, the same matrix, since $\mathbf{J} = \mathbf{W}^{1/2}\mathbf{J}_{\mathrm{ph}}$. Its inverse is the Laplace covariance of the state. The curvature term $\sum_\ell r_\ell\, \partial^2 r_\ell / \partial \mathbf{z}^2$ that Newton's method would add vanishes here, because every row is zero at $\hat{\mathbf{z}}$, so

$$\mathbf{\Sigma} = (\mathbf{J}^{\mathsf{T}}\mathbf{J})^{-1} = \begin{bmatrix} 0.23106 & 0.46192 & -0.04555 \\ 0.46192 & 0.92394 & -0.09111 \\ -0.04555 & -0.09111 & 0.00898 \end{bmatrix} \qquad (\sigma_p = 0.01\,\mathrm{K}). \tag{2.32}$$

The ridge of the posterior. The eigenvalues of this $\boldsymbol{\Sigma}$ are 1.164, 1.0×10^{-4} and 7.7×10^{-7}: one direction carries almost all of the uncertainty, and the other two shrink with σ_p^2. In the limit $\sigma_p \to 0$ the covariance is a single outer product of the vector $\mathbf{v} = (-Q_1/\hat{G}_0,\ -Q_2/\hat{G}_0,\ 1)^{\mathsf{T}}$,

$$\boldsymbol{\Sigma} \to \sigma_{\hat{g}_0}^2\, \mathbf{v}\mathbf{v}^{\mathsf{T}} = \begin{bmatrix} 0.23096 & 0.46192 & -0.04555 \\ 0.46192 & 0.92384 & -0.09111 \\ -0.04555 & -0.09111 & 0.00898 \end{bmatrix}, \qquad \mathbf{v} = \begin{bmatrix} -5.0701 \\ -10.1402 \\ 1 \end{bmatrix}, \tag{2.33}$$

where $\mathbf{v}$ is the direction along which both balance rows stay zero (it is orthogonal to the first two rows of $\mathbf{J}$) and $\sigma_{\hat{g}_0}^2$ the variance the third row leaves along it. The three entries move as one along $\mathbf{v}$: a unit rise of g cools the hall by 5.0701 K at the test load and 10.1402 K at full load. The first diagonal entry, $0.23096\,\mathrm{K}^2$, a spread of 0.48058 K, is the variance of the hall temperature that the commissioning fit of Chapter 1 leaves, carried back to T_1 by the balance; the second, $0.92384\,\mathrm{K}^2$, is four times it, because the full-load column of $\mathbf{v}$ is twice the test-load one. The limit is reached quickly: $\sqrt{\Sigma_{22}}$ is 1.3870 K at $\sigma_p = 1$ K, 0.9664 K at 0.1 K, 0.96122 K at 0.01 K, 0.96117 K at 0.001 K and 0.961165 K at 10^{-4} K, and $\sqrt{\Sigma_{33}} = 0.094787$ at every σ_p. The answers below are read off the limit.

The fit supplies the objects that question A was reduced to. The log-conductance is the third entry of the state, $\hat{z}_3 = -0.013926$ with $\sqrt{\Sigma_{33}} = 0.094787$, and the conductance has the gradient and band

$$\mathbf{g} = \frac{\partial G}{\partial \mathbf{z}} = \begin{bmatrix} 0 \\ 0 \\ G \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 5.9170 \end{bmatrix} \text{kW/K}, \qquad \sqrt{\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}} = 5.9170 \times 0.094787 = 0.5609\,\text{kW/K},$$

so that the reported answer is $G = 5.917 \pm 0.561$ kW/K, 9.48 % of its value, with the central 95 % band from 4.818 to 7.016 kW/K. The exact lognormal summaries follow from the normal of the third entry and stand beside it: the median is $G_{\mathrm{ref}}\, \mathrm{e}^{\hat{z}_3} = 5.9170$ kW/K, the location above; the mean, by the moment-generating function of the normal, is $G_{\mathrm{ref}}\, \mathrm{e}^{\hat{z}_3 + \Sigma_{33}/2} = 5.9170\, \mathrm{e}^{0.004492} = 5.9437$ kW/K; the standard deviation is 0.5647 kW/K, a gap of 0.0038 from the band; and the central 95 % credible interval is $5.9170\, \mathrm{e}^{\pm 1.960 \times 0.094787}$, from 4.914 to 7.125 kW/K, whose ends lie 0.096 and 0.109 kW/K above those of the band.

Read against the design sheet, the health index has the gradient $(0,\ 0,\ h)^{\mathsf{T}}$ with $h = 0.9862$, and the reported answer is $h = 0.986 \pm 0.0935$, a commissioned value just below the design value of one, with the delta band from 0.8030 to 1.1694; under the row it is lognormal with median 0.9862, mean 0.9906 and central interval from 0.8190 to 1.1875. The probability that the hall was commissioned below its design sheet is the estimation form at $f^* = 1$,

$$\Pr(h < 1 \mid \mathbf{y}) = \Pr(g < 0 \mid \mathbf{y}) = \Phi\!\left(\frac{0 - \hat{z}_3}{\sqrt{\Sigma_{33}}}\right) = \Phi(0.1469) = 0.558,$$

and the normal of h itself gives 0.5588. The exact posterior of Chapter 1, which keeps the skew the row drops, gives 0.5232: with one reading of 0.5 K, whether the air path was built to its sheet is close to an even question by either route. Each of these is a summary of a distribution; the answer to question A is the distribution itself, equation (2.9) with $f(\mathbf{z}) = G$ or $f(\mathbf{z}) = h$, and any single number reported from it must say which summary it is.

2.4.2 The threshold probability of the hall at full load

Question B asks whether the hall will remain below $T_{\max} = 27.5$°C at the requested load, that is, for

$$\Pr(T_2 < 27.5^{\circ}\mathrm{C} \mid \mathbf{y}).$$

It is a prediction under a threshold, of the general form (2.12). The requirement is the upper limit $T_2 < T_{\max}$ on the second entry of the state, so the form applies with $f(\mathbf{z}) = -T_2 = -z_2$ and $f^* = -T_{\max}$; equivalently, the probability of the requirement is one minus that of the exceedance, $1 - \Pr(T_2 \geq T_{\max} \mid \mathbf{y})$, and the assurance rule is applied to the requirement. The asset enters through the scenario node: the event is a bound on an entry that no sensor reads, tied to the third entry by the full-load balance row. The tail is read from the fit. Under the Laplace approximation the second entry is normal with center $\hat{z}_2$ and variance Σ_{22}, the second diagonal entry of $\boldsymbol{\Sigma}$, so the probability is a normal tail in those two numbers, read off $\hat{\mathbf{z}}$ and the limit of equation (2.33), and this is the value reported:

$$\Pr(T_2 < T_{\max} \mid \mathbf{y}) \approx \Phi\Big(\frac{T_{\max} - \hat{z}_2}{\sqrt{\Sigma_{22}}}\Big) = \Phi\Big(\frac{27.5 - 28.140}{0.9612}\Big) = \Phi(-0.6661) = 0.253. \tag{2.34}$$

The tail of g under the row. A second value stands beside it, through the ridge. Along $\mathbf{v}$ the full-load temperature is a decreasing function of the third entry, $T_2 = T_c + (Q_2/G_{\mathrm{ref}})\,\mathrm{e}^{-g}$, so the bound on T_2 is a bound on g, at the conductance where the second row meets the limit,

$$\Pr(T_2 < 27.5^{\circ}\mathrm{C} \mid \mathbf{y}) = \Pr(G > G^* \mid \mathbf{y}), \qquad G^* = \frac{Q_2}{T_{\max} - T_c} = 6.3158\,\mathrm{kW/K}, \tag{2.35}$$

that is $g > \ln(G^*/G_{\mathrm{ref}}) = 0.051293$, a required health index $h^* = 1.0526$: the requested load needs the air path 5.3 % above its design sheet, 0.6881 standard deviations above the commissioned value. The third entry is normal under the row, so the event is a tail of g alone, the route Chapter 1 reports as its Laplace value:

$$\Pr(T_2 < 27.5^{\circ}\mathrm{C} \mid \mathbf{y}) = \Phi\Big(\frac{\hat{z}_3 - \ln(G^*/G_{\mathrm{ref}})}{\sqrt{\Sigma_{33}}}\Big) = \Phi\Big(\frac{-0.013926 - 0.051293}{0.094787}\Big), \tag{2.36}$$

$$\Pr(T_2 < 27.5^{\circ}\mathrm{C} \mid \mathbf{y}) = \Phi(-0.6881) = 0.246. \tag{2.37}$$

The normal tail of the entry and the tail of g under the row differ by 0.0070, because T_2 is not linear in g along the ridge. The probability that the hall stays below the limit at the requested load is 25.3 %, with 24.6 % by the tail of g beside it, and the probability that it exceeds the limit is 74.7 %, with 75.4 % by the tail of g, conditional on the commissioning row of equation (2.29), the operating conditions and the steady-state thermal model. The exact value is that of the exact posterior of Chapter 1, which keeps the skew the Gaussian row drops: 0.2815, an exceedance of 0.719. The row moves the answer by more than the two routes on it differ, and it moves it in the direction that makes the request look worse. Whether a 75 % probability of exceeding the limit is acceptable is not a probability question; it depends on the consequences of exceedance and the operating criteria that apply. The assurance rule of equation (2.12) states such a criterion as a number, applied to the probability that the requirement holds: at an assurance level $\kappa = 0.95$ the requirement $T_2 < 27.5^{\circ}\mathrm{C}$ is not accepted by any of the three values, 0.253, 0.246 and 0.2815.

Figure 2.3 shows the two questions side by side. In panel (a) the horizontal axis is the conductance G and the vertical axis its posterior density under the row; the shaded area to the right of the dashed line at G^* is the tail of g, 0.246. In panel (b) the horizontal axis is the full-load temperature T_2 and the vertical axis its posterior density; the solid curve is the density of T_2 that the lognormal G implies under the row, the dashed curve the normal of the entry of equation (2.34), whose area to the left of the limit is the reported 0.253, marked on the panel, and the shaded area under the solid curve to the left of the limit is the same 0.246. The two shaded areas are one probability, seen once on the conductance and once on the temperature: the event $T_2 < 27.5^{\circ}\mathrm{C}$ and the event $G > G^*$ are the same set of states.

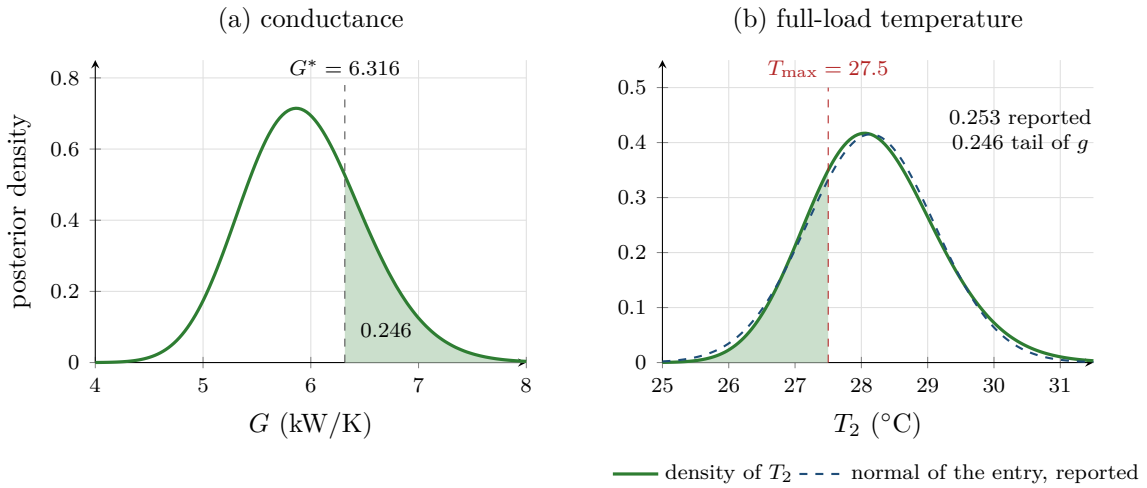


Figure 2.3: One posterior, two questions. (a) Lognormal posterior of the conductance under the commissioning row, density $(G\,\sigma_{\ln G}\sqrt{2\pi})^{-1}\exp[-(\ln G - \ln 5.9170)^2/(2\sigma^2_{\ln G})]$ with $\sigma_{\ln G} = \sqrt{\Sigma_{33}} = 0.094787$; shaded, $\Pr(G > 6.316\,\text{kW/K} \mid \mathbf{y}) = 0.246$. (b) Posterior of the full-load temperature $T_2 = 18 + 60/G$, lognormal in $T_2 - 18$ with median 10.140 K and log-spread 0.094787 (solid), and the normal of the entry, $\mathcal{N}(\hat{z}_2, \Sigma_{22}) = \mathcal{N}(28.140, 0.9612^2)$ (dashed), whose tail is the reported 0.253; shaded, the tail of g under the row, $\Pr(T_2 < 27.5°\text{C} \mid \mathbf{y}) = 0.246$.

The check against the tail of g and against draws. Equation (2.37) holds without approximation under the normal row on g, because the map from g to T_2 is monotone along the ridge; equation (2.34) is the route that every derived quantity has. Had T_2 not been carried in the state, the same band would come from the gradient of $f(\mathbf{z}) = T_c + (Q_2/G_{\text{ref}})\,e^{-g}$, $\mathbf{g} = (0,\ 0,\ -10.1402)^{\mathsf{T}}$ K, and it agrees with Σ_{22} because in the limit of equation (2.33) $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g} = \sigma^2_{\hat{g}_0}(\mathbf{g}^{\mathsf{T}}\mathbf{v})^2$ and $\mathbf{g}$ has the same projection on $\mathbf{v}$, -10.1402, as the unit vector on T_2. Table 2.6 compares the reported probability and the tail of g for the limit of the example and four others. The first column is the limit, the second the equivalent conductance threshold of equation (2.35), the third the health index it requires, the fourth the standardized distance in g, and the last two the reported normal of the entry, with mean $\hat{z}_2$ and variance Σ_{22}, and the tail of g under the row. The two differ by at most 0.0136 over these limits and by 0.0070 at 27.5°C, because the curvature of $60/G$ across a posterior of 9.5 % spread bends the right tail of T_2; panel (b) of Figure 2.3 shows the right skew the normal misses. Small as it is, the gap can cross a decision level: at $T_{\max} = 29.75$°C and $\kappa = 0.95$ the normal gives 0.9530, which accepts the requirement, and the tail of g gives 0.9400, which does not. Both values are then stated, and the asset manager sees that the requirement sits on the level. A second check draws 200 000 values of g from its posterior, computes T_2 for each, and counts the fraction below the limit: 0.2444, against the tail of g, 0.2457, within two sampling standard deviations of 0.0010.

What the two questions share. Both questions used the same posterior. Estimation, $p(G \mid \mathbf{y})$, describes the uncertainty about the conductance; the threshold probability, $\Pr(T_2 < T_{\max} \mid \mathbf{y})$, integrates that uncertainty over the states that satisfy the temperature limit. Neither required a new posterior; each asked a different question of the same one. That reuse is one of the principal advantages of a probabilistic model of the asset: the expensive step, the posterior, is paid once, and each new question is an integral over it.

Table 2.6: The threshold probability $\Pr(T_2 < T_{\max} \mid \mathbf{y})$ at $Q_2 = 60\,\mathrm{kW}$, $T_c = 18°\mathrm{C}$, $g \mid \mathbf{y} \sim \mathcal{N}(-0.013926, 0.094787^2)$: reported, by the normal $\mathcal{N}(\hat{z}_2, \Sigma_{22})$ of the entry T_2 (equation (2.34)), and the tail of g under the commissioning row (equation (2.36)).

$T_{\max}$ [°C]	$G^* = Q_2/(T_{\max} - T_c)$ [kW/K]	h^* [–]	$(\hat{z}_3 - \ln(G^*/G_{\mathrm{ref}}))/\sqrt{\Sigma_{33}}$ [–]	reported, normal of T_2 [–]	tail of g [–]
27.00	6.6667	1.1111	−1.2585	0.1178	0.1041
27.50	6.3158	1.0526	−0.6881	0.2527	0.2457
28.00	6.0000	1.0000	−0.1469	0.4420	0.4416
29.00	5.4545	0.9091	0.8586	0.8145	0.8047
29.75	5.1064	0.8511	1.5545	0.9530	0.9400

2.4.3 The posterior of the corrective work-order rate

A different record of the same asset: the work-order log of its coil design. Units of this design have logged $k = 2$ corrective work orders in $T_{\mathrm{tot}} = 10$ unit-years of service, and the asset manager asks for the probability that the hall's coil passes the coming summer season, 90 days, without one. This is a prediction question, not only a question about the work-order rate. Let λ be the unknown rate, in work orders per unit-year, and $\tau = 90$ days $= 0.2466\,\mathrm{yr}$ the length of the season. The work orders are taken to follow a homogeneous Poisson process, so that the count in the exposure is Poisson with mean λT_{tot},

$$K(T_{\mathrm{tot}}) \mid \lambda \sim \mathrm{Poisson}(\lambda T_{\mathrm{tot}}), \qquad \Pr(k \mid \lambda, T_{\mathrm{tot}}) = \frac{(\lambda T_{\mathrm{tot}})^k \, \mathrm{e}^{-\lambda T_{\mathrm{tot}}}}{k!}.$$

This is a modeling choice. It is what lets the log be summarized by k and T_{tot} alone, and it fails when the units of the design are not alike in age or duty, or when the rate rises with age (Chapter 16, in Volume II), or when one fault raises several work orders, a repeat visit for the same leak; the count here is of faults, with repeat visits merged into the first. As a function of λ, the likelihood is proportional to $\lambda^k \, \mathrm{e}^{-\lambda T_{\mathrm{tot}}}$; the factor $T_{\mathrm{tot}}^k/k!$ does not depend on the rate. The prior is a gamma distribution in the shape–rate parameterization, $\lambda \sim \mathrm{Gamma}(a, b)$, with density proportional to $\lambda^{a-1} \, \mathrm{e}^{-b\lambda}$ and, here,

$$a = 2, \qquad b = 5\,\mathrm{yr},$$

whose prior mean $a/b = 0.4$ work orders a year, an MTBF of 2.50 years, stands for the maintainer's records of the design. (The book writes the gamma parameters as a and b; the context tells the shape a from an action a.) Bayes' rule gives the posterior up to a constant as likelihood times prior, $p(\lambda \mid k, T_{\mathrm{tot}}) \propto \Pr(k \mid \lambda, T_{\mathrm{tot}}) \, p(\lambda)$.

The rate enters the state, as every rate does in this book, through its log-rate $\omega = \ln(\lambda/\lambda_{\mathrm{ref}})$ with $\lambda_{\mathrm{ref}} = 0.1\,\mathrm{yr}^{-1}$, so that the state is the one number

$$\omega = \ln \frac{\lambda}{\lambda_{\mathrm{ref}}}, \qquad \lambda = \lambda_{\mathrm{ref}} \, \mathrm{e}^{\omega}.$$

In the general forms $\mathbf{z}$ is here the one number ω, and every quantity of the fit is a number. The rate is asked for first, an estimation question of the general form (2.9) with $f(\omega) = \lambda = \lambda_{\mathrm{ref}} \, \mathrm{e}^{\omega}$, increasing in ω. It is read from the MAP log-rate $\hat{\omega}$, its variance σ_ω^2 and the derivative $\mathrm{d}\lambda/\mathrm{d}\omega = \hat{\lambda}$, whose band is $\hat{\lambda} \, \sigma_\omega$; the exact gamma posterior that the objective identifies is set beside it. Neither row is Gaussian, and each enters the objective as its negative log-probability term. The count row is the negative log of the Poisson probability; the prior row is the negative log of the gamma density carried over to ω, where the change of variable multiplies the density by

$\mathrm{d}\lambda/\mathrm{d}\omega = \lambda$ and so adds $-\ln\lambda$ (the rule, that a density carried to a new coordinate is multiplied by the derivative of the old coordinate with respect to the new, is derived in Section 4.3.3):

$$\mathcal{L}(\omega) = \underbrace{T_{\mathrm{tot}}\lambda - k\ln\lambda}_{\text{count row, } \mathcal{L}_c} + \underbrace{b\lambda - a\ln\lambda}_{\text{prior row, } \mathcal{L}_0} + \text{const.} \tag{2.38}$$

With primes for derivatives in ω and $\mathrm{d}\lambda/\mathrm{d}\omega = \lambda$, each row has a slope and a curvature, and they add:

$$\mathcal{L}_c' = T_{\mathrm{tot}}\lambda - k, \quad \mathcal{L}_c'' = T_{\mathrm{tot}}\lambda, \qquad \mathcal{L}_0' = b\lambda - a, \quad \mathcal{L}_0'' = b\lambda, \qquad \mathcal{L}' = \mathcal{L}_c' + \mathcal{L}_0', \quad \mathcal{L}'' = \mathcal{L}_c'' + \mathcal{L}_0''.$$

The Newton step $\Delta\omega = -\mathcal{L}'/\mathcal{L}''$ moves to the minimum of the parabola with the current slope and curvature. The table below lists the iterations from the maximum likelihood rate $k/T_{\mathrm{tot}} = 0.2\,\mathrm{yr}^{-1}$, that is $\omega = \ln 2$, where the count row's slope is zero: the columns are the iteration, the log-rate, the rate, the slopes of the two rows, the slope and curvature of the objective, and the step.

iteration j [–]	ω_j [–]	λ [yr^{-1}]	$\mathcal{L}_c'$ [–]	$\mathcal{L}_0'$ [–]	$\mathcal{L}'$ [–]	$\mathcal{L}''$ [–]	$\Delta\omega$ [–]
0	0.69315	0.200000	0.0000	−1.0000	−1.0000	3.0000	0.333
1	1.02648	0.279122	0.7912	−0.6044	0.1868	4.1868	−0.0446
2	0.98186	0.266940	0.6694	−0.6653	0.004107	4.0041	−0.00103
3	0.98083	0.266667	0.6667	−0.6667	2×10^{-6}	4.0000	-5×10^{-7}

The slope vanishes at $\lambda = (a + k)/(b + T_{\mathrm{tot}})$, so the mode is $\hat{\omega} = \ln(8/3) = 0.98083$ and $\lambda_{\mathrm{ref}}\,\mathrm{e}^{\hat{\omega}} = 4/(15\,\mathrm{yr}) = 0.266667\,\mathrm{yr}^{-1}$. The curvature there is the information, $\Lambda = \mathcal{L}''(\hat{\omega}) = T_{\mathrm{tot}}\hat{\lambda} + b\hat{\lambda} = 2.6667 + 1.3333 = 4 = a + k$: the count row contributes the expected number of work orders in the exposure, the prior row $b\hat{\lambda}$, and the two add as the information of independent rows does. Then $\sigma_\omega^2 = 1/\Lambda = 0.25$, the log-rate has $\sigma_\omega = 0.5$, and the rate, a derived quantity with derivative $\mathrm{d}\lambda/\mathrm{d}\omega = \hat{\lambda}$, has $\hat{\lambda}\,\sigma_\omega = 0.266667 \times 0.5 = 0.133333\,\mathrm{yr}^{-1}$. The reported answer is the work-order rate 0.267 ± 0.133 a year, an MTBF of 3.75 years at its center.

The objective also identifies the exact posterior. Back in λ, without the $-\ln\lambda$ of the change of variable, the negative log posterior is $(b + T_{\mathrm{tot}})\lambda - (a + k - 1)\ln\lambda$, the negative log of the kernel $\lambda^{a+k-1}\,\mathrm{e}^{-(b+T_{\mathrm{tot}})\lambda}$ of a gamma density. Therefore

$$\lambda \mid k, T_{\mathrm{tot}} \sim \mathrm{Gamma}(a', b'), \qquad a' = a + k, \quad b' = b + T_{\mathrm{tot}}. \tag{2.39}$$

For the coil design, $a' = 2 + 2 = 4$ and $b' = 5 + 10 = 15\,\mathrm{yr}$, so $\lambda \mid \mathbf{y} \sim \mathrm{Gamma}(4, 15\,\mathrm{yr})$, with posterior mean and standard deviation

$$\mathbb{E}[\lambda \mid \mathbf{y}] = \frac{a'}{b'} = \frac{4}{15\,\mathrm{yr}} = 0.266667\,\mathrm{yr}^{-1}, \qquad \sqrt{\mathrm{Var}[\lambda \mid \mathbf{y}]} = \frac{\sqrt{a'}}{b'} = 0.133333\,\mathrm{yr}^{-1}.$$

Set like with like, the reported posterior and the exact one differ, because the Laplace normal of the log-rate makes the rate lognormal: its median is $0.266667\,\mathrm{yr}^{-1}$ against the gamma's $0.244804\,\mathrm{yr}^{-1}$, a gap of $0.021863\,\mathrm{yr}^{-1}$; its mean $\hat{\lambda}\,\mathrm{e}^{\sigma_\omega^2/2} = 0.302173\,\mathrm{yr}^{-1}$ against the gamma's $0.266667\,\mathrm{yr}^{-1}$, a gap of $0.035506\,\mathrm{yr}^{-1}$; and its central 95 % interval runs from 0.100083 to $0.710522\,\mathrm{yr}^{-1}$ against the gamma's 0.072658 to $0.584485\,\mathrm{yr}^{-1}$. The pair the Newton fit returned coincides with the gamma's mean and standard deviation: the mode of the log-rate maps onto the gamma mean and the linearized band onto the gamma standard deviation, for every gamma posterior, but that compares different summaries, a median of one posterior with a mean of the other. The gamma mean, an MTBF of 3.75 years, lies between the prior mean $a/b = 0.4\,\mathrm{yr}^{-1}$, an MTBF of 2.50 years, and the maximum likelihood estimate $k/T_{\mathrm{tot}} = 0.2\,\mathrm{yr}^{-1}$, an MTBF of 5.00 years, as the weighted form $a'/b' = \frac{b}{b'}\,\frac{a}{b} + \frac{T_{\mathrm{tot}}}{b'}\,\frac{k}{T_{\mathrm{tot}}} = \frac{1}{3}(0.4) + \frac{2}{3}(0.2)$ shows. An estimation question has been answered. The asset manager's question, a season without a work order, has not.

2.4.4 The probability of a summer season without a work order

Let T be the time to the coil's next corrective work order, counted from the start of the season. (In this example T, in years, is a time; in the hall example, T_1 and T_2 are hall temperatures in °C. The units distinguish them.) Under the Poisson process the time to the next work order does not depend on how long ago the last one was, so the question is

$$\Pr(T > \tau \mid \mathbf{y}), \qquad \tau = 90 \text{ days} = 0.2466\,\text{yr}. \tag{2.40}$$

For a known constant rate the probability of no work order is exponential, $\Pr(T > \tau \mid \lambda) = \mathrm{e}^{-\lambda\tau}$. The rate is not known, and the event is a future one that stays random even when the rate is known, so the question is a prediction of the general form (2.13), in its event form (2.8): y_{new} is the event $T > \tau$, and $p(y_{\text{new}} \mid \mathbf{z}) = \mathrm{e}^{-\lambda\tau}$ with $\lambda = \lambda_{\text{ref}}\,\mathrm{e}^{\omega}$. The integral is a function of the one number ω averaged over its posterior, and under the Laplace approximation that posterior is $\mathcal{N}(\hat{\omega}, \sigma_\omega^2)$, so the value read from the state is a normal expectation in $\hat{\omega}$ and σ_ω^2:

$$\Pr(T > \tau \mid \mathbf{y}) \approx \mathbb{E}_{\omega\sim\mathcal{N}(\hat{\omega},\sigma_\omega^2)}\big[\mathrm{e}^{-\lambda_{\text{ref}}\mathrm{e}^{\omega}\tau}\big] \approx \sum_{j=1}^{5} w_j\,\mathrm{e}^{-\lambda_{\text{ref}}\mathrm{e}^{\omega_j}\tau}, \qquad \omega_j = \hat{\omega} + \sqrt{2}\,\sigma_\omega\,x_j. \tag{2.41}$$

The sum is the five-point Gauss–Hermite rule, which replaces a normal expectation by a weighted sum at fixed nodes x_j with weights w_j and is exact for a polynomial of degree nine. The table below evaluates it for the season, with $\hat{\omega} = 0.98083$ and $\sigma_\omega = 0.5$: the columns are the node, its weight, the log-rate and rate there, the probability of no work order at that rate, and the weighted term.

x_j [–]	w_j [–]	ω_j [–]	λ_j [yr^{-1}]	$\mathrm{e}^{-\tau\lambda_j}$ [–]	term [–]
−2.02018	0.01126	−0.44766	0.063912	0.98436	0.01108
−0.95857	0.22208	0.30302	0.135394	0.96717	0.21478
0	0.53333	0.98083	0.266667	0.93636	0.49939
0.95857	0.22208	1.65864	0.525218	0.87853	0.19510
2.02018	0.01126	2.40931	1.112633	0.76007	0.00856

The terms sum to 0.9289, and a direct numerical integration agrees to six decimals. The predicted probability that the coil passes the summer season without a corrective work order is 92.89 %, and the probability of at least one is 0.0711; these are the values reported.

The exact value stands beside them, because the objective identified the posterior as a gamma and the same integral then has a closed form. The conditional probability of a season without a work order is averaged over the posterior of the rate,

$$\Pr(T > \tau \mid \mathbf{y}) = \int_0^\infty \mathrm{e}^{-\lambda\tau}\,p(\lambda \mid \mathbf{y})\,\mathrm{d}\lambda.$$

Substitute the density of the gamma posterior of equation (2.39),

$$p(\lambda \mid \mathbf{y}) = \frac{b'^{\,a'}}{\Gamma(a')}\,\lambda^{a'-1}\,\mathrm{e}^{-b'\lambda},$$

and combine the two exponentials:

$$\Pr(T > \tau \mid \mathbf{y}) = \frac{b'^{\,a'}}{\Gamma(a')}\int_0^\infty \lambda^{a'-1}\,\mathrm{e}^{-(b'+\tau)\lambda}\,\mathrm{d}\lambda.$$

The remaining integral is the standard gamma integral, $\int_0^\infty \lambda^{a'-1}\mathrm{e}^{-\ell\lambda}\,\mathrm{d}\lambda = \Gamma(a')/\ell^{a'}$, obtained from the definition of the gamma function by the substitution $x = \ell\lambda$; with $\ell = b' + \tau$ the gamma functions cancel, and

$$\Pr(T > \tau \mid \mathbf{y}) = \left(\frac{b'}{b'+\tau}\right)^{a'}. \tag{2.42}$$

For the season,

$$\Pr(T > 0.2466\,\mathrm{yr} \mid \mathbf{y}) = \left(\frac{15}{15.2466}\right)^4 = 0.9369,$$

and the probability of at least one work order is 0.0631. The calculation needed no Gaussian or Laplace approximation: with a gamma posterior and an exponential probability of no work order, the predictive probability has an exact closed form. The gap between the two is 0.0079 for the season, and over a full year the state gives 0.7481 against the exact 0.7725, a gap of 0.0244. The state route errs low because the lognormal it implies for λ has the larger mean $0.302173\,\mathrm{yr}^{-1}$ and a heavier upper tail than the gamma. No decision of this chapter sits between the two values for the season; an assurance level of 0.93 would, and both would then be stated, the one read from the state failing it and the exact one meeting it.

2.4.5 The error of the plug-in MTBF

A common shortcut substitutes the posterior mean rate, the reciprocal of the MTBF, into the exponential:

$$\Pr_{\text{plug}}(T > 0.2466\,\mathrm{yr}) = \mathrm{e}^{-\mathbb{E}[\lambda|\mathbf{y}]\times 0.2466\,\mathrm{yr}} = \mathrm{e}^{-0.0658} = 0.9364,$$

against the exact predictive 0.9369; over a full year it gives $\mathrm{e}^{-0.2667} = 0.7659$ against 0.7725. The two differ because averaging and exponentiation do not commute,

$$\mathbb{E}\big[\mathrm{e}^{-\lambda\tau} \mid \mathbf{y}\big] \neq \mathrm{e}^{-\mathbb{E}[\lambda|\mathbf{y}]\,\tau} \quad \text{in general,} \tag{2.43}$$

since the probability of no work order is a nonlinear function of the rate. The direction of the difference follows from the shape of that function. For $\tau > 0$ the map $\lambda \mapsto \mathrm{e}^{-\lambda\tau}$ has second derivative $\tau^2\mathrm{e}^{-\lambda\tau} > 0$, so it is convex, and Jensen's inequality, the standard result that the expectation of a convex function is at least the function of the expectation, gives $\mathbb{E}[\mathrm{e}^{-\lambda\tau} \mid \mathbf{y}] \geq \mathrm{e}^{-\mathbb{E}[\lambda|\mathbf{y}]\tau}$: the plug-in value understates the exact probability. The low rates the posterior admits raise that probability by more than the high rates lower it. In the log-rate the same function is concave while $\lambda\tau < 1$, so the reported average (2.41) starts below the plug-in and crosses it at 3.56 years. For one season the plug-in sits 0.0005 from the exact value, closer than the reported 0.9289. The plug-in value is wrong not because it is far from the predictive value, but because it answers a different question: the probability of a quiet season for a coil whose rate is known to be $0.266667\,\mathrm{yr}^{-1}$, which is not the coil in question, and it carries no band.

Table 2.7 shows how the values change with the horizon. The first column is the horizon; the second the reported average of equation (2.41); the third the exact predictive probability of no work order, equation (2.42); the fourth their gap; the fifth the plug-in value at the posterior mean; the sixth the exact value minus the plug-in; and the last the plug-in value at the maximum likelihood estimate $0.2\,\mathrm{yr}^{-1}$, which ignores the prior as well as the uncertainty. For one season the exact and plug-in values differ by 0.0005. Their ratio grows with the horizon, from 1.0005 for the season to 1.87 at 10 years and 7.0 at 20 years; the absolute difference peaks near 8.08 years, at 0.0625, and then falls, to 0.0289 at 20 years, because both values tend to zero; the gap of the reported average peaks at the middle rows, at 0.0387 among the rows of the table. Figure 2.4 draws the ingredients. In panel (a) the horizontal axis is the work-order rate in work orders per unit-year and the vertical axis the probability density per unit of that axis; the blue curve is the prior, the green curve the posterior, and the dotted lines mark the maximum likelihood estimate, the posterior mean and the prior mean. In panel (b) the horizontal axis is the horizon and the vertical axis the probability of no work order; the green curve is the exact predictive, the blue dash-dotted curve the reported average over the normal of the log-rate and the dashed red curve the plug-in, and the vertical dotted lines mark the 90-day season and one year. The exact curve lies above the plug-in curve at every horizon, and their ratio grows as the

horizon lengthens, because the longer the horizon, the more the probability of no work order bends across the spread of the posterior; the blue curve falls below both at first and crosses the plug-in at 3.56 years.

Table 2.7: Probability that the coil passes a horizon τ without a corrective work order: the reported average over the normal of the log-rate (equation (2.41), $\hat{\omega} = \ln(8/3)$, $\sigma_\omega = 0.5$); exact, from the posterior Gamma$(4, 15\,\text{yr})$, $(15/(15+\tau))^4$; plug-in $e^{-4\tau/15}$; plug-in at the maximum likelihood estimate $e^{-0.2\tau}$, with τ in years. Gaps are computed before rounding.

τ [yr]	reported, from the state [–]	exact [–]	gap [–]	plug-in, posterior mean [–]	exact minus plug-in [–]	plug-in, maximum likelihood [–]
0.2466	0.9289	0.9369	0.0079	0.9364	0.0005	0.9519
1	0.7481	0.7725	0.0244	0.7659	0.0065	0.8187
5	0.2777	0.3164	0.0387	0.2636	0.0528	0.3679
10	0.1015	0.1296	0.0281	0.0695	0.0601	0.1353
20	0.0205	0.0337	0.0132	0.0048	0.0289	0.0183

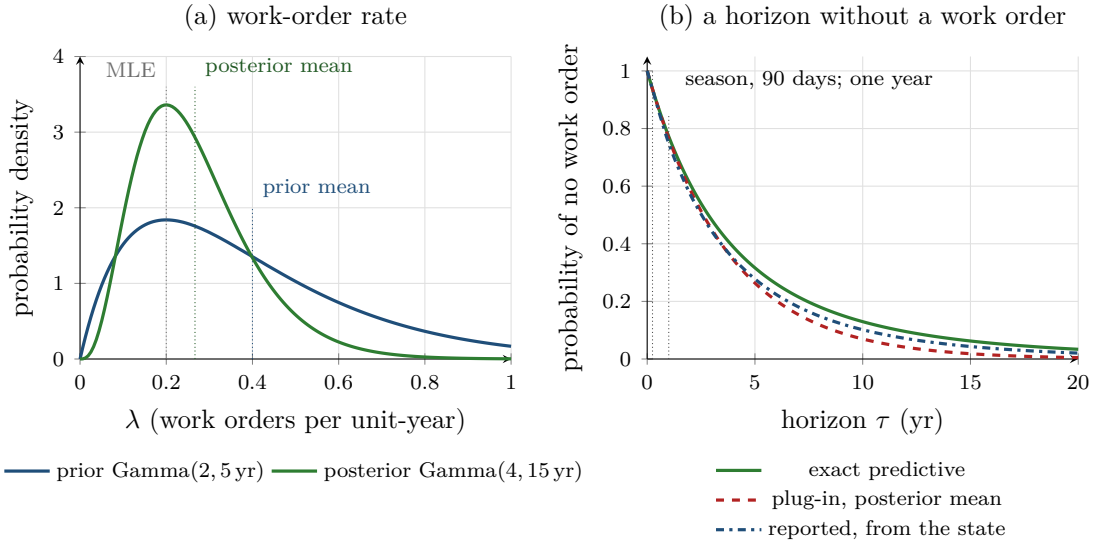


Figure 2.4: The work-order rate of the coil design and the probability of a horizon without a corrective work order. (a) Prior $25\,\lambda\,e^{-5\lambda}$ and posterior $(15^4/3!)\,\lambda^3 e^{-15\lambda}$; dotted lines at the maximum likelihood estimate 0.2, the posterior mean 0.2667 and the prior mean 0.4 work orders per unit-year. (b) Exact predictive $(15/(15+\tau))^4$ of equation (2.42), plug-in $e^{-4\tau/15}$ and the reported average of equation (2.41); for the 90-day season they are 0.9369, 0.9364 and 0.9289, at one year 0.7725, 0.7659 and 0.7481.

Principle 2.2. *Predicting a season without a work order requires averaging the probability of no work order over the uncertainty in the rate. Substituting a point estimate of the rate, the plug-in MTBF, answers the question for a different coil, one whose rate is known.*

2.4.6 Two replacements in one year as a joint event

The hall's cooling unit has a fan motor and a valve actuator, and the year's maintenance budget covers one replacement but not two. Figure 2.5 draws the arrangement: the white box on the

left is the unit's control-power supply, which both components draw on; the two green boxes are the fan motor and the actuator, each with its own wear; the gray box on the right is the cooling unit in service, which needs both; the solid arrows are the paths along which power and service pass. Define the events

$$F_A = \{\text{fan motor replaced this year}\}, \qquad F_B = \{\text{valve actuator replaced this year}\}.$$

The budget fails only if both need replacing, so the event the asset manager fears is the intersection

$$F_{\text{both}} = F_A \cap F_B, \qquad \Pr(F_{\text{both}} \mid \mathbf{y}) = \Pr(F_A \cap F_B \mid \mathbf{y}).$$

It is the general form (2.12) with the indicator of an intersection in place of the indicator of one bound, here with the marginal probabilities given directly rather than computed from a posterior, so there is no state route to report and no gap to state; every probability of this example is conditional on the same evidence, and the bar is left out below to keep the products readable. Suppose $\Pr(F_A) = 0.05$ and $\Pr(F_B) = 0.03$. If the two replacements are independent, the probability of the intersection is the product of the marginal probabilities,

$$\Pr(F_A \cap F_B) = \Pr(F_A)\,\Pr(F_B) = 0.05 \times 0.03 = 0.0015,$$

and the probability that the year needs at most one of the two replacements, the probability that the budget holds, is the probability of the complement,

$$\Pr(\text{at most one} \mid \mathbf{y}) = 1 - 0.0015 = 0.9985, \tag{2.44}$$

or 99.85 %; the probability that neither is needed is $0.95 \times 0.97 = 0.9215$.

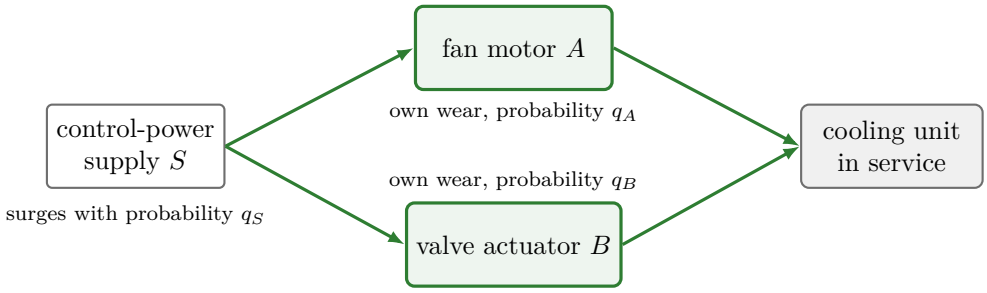


Figure 2.5: The fan motor and the valve actuator of the hall's cooling unit on one control-power supply. A surge on the supply (probability q_S in a year) damages both at once; the components' own wear (probabilities q_A, q_B) is independent given no surge. The year's budget covers one replacement; the event of concern is that both are needed. With $q_S = 0$ the two replacements are independent.

2.4.7 A shared cause of the two replacements

Both components draw on the same control-power supply. Then independence may not hold, and the probability of the intersection in the form (2.12) must be written with the product rule, which holds without any independence assumption:

$$\Pr(F_A \cap F_B) = \Pr(F_A)\,\Pr(F_B \mid F_A). \tag{2.45}$$

The conditional probability $\Pr(F_B \mid F_A)$ may be much larger than the marginal $\Pr(F_B)$: a failed fan motor raises the probability that the supply has surged, and a surge takes the actuator with it. A shared cause can make the two replacements strongly dependent.

An illustrative model shows by how much. Let the supply surge during the year with probability q_S, which damages both components, and let each component otherwise wear out by its own mechanisms, independently, with probabilities q_A and q_B. The law of total probability over the two states of the supply, the integral of the form over a state with two values, gives the marginal and the joint probabilities, exactly, with no second route beside them:

$$\Pr(F_A) = q_S + (1 - q_S)\, q_A, \qquad \Pr(F_B) = q_S + (1 - q_S)\, q_B,$$
$$\Pr(F_A \cap F_B) = q_S \cdot 1 + (1 - q_S)\, q_A q_B. \tag{2.46}$$

Holding the marginals at 0.05 and 0.03, the own-wear probabilities are $q_A = (0.05 - q_S)/(1 - q_S)$ and $q_B = (0.03 - q_S)/(1 - q_S)$. For $q_S = 0.01$, one surge in a hundred years, $q_A = 0.04/0.99 = 0.040404$ and $q_B = 0.02/0.99 = 0.020202$, so

$$\Pr(F_A \cap F_B) = 0.01 + 0.99 \times 0.040404 \times 0.020202 = 0.01 + 0.000808 = 0.010808,$$

and equation (2.45) gives $\Pr(F_B \mid F_A) = 0.010808/0.05 = 0.216$, seven times the marginal 0.03. Table 2.8 repeats the calculation for five surge probabilities. The first column is q_S; the next two the own-wear probabilities that keep the marginals fixed; then the probability that both are needed, the conditional probability of the second replacement given the first, the ratio of the joint probability to the independent value 0.0015, and the probability $\Pr(\text{at most one} \mid \mathbf{y})$ that the budget holds. With the marginals unchanged, a surge probability of 0.01 multiplies the probability that the budget fails by 7.2, and one of 0.02 by 13.5: the supply, not the components' own wear, then governs it.

Table 2.8: Two replacements in one year with a shared control-power supply, the marginals held at $\Pr(F_A) = 0.05$ and $\Pr(F_B) = 0.03$ (illustrative model of equation (2.46)).

q_S [–]	q_A [–]	q_B [–]	$\Pr(F_A \cap F_B)$ [–]	$\Pr(F_B \mid F_A)$ [–]	ratio to 0.0015 [–]	at most one [–]
0	0.05000	0.03000	0.001500	0.0300	1.00	0.99850
0.001	0.04905	0.02903	0.002422	0.0484	1.61	0.99758
0.005	0.04523	0.02513	0.006131	0.1226	4.09	0.99387
0.01	0.04040	0.02020	0.010808	0.2162	7.21	0.98919
0.02	0.03061	0.01020	0.020306	0.4061	13.54	0.97969

The example shows why the wording "both need replacing" matters: it names an intersection of events, and an intersection directs the formulation to joint probability and to the modeling of dependence, not to a product of marginals. A unit can also fail its duty through several pathways, and in larger systems the event of concern is a union of several overlapping events, and a shared cause is then carried as a shared entry of the state.

2.5 Model comparison, diagnosis and decisions

The questions so far were asked of the state of one model. A question about which model holds or which fault is present is put to several models at once, and its answer is a ratio or a ranking of evidences; a question about which action to take is put to several actions, and its answer is a comparison of expected consequences.

2.5.1 Posterior probabilities of competing models

Consider an air path whose condition tests read warmer than its commissioned model expects, and two hypotheses about it,

$$M_{\text{clean}} : \text{the air path is clean}, \qquad M_{\text{fouled}} : \text{the air path is fouled}.$$

The question is which explanation the evidence supports better. This is not a threshold probability on a fixed, known model: the asset manager is comparing models. It is a hypothesis comparison, of the general form (2.14) with $M_A = M_{\text{fouled}}$ and $M_B = M_{\text{clean}}$. The asset enters through the two drawings: each is the graph of the air path with its own rows, and M_{fouled} carries an entry for the fouling, a reduction of the conductance of the path, that M_{clean} pins at zero. Each evidence is the integral (2.26) over its own model's state, read under the Laplace approximation from that model's own fit, its MAP state $\hat{\mathbf{z}}_j$ and covariance $\boldsymbol{\Sigma}_j$. The extra entry of M_{fouled} is not rewarded merely for fitting the readings better: the evidence multiplies the height of the fit by the posterior volume $\det(\boldsymbol{\Sigma}_j)^{1/2}$, and the prior density $p(\hat{\mathbf{z}}_j \mid M_j)$ in that height is low when the prior spreads the entry over a wide range. Treat the model index as an unknown taking the values $M_{\text{clean}}, M_{\text{fouled}}, \ldots$, with prior probabilities $\Pr(M_j)$. Bayes' rule for a discrete unknown, with the law of total probability in the denominator, gives

$$\Pr(M_j \mid \mathbf{y}) = \frac{p(\mathbf{y} \mid M_j)\,\Pr(M_j)}{\sum_\ell p(\mathbf{y} \mid M_\ell)\,\Pr(M_\ell)}, \tag{2.47}$$

where $\Pr(M_j)$ is the prior probability of model M_j, $\Pr(M_j \mid \mathbf{y})$ its posterior probability, and $p(\mathbf{y} \mid M_j)$ its evidence, equation (2.26). Dividing equation (2.47) for M_{fouled} by the same equation for M_{clean} cancels the denominator and gives the odds form,

$$\frac{\Pr(M_{\text{fouled}} \mid \mathbf{y})}{\Pr(M_{\text{clean}} \mid \mathbf{y})} = B_{\text{fc}}\,\frac{\Pr(M_{\text{fouled}})}{\Pr(M_{\text{clean}})}, \qquad B_{\text{fc}} = \frac{p(\mathbf{y} \mid M_{\text{fouled}})}{p(\mathbf{y} \mid M_{\text{clean}})}\,: \tag{2.48}$$

posterior odds equal the Bayes factor (2.14) times the prior odds.

2.5.2 A numerical comparison of two explanations

Suppose the computed evidences are $p(\mathbf{y} \mid M_{\text{clean}}) = 0.01$ and $p(\mathbf{y} \mid M_{\text{fouled}}) = 0.04$, and the prior probabilities are $\Pr(M_{\text{clean}}) = 0.8$ and $\Pr(M_{\text{fouled}}) = 0.2$, illustrative values. Equation (2.47) gives

$$\Pr(M_{\text{fouled}} \mid \mathbf{y}) = \frac{0.04 \times 0.2}{0.04 \times 0.2 + 0.01 \times 0.8} = \frac{0.008}{0.016} = 0.50.$$

The Bayes factor is

$$B_{\text{fc}} = \frac{p(\mathbf{y} \mid M_{\text{fouled}})}{p(\mathbf{y} \mid M_{\text{clean}})} = \frac{0.04}{0.01} = 4, \qquad \ln B_{\text{fc}} = \ln 4 = 1.386 \text{ nats},$$

and the odds form checks the result: posterior odds $4 \times (0.2/0.8) = 1$, a posterior probability of $1/(1+1) = 0.50$. The evidence is four times more probable under the fouling model, yet its posterior probability is only 50 %, because the prior probability assigned to fouling was low. The Bayes factor expresses the relative support that the evidence provides; it is not by itself a posterior model probability, and turning it into one requires the prior probabilities of the models.

Figure 2.6 makes the distinction visible. In panel (a) the two pairs of bars are the prior and the posterior probabilities of the two models: the evidence moves probability from M_{clean} to M_{fouled} until they are equal. In panel (b) the horizontal axis is the prior probability of the fouling model and the vertical axis its posterior probability; the green curve is $\Pr(M_{\text{fouled}} \mid \mathbf{y}) = 4\,\Pr(M_{\text{fouled}})/(3\,\Pr(M_{\text{fouled}}) + 1)$, which equation (2.48) gives for $B_{\text{fc}} = 4$, and the dashed diagonal is the case $B_{\text{fc}} = 1$, evidence that does not discriminate. The same Bayes factor of 4 turns a prior of 0.05 into 0.174, a prior of 0.2 into 0.50 (the marked point) and a prior of 0.5 into 0.80: the factor is fixed by the evidence, the conclusion is not.

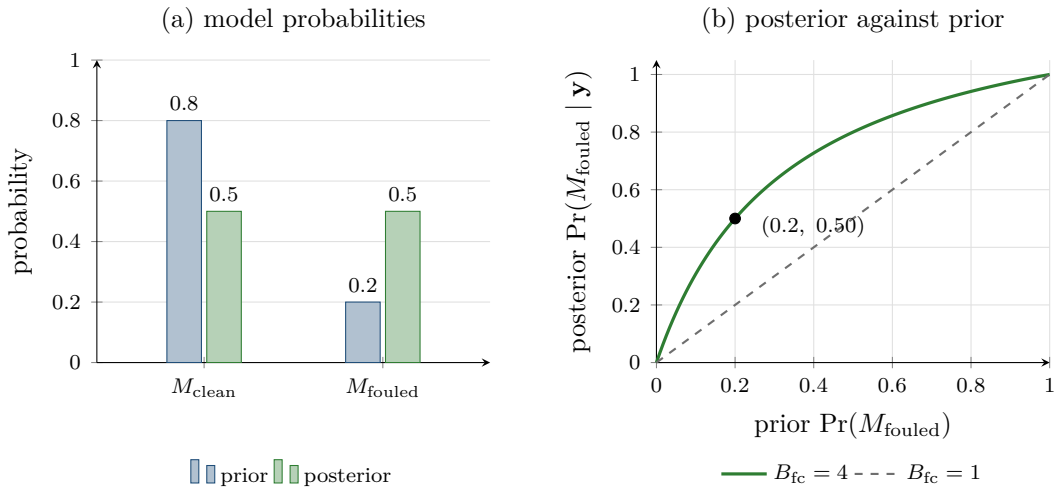


Figure 2.6: Model comparison for the air path. (a) Prior and posterior probabilities of the clean (M_{clean}) and fouled (M_{fouled}) models, with evidences 0.01 and 0.04. (b) Posterior probability of M_{fouled} as a function of its prior, $4x/(3x+1)$ for a Bayes factor of 4 (equation (2.48)); the dashed diagonal is a Bayes factor of 1; the point is the example.

2.5.3 Diagnosis over a family of candidate faults

Fault detection and isolation usually involves many competing explanations, not two. For a hall whose condition tests read warmer than at commissioning they might be a fouled coil, a clogged filter, a fan running slow, a biased hall sensor or a load meter reading low. The mathematical target is the general form (2.15), normalized as in equation (2.47): the posterior probability of each candidate,

$$\Pr(F_j \mid \mathbf{y}) = \frac{p(\mathbf{y} \mid F_j)\,\Pr(F_j)}{\sum_\ell p(\mathbf{y} \mid F_\ell)\,\Pr(F_\ell)},$$

where F_j is a candidate fault condition, which is the drawing M_j of equation (2.15): the plant graph with that fault written into its rows, so that $p(\mathbf{y} \mid F_j) = p(\mathbf{y} \mid M_j)$. The output is the ranking of the candidates by it, not a single verdict. A fault condition can be represented in two ways: as a separate model, whose evidence enters equation (2.47), or as a latent quantity within one common model, such as a bias entry of the state that is zero when the sensor is sound. The second representation means that diagnosis does not always require fitting a separate model for every possible fault. Diagnosis earns its own category for two reasons: the candidates are generated systematically from the plant graph, one for each component or reading that could be at fault, rather than written by hand; and the answer is a ranking. The critical step in either representation is to define the candidate fault events or hypotheses clearly enough that their probabilities can be calculated and compared.

For the hall of Figure 2.1 the family is built on the graph with its reading row kept as a row of its own, $(T_1 - y_T)/\sigma_T$, rather than folded, as in the fit of Section 2.4.1, into the commissioning row on g; a fault must have a row to act on. One candidate is generated for each element that could be at fault, each the extended graph with one entry added: M_{bias}, a bias entry Δ_T on the reading row, which becomes $(T_1 + \Delta_T - y_T)/\sigma_T$, the sensor drifting; M_{path}, a shift Δ_g of the log-conductance away from its commissioned value, an entry added to both balance rows, the fouling; and M_{load}, an offset Δ_Q of the test load Q_1 from its nominal value, an entry added to the test-load balance row, the load meter reading wrong. Each candidate is fitted to the same readings, its evidence is read from its own $\hat{\mathbf{z}}_j$ and $\boldsymbol{\Sigma}_j$, and the ranking orders them. Which candidates the readings can tell apart is read from the Jacobian before any fit. With one reading of the hall, each added entry can absorb the misfit of that reading alone, so the three candidates

fit it equally well and their ranking rests on their prior rows and prior probabilities. The hall's two tests of the opening do no better, because both were taken at 30 kW: the rise between them is absorbed equally by a drift of the bias, a loss of conductance or a change of the load offset. A test at another load separates them, because the three added entries then move the two readings in different proportions, which are different columns of $\mathbf{J}$: a bias by equal amounts, a shift of the path in proportion to the loads, and an offset of the test load at the test load only.

2.5.4 Expected loss of a filter change before the summer peak

Questions about an asset also lead to decisions. Before the summer peak the asset manager can change the cooling unit's filters, which restores some of the air path's conductance, or wait, and the probability of a shortfall is not enough to decide: the consequence of a shortfall and the cost of the action enter as well. Let e be the event that the hall cannot be held at the summer peak, the hall air staying above its limit long enough for the IT load to be shed, with consequence $C_e = \text{R\$}\,100\,000$, and suppose that if the asset manager waits its probability is $\Pr(e \mid \mathbf{y}, \text{wait}) = 0.0015$, an illustrative value. A shortfall in this sense, at today's load and held long enough to shed load, is far rarer than the hall crossing the 27.5°C envelope at the requested 60 kW of Question B, whose probability is 0.747, with 0.754 by the tail of g under the commissioning row and 0.719 under the exact posterior. The shortfall is a future event that stays random when the state is known, so the loss of waiting given the state is itself an expectation, the derived quantity

$$L(\text{wait}, \mathbf{z}) = C_e \Pr(e \mid \mathbf{z}, \text{wait}) + 0 \cdot \Pr(\bar{e} \mid \mathbf{z}, \text{wait}) = C_e \Pr(e \mid \mathbf{z}, \text{wait}),$$

with no loss when the hall is held, and the expected loss is its location (2.10), an estimation question about a consequence, answered by the expected consequence (2.23) of waiting. Averaging over the posterior with equation (2.8) turns the conditional probability into its predictive one, and the expected loss is the probability-weighted average of the two outcomes:

$$\mathcal{R}(\text{wait}) = \mathbb{E}[L(\text{wait}, \mathbf{z}) \mid \mathbf{y}] = \Pr(e \mid \mathbf{y}, \text{wait})\, C_e + \Pr(\bar{e} \mid \mathbf{y}, \text{wait}) \cdot 0, \tag{2.49}$$

the second line of equation (2.23) with one event, consequence C_e, and no cost of action: probability times consequence. Substituting,

$$\mathcal{R}(\text{wait}) = 0.0015 \times \text{R\$}\,100\,000 + 0.9985 \times 0 = \text{R\$}\,150$$

for the summer, under these assumptions; the probability that the hall is held at the peak is 0.9985. It does not mean that the owner will lose exactly R$ 150; the actual loss is either zero or R$ 100 000. It means that R$ 150 is the probability-weighted average of the possible losses, the amount that would be lost per summer on average over many summers like this one.

Now suppose the filter change costs $c_{\text{change}} = \text{R\$}\,80$ in filters and labor (a cost, the c_a of equation (2.23)), and lowers the probability of a shortfall from 0.0015 to 0.0005. Its loss given the state is $L(\text{change}, \mathbf{z}) = c_{\text{change}} + C_e \Pr(e \mid \mathbf{z}, \text{change})$, and its expected loss is the certain cost of the action plus the expected loss that remains,

$$\mathcal{R}(\text{change}) = \text{R\$}\,80 + 0.0005 \times \text{R\$}\,100\,000 = \text{R\$}\,130,$$

against $\mathcal{R}(\text{wait}) = \text{R\$}\,150$. The expected net benefit of the filter change is $\text{R\$}\,150 - \text{R\$}\,130 = \text{R\$}\,20$ for the summer, and under these assumptions changing the filters is the preferred decision. The comparison rests on three numbers, two probabilities and one consequence, and its conclusion changes if any of them moves: the break-even cost of the action is $(0.0015 - 0.0005) \times \text{R\$}\,100\,000 = \text{R\$}\,100$, and, read the other way, changing and waiting cost the same when the probability of a shortfall if the asset manager waits is $0.0005 + 80/100\,000 = 0.0013$, the break-even probability.

A summer whose shortfall probability the posterior puts below 0.0013 is a summer to wait. The probabilities here stand for the predictives a posterior would supply; they are given, so the expected losses carry no gap of their own, and a gap in the shortfall probability that straddled 0.0013 would be stated with both values and both actions.

2.5.5 The general decision formulation

Let a be a maintenance action from a set of candidates, and $L(a, \mathbf{z})$ the loss of taking it when the state is $\mathbf{z}$. The expected loss of the action is its expected consequence (2.23), $\mathcal{R}(a) = \mathbb{E}[L(a, \mathbf{z}) \mid \mathbf{y}]$, the location (2.10) of the derived quantity $L(a, \mathbf{z})$, one for each candidate, and the preferred action is the one that minimizes it:

$$a^* = \arg\min_a \mathbb{E}[L(a, \mathbf{z}) \mid \mathbf{y}]. \tag{2.50}$$

The filter-change example of Section 2.5.4 is this rule with one event, the shortfall, and two actions. The change is an important one. The questions of the earlier sections asked what will happen; equation (2.50) asks what should be done, given what might happen. The transition from inference to decisions is what asset management is for: the asset manager wants a date and a decision, not a health index.

2.6 Measurements, the classification of questions and the derivation workflow

A design question asks whether more information is worth obtaining, and it is answered before the information exists.

2.6.1 Uncertainty reduction by a measurement

Suppose the asset manager is uncertain about the conductance because the record is short, one commissioning test, and another condition test before the next shutdown could reduce that uncertainty. The question "is the test worth its cost?" is a design question, of the general form (2.16). Its first part is how much the measurement would narrow the posterior: the form with U_{p} the posterior covariance, which for linear Gaussian rows is the preposterior covariance (2.17), with the information gain (2.18) as its summary. The asset enters through two objects: the Jacobian $\mathbf{J}_a$ of the candidate rows, which says which entries of the state the measurement acts on and how strongly, and the current covariance $\boldsymbol{\Sigma}$, which says along which directions the posterior can still shrink.

For the hall, consider a condition test of n readings of the hall temperature at the test load $Q_1 = 30\,\mathrm{kW}$, each with an error of standard deviation $\sigma_T = 0.5\,\mathrm{K}$, the hall sensor's, taken independently of one another. Each reading is a row on the first entry of the state of equation (2.30),

$$r_{a,j}(\mathbf{z}) = \frac{T_1 - \tilde{y}_j}{\sigma_T}, \qquad \mathbf{J}_{a,j} = \frac{\partial r_{a,j}}{\partial \mathbf{z}} = \begin{bmatrix} 1/\sigma_T & 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \end{bmatrix} \mathrm{K}^{-1},$$

so that n readings stack to an $n \times 3$ Jacobian $\mathbf{J}_a$ with $\boldsymbol{\Lambda}_a = \mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a = \mathrm{diag}(4n, 0, 0)$ in K^{-2}. Independence of the readings is a modeling choice: it is what lets the n rows add, and it fails when the errors share a cause, a sensor offset that every reading of one test carries (Chapter 3 treats that case). The current $\boldsymbol{\Sigma}$ is the ridge of equation (2.33), $\sigma^2_{\hat{g}_0}\mathbf{v}\mathbf{v}^{\mathsf{T}}$, so in the update form (2.27) every product runs through $\mathbf{v}$: $\boldsymbol{\Sigma}\mathbf{J}_{a,j}^{\mathsf{T}} = \sigma^2_{\hat{g}_0}\,\mathbf{v}\,(2v_1)$ and $\mathbf{J}_{a,j}\boldsymbol{\Sigma}\mathbf{J}_{a,\ell}^{\mathsf{T}} = 4\sigma^2_{\hat{g}_0}v_1^2 =$

$4\Sigma_{11} = 4 \times 0.23096 = 0.92384$ for every pair of readings. The update therefore removes a multiple of $\mathbf{v}\mathbf{v}^{\mathsf{T}}$ itself, and

$$\boldsymbol{\Sigma}_a = \frac{\boldsymbol{\Sigma}}{1 + 0.92384\,n}, \qquad \boldsymbol{\Sigma}_a = \frac{\boldsymbol{\Sigma}}{1.92384} = \begin{bmatrix} 0.12005 & 0.2401 & -0.02368 \\ 0.2401 & 0.48021 & -0.04736 \\ -0.02368 & -0.04736 & 0.00467 \end{bmatrix} \quad (n = 1). \tag{2.51}$$

The reading sees only T_1, yet every entry of $\boldsymbol{\Sigma}$ shrinks by the same factor, because the posterior has only the one direction $\mathbf{v}$ to shrink along: one reading removes 27.9 % of the spread of T_1, T_2 and g alike, so that $\sqrt{\Sigma_{a,33}}$ falls from 0.0948 to 0.0683, the spread of the health index from 0.0935 to 0.0674, and $\sqrt{\Sigma_{a,22}}$ from 0.961 to 0.693 K. The row $n = 1$ is the second hall sensor that step 7 of Chapter 1 ranks, the same reading row added to the same posterior. Table 2.9 gives the result for 0 to 64 readings. The columns are the number of readings, the precision $1/\Sigma_{a,33}$ of g, its posterior standard deviation $\sqrt{\Sigma_{a,33}}$, the standard deviation $\sqrt{\Sigma_{a,22}}$ of T_2, the standard deviation of the health index, the fraction of the spread removed, and the expected information gain. The precision of g is $111.30\,(1 + 0.92384\,n) = 111.30 + 102.82\,n$ by equation (2.51). The expected information gain is equation (2.18); in the limit both determinants are zero, and the determinant identity $\det\boldsymbol{\Sigma}/\det\boldsymbol{\Sigma}_a = \det(\mathbf{I} + \mathbf{J}_a\boldsymbol{\Sigma}\mathbf{J}_a^{\mathsf{T}})$, with every entry of the $n \times n$ matrix $\mathbf{J}_a\boldsymbol{\Sigma}\mathbf{J}_a^{\mathsf{T}}$ equal to 0.92384, gives $\mathrm{EIG} = \frac{1}{2}\ln(1 + 0.92384\,n)$. Equation (2.17) evaluated directly at $\sigma_p = 10^{-4}$ K, where $\boldsymbol{\Sigma}^{-1}$ exists, gives the same entries, the precision of 64 readings to within 0.02. The reduction is steep at first and then slows, because the precision grows linearly in n and the spread only as its inverse square root: halving the spread requires four times the precision, $n = 3.25$ readings, so four, and halving it again sixteen times the precision, $n = 16.24$, so seventeen.

Table 2.9: Preposterior spread of the hall after a condition test of n readings at $Q_1 = 30$ kW with $\sigma_T = 0.5$ K, from the posterior $\boldsymbol{\Sigma}$ of equation (2.33) and equation (2.51).

n [–]	precision $1/\Sigma_{a,33}$ [–]	$\sqrt{\Sigma_{a,33}}$ [–]	$\sqrt{\Sigma_{a,22}}$ [K]	sd of h [–]	spread removed [–]	EIG [nat]
0	111.3	0.0948	0.961	0.0935	0	0
1	214.1	0.0683	0.693	0.0674	27.9 %	0.327
4	522.6	0.0437	0.444	0.0431	53.9 %	0.773
16	1756.5	0.0239	0.242	0.0235	74.8 %	1.379
64	6692.1	0.0122	0.124	0.0121	87.1 %	2.048

2.6.2 The value of information

Reducing uncertainty is not necessarily the same as making a better decision. Suppose there are two possible actions, change the filters and wait. The best action now minimizes the expected loss, and the expected loss of acting now is

$$\bar{L}_{\mathrm{now}} = \min_a \mathbb{E}[L(a, \mathbf{z}) \mid \mathbf{y}].$$

(The book writes $\bar{L}$ for these minimum expected losses because C names a consequence.) If a candidate measurement is collected first, written m here because a ranges over the actions, with outcome $\tilde{y}_{\mathrm{m}}$, the best action can be chosen again after its value is known, with expected loss $\min_a \mathbb{E}[L(a, \mathbf{z}) \mid \mathbf{y}, \tilde{y}_{\mathrm{m}}]$. Before the measurement is collected its result is unknown, so that expected loss is averaged over the possible results, weighted by their predictive distribution given the present evidence:

$$\bar{L}_{\mathrm{after}} = \mathbb{E}_{\tilde{y}_{\mathrm{m}}|\mathbf{y}}\Big[\min_a \mathbb{E}[L(a, \mathbf{z}) \mid \mathbf{y}, \tilde{y}_{\mathrm{m}}]\Big].$$

This is the general form (2.16) for the candidate m, with U_{p} the minimum expected loss under the updated posterior and the candidate ranked smallest first, and here, unlike for the spread, U_{p} depends on the outcome itself and not only on its covariance. The expected value of sample information is the difference,

$$\mathrm{EVSI} = \bar{L}_{\mathrm{now}} - \bar{L}_{\mathrm{after}}, \tag{2.52}$$

the expected improvement in the decision, measured as a reduction in expected loss. It is never negative. For each fixed action, averaging the post-measurement expected loss over the results returns the present one, $\mathbb{E}_{\tilde{y}_{\mathrm{m}}|\mathbf{y}}\big[\mathbb{E}[L(a,\mathbf{z}) \mid \mathbf{y}, \tilde{y}_{\mathrm{m}}]\big] = \mathbb{E}[L(a,\mathbf{z}) \mid \mathbf{y}]$, by the law of total expectation; and the minimum over actions inside the average is at most the value of any one action, so $\bar{L}_{\mathrm{after}} \leq \min_a \mathbb{E}[L(a,\mathbf{z}) \mid \mathbf{y}] = \bar{L}_{\mathrm{now}}$. Equality holds when no possible result would change the action: a measurement that cannot change the decision has no value for it, however much it narrows the posterior. If the measurement costs more than its EVSI, obtaining it is not economically justified on this criterion.

An illustrative case shows the calculation. Suppose the shortfall probability 0.0015 of Section 2.5.4 is an average over two conditions of the air path: fouled, with probability 0.25 and shortfall probability 0.0045, and clean, with probability 0.75 and shortfall probability 0.0005; indeed $0.25 \times 0.0045 + 0.75 \times 0.0005 = 0.0015$. The filter change costs R$ 80 and leaves either path at 0.0005. An inspection of the coil and filters before the summer would reveal the condition exactly. Table 2.10 lists, for each condition and for the present state of knowledge, the expected loss of each action and the better one. Without the inspection the best action is to change the filters, at R$ 130. With it, a fouled path is serviced at R$ 130 and a clean one left alone at R$ 50, so

$$\bar{L}_{\mathrm{after}} = 0.25 \times \mathrm{R\$}\,130 + 0.75 \times \mathrm{R\$}\,50 = \mathrm{R\$}\,70, \qquad \mathrm{EVSI} = \mathrm{R\$}\,130 - \mathrm{R\$}\,70 = \mathrm{R\$}\,60.$$

An inspection that reveals the condition is worth up to R$ 60 for the summer; because it reveals the condition exactly, this is the expected value of perfect information, an upper bound on what any imperfect inspection or condition test could be worth.

Table 2.10: Expected loss for the summer of the two actions for an air path of unknown condition (illustrative): filter change R$ 80, shortfall consequence R$ 100 000.

Condition	probability [–]	Pr(e), wait [–]	wait [R$]	filter change [R$]	better action
fouled	0.25	0.0045	450	130	filter change
clean	0.75	0.0005	50	130	wait
unknown (present)	1	0.0015	150	130	filter change

Three quantities are now on the table, and they are related but not interchangeable. The reduction of the posterior covariance, equation (2.17), measures how much uncertainty disappears. The expected information gain, equation (2.18), measures how much is learned. The value of information, equation (2.52), measures how much better the resulting decision is expected to be. A measurement can score well on the first two and be worth nothing on the third, which is why the design of experiments and the optimization of decisions both need the preposterior but do not ask the same question of it. The value of information is where two computations meet: the preposterior supplies the outcomes and their weights, and the expected consequence, minimized over the actions, is what it averages.

2.6.3 Identifying the mathematical form of a question

The examples supply a systematic procedure: before computing anything, identify the mathematical objective of the question. Table 2.11 is the reference for that step. Each row is a

form of question in the asset manager's words, the general form it takes, by equation number, its mathematical form, what it derives to, and its category; the table mirrors the atlas of Appendix A, which collects the questions of every chapter with the chapter that derives each.

Table 2.11: Classification of questions by their mathematical form, with the general form of each category from Section 2.3.3 and the forms of a trend and of an expected consequence from Sections 2.3.4 and 2.3.5.

Asset manager's question	General form	Mathematical form	Derives to	Category
What condition is it in, and how well is that known?	(2.9), (2.10)	$p(z_i \mid \mathbf{y})$, $\mathbb{E}[z_i \mid \mathbf{y}]$	a distribution, a mean, a credible interval: $\hat{z}_i$ and Σ_{ii}; a health index through its gradient	estimation
Does a requirement hold, with the stated assurance?	(2.12)	$\Pr(f(\mathbf{z}) \geq f^* \mid \mathbf{y}) \geq \kappa$, f^* the requirement	the posterior integral of the indicator of the requirement; a normal tail under linearization	prediction under a threshold
What happens in the coming season?	(2.13)	$\int \Pr(A_{\text{future}} \mid \mathbf{z})\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}$	the posterior predictive, e.g. $(b'/(b'+\tau))^{a'}$	prediction under a threshold
How fast is it degrading, and when does it reach its limit?	(2.21)	$\Pr(\gamma < 0 \mid \mathbf{y})$; $\Pr(t^* \leq s \mid \mathbf{y}) = \Pr(\gamma < 0,\ g_s \leq g^* \mid \mathbf{y})$	the posterior read along time, $\mathbf{g}(t)^{\mathsf{T}}\hat{\mathbf{z}}$ and $\mathbf{g}(t)^{\mathsf{T}}\boldsymbol{\Sigma}\,\mathbf{g}(t)$; a crossing band by inversion	prediction under a threshold for the sign and a crossing by a date; estimation for a forecast and the distribution of a crossing time
Which explanation is correct?	(2.14), (2.15)	$\Pr(M_j \mid \mathbf{y}) \propto p(\mathbf{y} \mid M_j)\, \Pr(M_j)$	the evidences and the prior model probabilities; the Bayes factor	hypothesis comparison; diagnosis over a family of faults
What is the expected consequence?	(2.23)	$\mathcal{R}(a) = \mathbb{E}[L(a,\mathbf{z}) \mid \mathbf{y}] = \int L(a,\mathbf{z})\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}$	probability times consequence; $L(a,\hat{\mathbf{z}}) + \frac{1}{2}\operatorname{tr}(\mathbf{H}_a\boldsymbol{\Sigma})$	estimation of a derived quantity
What should be done or measured?	(2.23), (2.16)	$\arg\min_a \mathcal{R}(a)$; EVSI	the expected consequence of each action; the preposterior $\boldsymbol{\Sigma}_a$	design

Many variations of these forms exist. Importance measures, the forecasting of a health index, maintenance scheduling and spare-parts optimization each add mathematical structure of their own: gradients, integrals of failure probability over an interval, consequences that accrue over time. The rows of the table, over the general forms of the categories, of a trend and of an expected consequence, already cover a substantial range of inference problems, and a new question is usually classified by recognizing which of them it is, before any computation.

2.6.4 The general derivation workflow

The classification is one step of the derivation workflow. Figure 2.7 draws its eight steps: the top row runs from left to right, the bottom row from right to left, and the arrow beneath the bottom row is the return from step 7 to step 3 when a new measurement is added. In the words of the preface:

The plant graph (step 0). The physical system as it is, before any question is asked: its

nodes, paths and sources; which quantities are pinned and which are free; the physical relations that must hold among them. The free quantities form the state vector $\mathbf{z}$.

The question (step 1). A probability or an expectation of a function of the state, with the unknown on the left of the conditioning bar and everything treated as known on the right.

The graph extended for the question (step 2). Scenario nodes, a second drawing when two structures compete, the windows of the record with their trend rows, and the evidence attached as rows: readings, counts, handbook values and commissioning tests as priors, each with the modeling assumption it rests on, what that assumption buys and how it can fail.

The posterior of the state (step 3). The rows combined into one objective; its minimum, the MAP state; its curvature, the information; and the covariance, derived and then computed for the asset, with the exact posterior beside them where the model admits one. It is built once and serves every question.

The answers (step 4). Each question read from that posterior, from the asset manager's question through its general and specialized forms down to its number: an entry of the state or the gradient of a derived quantity, the quantity at a time asked included; when drawings compete, a ratio of evidences; when actions compete, an expected consequence.

Evaluation and checks (step 5). The closed forms evaluated, and their approximations tested against the full calculation.

The report (step 6). The probability, the Bayes factor or the expected consequence with its band, or a crossing date with its band by inversion, and the entry in the maintenance log.

The next measurement (step 7). If the band is too wide, the candidate measurements ranked by the width each would remove, and a return to step 3 with the new row.

The examples of this chapter sit in it as follows. Stating precisely what the asset manager needs to know or decide, and defining the events, uncertain quantities, future outcomes, hypotheses or actions that express it, is step 1; adding the full-load scenario node of the hall example, and attaching the commissioning fit, the readings and the priors as rows with the assumptions behind them, is step 2. Step 3 builds the posterior the questions are read from: a fit for the most probable state and its Laplace covariance, a numerical integration, the evidence of each competing model, or another operation appropriate to the problem; for the hall it is the Gauss–Newton fit of three rows, for the hall's two tests the fit of the straight line, and for the work-order log the Newton fit of the log-rate, whose objective is a gamma kernel and so gives the update in closed form as well. Step 4 writes each question in its general form and reduces it, through the specialized form, to posterior marginals, integrals, gradients, evidences or decision objectives, and evaluates them. Step 5 checks assumptions, identifiability, approximations, numerical convergence and the sensitivity of the answer to modeling choices, as the tolerance, exact-value and sampling checks of the hall example did. Step 6 communicates the answer, its uncertainty, its limitations and its maintenance implications, and step 7 is the design question of Section 2.6.1.

The workflow is broader than any particular software implementation. The derivation fixes what has to be calculated before a computational tool is chosen, and the tool then calculates it.

Principle 2.3. *The derivation tells the computational tool what must be calculated. A computational tool should not dictate which asset-management question is asked.*

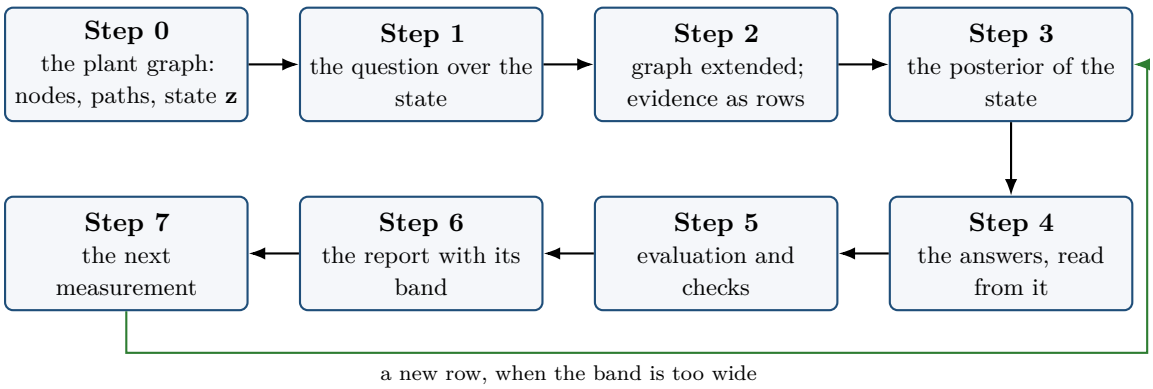


Figure 2.7: The derivation workflow of the book. Step 0 is built once per asset; steps 1 to 7 are one pass for one question; the green return arrow adds the measurement chosen at step 7 as a new row and rebuilds the posterior from step 3.

2.6.5 Common mistakes in probabilistic formulation

Six mistakes of formulation recur in asset management, and each is a confusion between two questions that look alike.

Common mistake

Confusing uncertainty about a rate with the probability of a quiet horizon. For a coil design with an unknown work-order rate λ, the probability of no work order over a horizon τ at a given rate is $\mathrm{e}^{-\lambda\tau}$. The question $\mathbb{E}[\mathrm{e}^{-\lambda\tau} \mid \mathbf{y}]$ is the posterior predictive probability: how likely the coming season is to pass without a work order. The question $\Pr(\mathrm{e}^{-\lambda\tau} > 0.99 \mid \mathbf{y})$ is the posterior probability that the design's underlying rate meets a requirement written as a 99 % chance of a quiet season: how much support there is for the requirement. They are different numbers and answer different questions.

Common mistake

Confusing prediction with diagnosis. When the hall runs warm, the prediction $\Pr(T > T_{\max} \mid \mathbf{y})$ concerns a temperature event, and the diagnosis $\Pr(M_{\text{path}} \mid \mathbf{y})$ concerns an explanation, a fouled air path against a drifting sensor. They can use the same readings but require different calculations: a tail of a posterior in one model against a comparison of evidences across models.

Common mistake

Ignoring the time horizon. $\Pr(T > 90 \text{ days} \mid \mathbf{y})$ and $\Pr(T > 15\,\text{yr} \mid \mathbf{y})$ are not interchangeable statements about the coil; for the work-order log of Section 2.4.4 the first, a summer season without a work order, is 0.929 (0.937 exact) and the second, the unit's fifteen-year design life without one, is 0.043 (exact, $(15/30)^4 = 0.0625$). A probability of a quiet horizon is incomplete without its horizon and its operating conditions.

Common mistake

Forgetting dependence. In general $\Pr(A \cap B) \neq \Pr(A)\,\Pr(B)$; the product is justified only under independence. A shared control-power supply, the weather, a common maintenance crew and the same batch of filters create dependence between components and units, and Table 2.8 shows that a small shared cause can multiply the probability of two replacements in one year several times over.

Common mistake

Treating the model assumptions as certain. A calculated probability may be numerically precise while depending strongly on uncertain modeling assumptions. A temperature-exceedance probability computed under the steady-state model of equation (2.2) may be unsuitable for the minutes after the cooling unit trips, when the hall air rises through a transient that the steady-state model does not describe and the supply air is no longer held at T_c. Every probability is conditional on its model, $\Pr(A \mid \mathbf{y}, M)$, and must be interpreted in that context.

Common mistake

Trending raw performance as condition. The dashboard of the opening read the approach $T - T_c$ as the health index: up 5.77 % in two years, "5.8 % lost". Read through the physics, the same two readings taken as exact are a fall of the conductance by 5.45 %, and the fit puts the health index at window 24 at 0.9091. The approach's straight line reaches its limit at year 5.33, 1.20 years after the plug-in date of the trend on g, year 4.14, beside its median, year 4.15, and carries no band, while the trend's band by inversion has no upper end: a mass of 0.268 never crosses. A raw-performance figure mixes the condition with the load and the supply temperature, and its line is the naive baseline, not the condition.

The questions of this chapter were all written with conditional probabilities and the posterior was taken as given; the next chapter, Chapter 3, derives conditional probability and Bayes' rule, separates aleatory variability from epistemic uncertainty, and shows how condition tests, inspection records and work-order logs change the uncertainty about the condition of an asset through marginalization, dependence and conditional independence on the plant graph.

2.7 What to retain

Definition 2.1 (Asset-management question)**.** An asset-management question is a probability or an expectation of a function of the state of the asset's plant graph, written with the unknown on the left of the conditioning bar and the evidence and the model on the right, $\Pr(A \mid \mathbf{y}, M) = \int \mathbf{1}_A(\mathbf{z})\, p(\mathbf{z} \mid \mathbf{y}, M)\, \mathrm{d}\mathbf{z}$ for a present event and $\int \Pr(A_{\text{future}} \mid \mathbf{z}, M)\, p(\mathbf{z} \mid \mathbf{y}, M)\, \mathrm{d}\mathbf{z}$ for a future one. Its form decides its category, estimation, prediction under a threshold, hypothesis comparison, diagnosis or design, each with its general forms, equations (2.9) to (2.16); the question decides which of five computations answers it, the posterior of the state, the trend, the evidence, the preposterior or the expected consequence (Table 2.3), two of which carry general forms of their own, the trend form (2.21) and the expected-consequence form (2.23); and the asset enters only in the specialized form, through the entry or function of the state, the event or family of models, and the objects of the fit that carry the answer. A health index is a parameter of that state read against its design or commissioned value, never a raw-performance figure.

Table 2.12 lists the chapter's terms with their expressions, units and the subsection that defines each, grouped by their role.

Table 2.12: The terms of asset-management questions and where each is defined.

Term	Expression	Units	Meaning	Defined in
Event	A, e.g. $T > 28\,^{\circ}\mathrm{C}$	–	a statement true or false	2.2.1
Conditional probability	$\Pr(A \mid \mathbf{y}, M)$	–	probability given evidence and model	2.2.3

Term	Expression	Units	Meaning	Defined in
Event probability from a density	$\int_5^7 p(G \mid \mathbf{y})\,\mathrm{d}G$	–	probability of the range $5 < G < 7$	2.2.4
Probability over the state	$\int \mathbf{1}_A(\mathbf{z})\,p(\mathbf{z} \mid \mathbf{y})\,\mathrm{d}\mathbf{z}$	–	posterior expectation of the indicator	2.3.1
Posterior predictive	$\int \Pr(A_{\text{future}} \mid \mathbf{z})\,p(\mathbf{z} \mid \mathbf{y})\,\mathrm{d}\mathbf{z}$	–	future event averaged over the state	2.3.2
General forms				
Estimation	$\Pr(f(\mathbf{z}) \le f^* \mid \mathbf{y})$; $\mathbb{E}[f \mid \mathbf{y}]$, $\mathrm{Var}(f \mid \mathbf{y})$	–	posterior of a derived quantity	2.3.3
Credible interval	$\Pr(f \le f_L \mid \mathbf{y}) = 0.025$, $\Pr(f \le f_U \mid \mathbf{y}) = 0.975$	–	central 95 % band; quantiles map through monotone f	2.3.3
Threshold and assurance	$\Pr(f(\mathbf{z}) \ge f^* \mid \mathbf{y}) \ge \kappa$	–	probability that the requirement $f \ge f^*$ holds, against κ	2.3.3
Predictive	$\int p(y_{\text{new}} \mid \mathbf{z})\,p(\mathbf{z} \mid \mathbf{y})\,\mathrm{d}\mathbf{z}$	–	future observation averaged over the state	2.3.3
Comparison	$B_{AB} = p(\mathbf{y} \mid M_A)/p(\mathbf{y} \mid M_B)$	–	evidence ratio of two drawings	2.3.3
Diagnosis	$\Pr(M_j \mid \mathbf{y}) \propto p(\mathbf{y} \mid M_j)\ \Pr(M_j)$, ranked	–	comparison over a family of faults	2.3.3
Design	$U(a) = \mathbb{E}_{\tilde{y}_a \mid \mathbf{y}}[U_{\mathrm{p}}(p(\mathbf{z} \mid \mathbf{y}, \tilde{y}_a))]$	varies	preposterior expectation	2.3.3
Preposterior covariance	$\boldsymbol{\Sigma}_a = (\boldsymbol{\Sigma}^{-1} + \mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a)^{-1}$	varies	spread after a future measurement	2.3.3
Expected information gain	$\frac{1}{2}\ln(\det \boldsymbol{\Sigma}/\det \boldsymbol{\Sigma}_a)$	nat	what a measurement teaches	2.3.3
Trend	$\Pr(\gamma > 0 \mid \mathbf{y})$; $\Pr(t^* \le s \mid \mathbf{y}) = \Pr(\gamma > 0,\ f_s(\mathbf{z}) \ge f^* \mid \mathbf{y})$; $p(f_t(\mathbf{z}) \mid \mathbf{y})$; read with $\mathbf{g}(t)^{\mathsf{T}}\hat{\mathbf{z}}$, $\mathbf{g}(t)^{\mathsf{T}}\boldsymbol{\Sigma}\,\mathbf{g}(t)$	–	sign, crossing and forecast: the posterior read along time	2.3.4
Falling health index	$f_t = -g_t$, $f^* = -g^*$; sign $\Pr(\gamma < 0 \mid \mathbf{y})$, crossing $\Pr(\gamma < 0,\ g_s \le g^* \mid \mathbf{y})$, never-crossing mass $\Pr(\gamma \ge 0 \mid \mathbf{y})$	–	the trend form for a quantity that falls to its level	2.3.4
Rows of a record in windows	$\mathbf{J} = \mathbf{D}\mathbf{X}$, $[\boldsymbol{\Lambda}]_{\gamma\gamma} = \sum_\ell d_\ell^2 (t_\ell - \bar{t})^2$	varies	a window's distance enters squared, windows add linearly	2.3.4
Expected consequence	$\mathcal{R}(a) = \mathbb{E}[L(a, \mathbf{z}) \mid \mathbf{y}]$, $a^* = \arg\min_a \mathcal{R}(a)$; spread $\sqrt{\mathbf{g}_a^{\mathsf{T}}\boldsymbol{\Sigma}\,\mathbf{g}_a}$	varies	probability times consequence, minimized over actions	2.3.5
Curvature term	$L(a, \hat{\mathbf{z}}) + \frac{1}{2}\,\mathrm{tr}(\mathbf{H}_a\boldsymbol{\Sigma})$	varies	what the plug-in loss misses, in the state's coordinates	2.3.5
Computations				
Posterior	$p(\mathbf{z} \mid \mathbf{y}) \propto \mathrm{e}^{-\mathcal{L}(\mathbf{z})}$; $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$	varies	what estimation and prediction integrate	2.3.6
Evidence	$p(\mathbf{y} \mid M) = \int p(\mathbf{y} \mid \mathbf{z}, M)\,p(\mathbf{z} \mid M)\,\mathrm{d}\mathbf{z}$	–	normalizing constant of a model	2.3.6
Preposterior	$p(\mathbf{z} \mid \mathbf{y}, \tilde{y}_a)$ averaged over $p(\tilde{y}_a \mid \mathbf{y})$	varies	the expected posterior under a measurement not yet taken	2.3.6
Information adds	$\mathbf{J}^{\mathsf{T}}\mathbf{J} + \mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a$	varies	stacked rows of a candidate	2.3.6

Term	Expression	Units	Meaning	Defined in
Update form	$\boldsymbol{\Sigma} - \boldsymbol{\Sigma}\mathbf{J}_a^{\mathsf{T}}(\mathbf{I} + \mathbf{J}_a\boldsymbol{\Sigma}\mathbf{J}_a^{\mathsf{T}})^{-1}\mathbf{J}_a\boldsymbol{\Sigma}$	varies	$\boldsymbol{\Sigma}_a$ when $\boldsymbol{\Sigma}$ is singular	2.3.6
Building blocks of the fit				
State	$\mathbf{z} = (T_1, T_2, g)^{\mathsf{T}}$	varies	every free quantity of the graph at once	2.4.1
Commissioning row	$(g - \hat{g}_0)/\sigma_{\hat{g}_0}$, $\hat{g}_0 = -0.013926$, $\sigma_{\hat{g}_0} = 0.094787$	–	the commissioning fit of Chapter 1 as one row	2.4.1
Rows and objective	$r_\ell(\mathbf{z})/\sigma_\ell$; $\mathcal{L} = \frac{1}{2} r^{\mathsf{T}} r$	–	physics, reading and prior rows in standard deviations	2.4.1
Jacobian	$\mathbf{J} = \partial r / \partial \mathbf{z}$	varies	sensitivity of each row to each entry	2.4.1
Normal equations	$(\mathbf{J}^{\mathsf{T}}\mathbf{J})\,\Delta\mathbf{z} = -\mathbf{J}^{\mathsf{T}} r$	varies	Gauss–Newton step to $\hat{\mathbf{z}}$	2.4.1
Laplace covariance	$\boldsymbol{\Sigma} = (\mathbf{J}^{\mathsf{T}}\mathbf{J})^{-1}$	varies	band of every entry of the state	2.4.1
Ridge of exact rows	$\boldsymbol{\Sigma} \to \sigma_{\hat{g}_0}^2 \mathbf{v}\mathbf{v}^{\mathsf{T}}$ as $\sigma_p \to 0$	varies	entries tied by the physics move as one	2.4.1
Band of a derived quantity	$\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$, $\mathbf{g} = \partial f / \partial \mathbf{z}$	varies	delta method	2.4.1
Health index	$h = \mathrm{e}^{g} = 0.9862 \pm 0.0935$	–	the conductance against the design sheet, through $\mathbf{g} = (0,\ 0,\ h)^{\mathsf{T}}$	2.4.1
Specialized forms and answers				
Requirement probability of the hall	$\Phi\big((T_{\max} - \hat{z}_2)/\sqrt{\Sigma_{22}}\big) = 0.253$	–	reported, the normal tail of the entry	2.4.2
The tail of g	$\Phi\big((\hat{z}_3 - \ln(G^*/G_{\mathrm{ref}}))/\sqrt{\Sigma_{33}}\big) = 0.246$	–	the bound on T_2 mapped along the ridge to the third entry; gap 0.0070; the exact posterior gives 0.2815	2.4.2
Work-order rate	$\hat{\lambda} \pm \hat{\lambda}\,\sigma_\omega = 0.267 \pm 0.133$	yr^{-1}	reported, from the log-rate; MTBF 3.75 years	2.4.3
Exact posterior of the rate	$\mathrm{Gamma}(a + k,\ b + T_{\mathrm{tot}})$	–, yr	the same work-order log, conjugate	2.4.3
A season without a work order	$\sum_j w_j\, \mathrm{e}^{-\lambda_{\mathrm{ref}} \mathrm{e}^{\omega_j} \tau} = 0.9289$	–	reported, averaged over the normal of the log-rate	2.4.4
Its exact value	$(b'/(b' + \tau))^{a'} = 0.9369$	–	the same average over the gamma posterior; gap 0.0079	2.4.4
Plug-in MTBF	$\mathrm{e}^{-\mathbb{E}[\lambda \mid \mathbf{y}]\tau}$	–	the same at a point estimate; below the exact value	2.4.5
Two replacements in a year	$\Pr(F_A \mid \mathbf{y})\ \Pr(F_B \mid F_A, \mathbf{y})$	–	intersection; product rule; a shared supply	2.4.7
Model posterior	$\Pr(M_j \mid \mathbf{y}) \propto p(\mathbf{y} \mid M_j)\ \Pr(M_j)$	–	probability of a model	2.5.1
Bayes factor of the air path	$B_{AB} = p(\mathbf{y} \mid M_{\mathrm{fouled}})/p(\mathbf{y} \mid M_{\mathrm{clean}})$, $M_A = M_{\mathrm{fouled}}$, $M_B = M_{\mathrm{clean}}$	–	relative support from the evidence	2.5.1
Expected loss of an action	$\mathcal{R}(a) = c_a + C_e\ \Pr(e \mid \mathbf{y}, a)$	R$	a filter change against waiting	2.5.4

Term	Expression	Units	Meaning	Defined in
Break-even probability	$\Pr(e \mid \mathbf{y}, \text{change}) + c_{\text{change}}/C_e$	–	the probability at which acting and waiting cost the same	2.5.4
Optimal action	$a^* = \arg\min_a \mathbb{E}[L(a, \mathbf{z}) \mid \mathbf{y}]$	–	decision by expected loss	2.5.5
Readings of a condition test	$\mathbf{\Sigma}_a = \mathbf{\Sigma}/(1 + 0.92384\, n)$	varies	every entry shrinks along the ridge	2.6.1
Value of information	$\text{EVSI} = \bar{L}_{\text{now}} - \bar{L}_{\text{after}}$	R$	what an inspection improves	2.6.2
Raw-performance trend	the approach $T - T_c$ trended as condition	K	the naive baseline; no band, mixes load and supply temperature	2.6.5
The book's questions				
Drill questions	twenty-one asset-management questions	–	general form, category, computation and chapter of each	2.3.8

2.8 Exercises

Exercise 2.1 (Formulation)**.** For each statement, name the category of the question, the expression that answers it and the computation beneath it. (a) What is the probability that a coil passes the next 12 months without a corrective work order? ($\Pr(T > 12 \text{ months} \mid \mathbf{y})$, a posterior predictive probability over the stated horizon; prediction under a threshold; the posterior.) (b) What is the work-order rate of a coil design, given its maintenance records? ($p(\lambda \mid \mathbf{y})$ or a summary of it; estimation; the posterior.) (c) Are two replacements in the same year more likely to have been caused by a common fault on the control supply? ($\Pr(M_{\text{common}} \mid \mathbf{y})$ against $\Pr(M_{\text{independent}} \mid \mathbf{y})$, or their Bayes factor; hypothesis comparison; the evidence.) (d) Which inspection interval minimizes the expected maintenance cost? ($\arg\min_a \mathbb{E}[L(a, \mathbf{z}) \mid \mathbf{y}]$ over the candidate intervals; a decision, in the design category; the expected consequence.) (e) Is the hall's health index falling, and in which year will it reach its functional limit h_F? ($\Pr(\gamma < 0 \mid \mathbf{y})$ and $\Pr(t^* \le s \mid \mathbf{y}) = \Pr(\gamma < 0,\ g_s \le \ln h_F \mid \mathbf{y})$ over a state that carries the log-conductance and its trend; prediction under a threshold; the trend.) (f) Of the questions A to D of Table 2.1, which two are answered from the same posterior without a new calculation of it? (A and B: an estimation and a threshold probability read from one posterior.)

Exercise 2.2 (Formulation)**.** The coil design of Section 2.4.3 has a work-order rate λ with a known posterior, and the owner requires at least a 99 % probability of passing the 90-day summer season, $\tau = 0.2466\,\text{yr}$, without a corrective work order. (a) Formulate the posterior predictive probability of a quiet season and the probability that the design's underlying rate meets the requirement. ($\mathbb{E}[\mathrm{e}^{-\tau\lambda} \mid \mathbf{y}] = \int_0^\infty \mathrm{e}^{-\tau\lambda} p(\lambda \mid \mathbf{y})\,\mathrm{d}\lambda$; $\Pr(\mathrm{e}^{-\tau\lambda} \ge 0.99 \mid \mathbf{y}) = \Pr(\lambda \le -\ln 0.99/\tau \mid \mathbf{y}) = \Pr(\lambda \le 0.04076\,\text{yr}^{-1} \mid \mathbf{y})$.) (b) Evaluate both for the posterior of the log-rate, $\hat{\omega} = \ln(8/3)$ and $\sigma_\omega = 0.5$, and beside it for the exact Gamma$(4, 15\,\text{yr})$, using for the second the integer-shape identity $\Pr(\lambda \le x \mid \mathbf{y}) = 1 - \sum_{j=0}^{a'-1} \mathrm{e}^{-b'x}(b'x)^j/j!$. (From the state, the season is 0.9289 by equation (2.41), and $\Pr(\lambda \le x \mid \mathbf{y}) = \Phi((\ln(0.40760) - 0.98083)/0.5) = \Phi(-3.757) = 0.00009$; exactly, $(15/15.2466)^4 = 0.9369$, and $b'x = 0.6114$ with the four Poisson terms 0.5426, 0.3317, 0.1014 and 0.0207 leaves 0.0036.) (c) Explain why the two questions differ. (The first averages the probability of a quiet season over the uncertainty in the rate and asks how likely this season is to pass; the second asks how much posterior probability supports the requirement on the rate itself. Here the season is likely to pass, 0.9289, while the requirement is almost certainly not met, because the posterior puts almost no mass below $0.04076\,\text{yr}^{-1}$: 0.00009 from the state and 0.0036 exactly.)

Exercise 2.3 (Derivation)**.** The hall obeys $T = T_c + Q/G$ with $T_c = 18°\mathrm{C}$ and $Q = 60\,\mathrm{kW}$. (a) With the state, rows and $\mathbf{\Sigma}$ of equations (2.30) to (2.33), find the probability that the full-load temperature remains below 29.75°C from the normal of the entry T_2, then the tail of g under the commissioning row from the decreasing map between G and T_2 along the ridge, and say what an assurance level $\kappa = 0.95$ decides. ($\Pr(T_2 < 29.75°\mathrm{C} \mid \mathbf{y}) \approx \Phi\big((29.75 - \hat{z}_2)/\sqrt{\Sigma_{22}}\big) = \Phi\big((29.75 - 28.1402)/0.9612\big) = \Phi(1.675) = 0.953$. By the tail of g, along $\mathbf{v}$ the bound on T_2 is $G > Q_2/(29.75 - T_c) = 60/11.75 = 5.1064\,\mathrm{kW/K}$, that is $g > -0.161268$ on the third entry, so $\Pr(T_2 < 29.75°\mathrm{C} \mid \mathbf{y}) = \Phi\big((\hat{z}_3 + 0.161268)/\sqrt{\Sigma_{33}}\big) = \Phi(1.554) = 0.940$; the normal accepts at $\kappa = 0.95$ and the tail of g does not, and both columns of its 29.75°C row of Table 2.6 are stated.) (b) Now let both G and the supply air T_c be uncertain, as when the discharge controller saturates, with a joint posterior $p(G, T_c \mid \mathbf{y})$, and derive an integral for $\Pr(T > T_{\max} \mid \mathbf{y})$ without assuming that G and T_c are independent. (The event is $T_c + Q/G > T_{\max}$, so $\Pr(T > T_{\max} \mid \mathbf{y}) = \iint \mathbf{1}\big[T_c + Q/G > T_{\max}\big]\, p(G, T_c \mid \mathbf{y})\, \mathrm{d}G\, \mathrm{d}T_c$; the supply air becomes an entry of the state, and the joint density is required because the readings can leave it and the conductance correlated, a nonzero off-diagonal entry of $\mathbf{\Sigma}$.)

Exercise 2.4 (Derivation)**.** A sister unit's fan motor and valve actuator have probabilities $\Pr(F_A) = 0.08$ and $\Pr(F_B) = 0.04$ of needing replacement in the coming year. (a) Assuming independence, derive the probability that both are needed and the probability that the year needs at most one. ($\Pr(F_A \cap F_B) = 0.08 \times 0.04 = 0.0032$; $\Pr(\text{at most one} \mid \mathbf{y}) = 1 - 0.0032 = 0.9968$.) (b) Now let both share a control-power supply that surges with probability $q_S = 0.01$ a year, with the marginals unchanged. Derive the joint probability from equation (2.46), the conditional probability $\Pr(F_B \mid F_A)$ and the probability of at most one replacement. ($q_A = 0.07/0.99 = 0.07071$, $q_B = 0.03/0.99 = 0.03030$; $\Pr(F_A \cap F_B) = 0.01 + 0.99 \times 0.07071 \times 0.03030 = 0.01212$; $\Pr(F_B \mid F_A) = 0.01212/0.08 = 0.152$; $\Pr(\text{at most one} \mid \mathbf{y}) = 0.98788$, the joint probability 3.79 times the independent value.)

Exercise 2.5 (Derivation)**.** The log-conductance of the hall is carried as a straight line in time, $g_t = \alpha + \gamma(t - \bar{t})$ with $\bar{t} = 1$ year, and the fit to the hall's two tests of Section 2.3.4 gives $\hat{\mathbf{z}} = (-0.054618,\ -0.040692)^{\mathsf{T}}$ and $\mathbf{\Sigma}$ with entries $\Sigma_{\alpha\alpha} = 4.312270 \times 10^{-3}$, $\Sigma_{\alpha\gamma} = -1.800393 \times 10^{-4}$ and $\Sigma_{\gamma\gamma} = 4.312270 \times 10^{-3}$. The functional limit is $g_F = \ln(5/6) = -0.182322$. (a) From the trend form (2.21) for a falling quantity under the Laplace posterior, find the probability that the conductance is falling. ($\Phi(0.040692/\sqrt{4.312270 \times 10^{-3}}) = \Phi(0.6197) = 0.7323$.) (b) Find the crossing year at the MAP state and say why it lies near, but not at, the median of the crossing time. ($t^*(\hat{\mathbf{z}}) = \bar{t} + (g_F - \hat{\alpha})/\hat{\gamma} = 4.14$ years. The second argument of equation (2.22) is zero there, so the single tail is one half; the first passage subtracts the states that stand below the limit with $\gamma \geq 0$, a mass of 0.0010, and is 0.4990 there, so the median is year 4.15.) (c) Find the probability that the hall has crossed the limit by year 5. ($\mathbf{g}(5) = (1,\ 4)^{\mathsf{T}}$, $\mathbf{g}(5)^{\mathsf{T}}\hat{\mathbf{z}} = -0.217387$, $\mathbf{g}(5)^{\mathsf{T}}\mathbf{\Sigma}\,\mathbf{g}(5) = 4.312270 \times 10^{-3} + 2 \times 4 \times (-1.800393 \times 10^{-4}) + 4^2 \times 4.312270 \times 10^{-3} = 7.186828 \times 10^{-2}$, and the covariance of γ and g_5 is $\Sigma_{\alpha\gamma} + 4\,\Sigma_{\gamma\gamma} = 1.706904 \times 10^{-2}$, a correlation of 0.9696. The single tail is $\Phi\big((-0.182322 + 0.217387)/0.26808\big) = \Phi(0.1308) = 0.5520$, and the first passage is $\Pr(\gamma < 0,\ g_5 \leq g_F \mid \mathbf{y}) = \Phi_2(0.6197,\ 0.1308;\ 0.9696) = 0.5512$.) (d) Explain why the crossing time has no posterior mean, and why its band by inversion has no upper end. ($t^* = \bar{t} + (g_F - \alpha)/\gamma$ divides by the trend entry, whose Gaussian posterior has a positive density at $\gamma = 0$, and the integral of $1/\gamma$ against such a density diverges; and the distribution function rises only toward $\Pr(\gamma < 0 \mid \mathbf{y}) = 0.7323$, so its 0.975 point does not exist, a mass 0.2677 never crossing.)

Exercise 2.6 (Application)**.** The hall has a probability of 0.004 of a forced outage before the next planned shutdown, and each forced outage costs R$ 250 000. A work order on the cooling unit now costs R$ 400 and reduces that probability to 0.001. (a) Derive the expected total cost with and without the work order, and state which action minimizes it. (The expected consequence (2.23) of each action, probability times consequence plus the certain cost. Without:

$0.004 \times \text{R\$}\,250\,000 = \text{R\$}\,1000$; with: $\text{R\$}\,400 + 0.001 \times \text{R\$}\,250\,000 = \text{R\$}\,650$; the work order minimizes expected cost, with an expected saving of R$ 350.) (b) Find the cost of the work order at which the decision would reverse. ($(0.004 - 0.001) \times \text{R\$}\,250\,000 = \text{R\$}\,750$.)

Exercise 2.7 (Application)**.** Two models are competing explanations of the hall's warmer condition test, fouling of the air path (M_{fouled}) and a bias of the hall sensor (M_{bias}), with $\Pr(M_{\text{bias}}) = \Pr(M_{\text{fouled}}) = 0.5$ and evidences $p(\mathbf{y} \mid M_{\text{bias}}) = 0.02$ and $p(\mathbf{y} \mid M_{\text{fouled}}) = 0.08$. (a) Derive the Bayes factor B_{fb} and the posterior probability $\Pr(M_{\text{fouled}} \mid \mathbf{y})$. ($B_{\text{fb}} = 0.08/0.02 = 4$, $\ln B_{\text{fb}} = 1.386$ nats; $\Pr(M_{\text{fouled}} \mid \mathbf{y}) = 0.08 \times 0.5/(0.08 \times 0.5 + 0.02 \times 0.5) = 0.80$, assuming that the two models exhaust the possibilities under consideration.) (b) Repeat with $\Pr(M_{\text{fouled}}) = 0.2$, and say what the comparison of the two answers shows. (The Bayes factor is the same 4 in both cases, because it depends only on the evidence, while the posterior model probability also depends on the prior: $0.08 \times 0.2/(0.08 \times 0.2 + 0.02 \times 0.8) = 0.50$.)

Chapter 3

Uncertainty and Conditional Probability

3.1 Uncertainty and evidence in asset-management questions

The air path of the hall is handed over with one commissioning test, window 0 of its condition record. The test runs the hall at its 30 kW with the supply air at its 18°C set point, and the hall sensor reads 23.2°C. The health-index dashboard takes the reading as exact, $G = 30/(23.2 - 18) = 5.769\,\text{kW/K}$, and sets it against the 6 kW/K of the design sheet: an index of 0.962, "3.8 % short of design". Its condition flag fires before handover, and the flag's data sheet says that it catches nine fouled coils in ten. Read from the dashboard alone, the path was handed over 3.8 % short of design, the coil is fouled with probability 0.9, and a cleaning work order goes out before handover. The fan sits behind the same filter as the coil, and the planner wonders whether the crew should open the fan too.

The dashboard settles none of this. One test scatters: a second test of the same coil would read another temperature and report another index. The data sheet states how often the flag fires on a fouled coil, Pr(flag | fouled), which is not how often a flagged coil is fouled, Pr(fouled | flag); and nothing on the sheet or the screen says how often coils of this design arrive fouled, or what the coil and the fan share. Chapter 1 represented the hall as a plant graph whose free quantities form the state vector $\mathbf{z}$, and Chapter 2 wrote the asset manager's questions as probabilities over that state, with the unknown on the left of the conditioning bar and the evidence on the right. The question of this chapter is the mechanism between the two:

how new evidence changes what is known about an asset, and how dependencies between components and between physical quantities affect a probability calculation.

Jargon: aleatory and epistemic uncertainty, conditional probability, prior, likelihood, evidence, posterior, marginalization, nuisance quantity, conditional independence, common cause

Aleatory uncertainty is the variability of outcomes that the model treats as random even when its parameters are known, such as the temperature the next condition test of an unchanged coil will read, or the year in which a unit with a known rate next needs a corrective work order. *Epistemic* uncertainty is incomplete knowledge of quantities that have one true value, such as the conductance of the coil's air path or the calibration offset of the hall sensor; more evidence can reduce it. A *conditional probability* $\Pr(A \mid B)$ is the probability of event A among the situations in which event B occurs. In Bayes' theorem the *prior* is the probability of a hypothesis before an observation, the *likelihood* is the probability of the observation under the hypothesis, the *evidence* (also *marginal likelihood*) is the probability of the observation averaged over the hypotheses, and the *posterior* is the probability of the hypothesis after the observation. *Marginalization* removes an unknown from a joint distribution by summing or integrating over it; an unknown removed this way because the question does not ask about it is a *nuisance quantity*. Two quantities are *conditionally independent* given a third when, once the third is known, learning one says nothing more about the other. A *common cause* is a shared condition, such as a missed filter change, that raises the probability of wear in several components at once.

Key idea

Conditioning on evidence restricts attention to the situations consistent with it and renormalizes; Bayes' theorem does this for a hypothesis, marginalization averages over every unknown the question does not name, and the dependence among unknowns, through a common cause or a shared physical relation, is what lets one observation inform another quantity and what an independence assumption silently discards.

An asset record carries two kinds of not knowing: what the next condition test will read, and what the asset itself is. Keeping them apart changes the probability that a test reads a coil below its requirement, and each kind has its own distribution.

3.1.1 The scatter of a condition test

Suppose the coil were unchanged since commissioning and its conductance known exactly: take the commissioned value of Chapter 1, $G = 5.917\,\mathrm{kW/K}$ ($h = 0.986$), as if it were the truth, so that at 30 kW the hall air is $T = 18 + 30/5.9171 = 23.07^\circ\mathrm{C}$. Even then the reading of the next condition test cannot be predicted. Let y_T be the reading of one test. The sensor adds an error that changes from test to test, $y_T \mid T \sim \mathcal{N}(T, \sigma_T^2)$ with $\sigma_T = 0.5\,\mathrm{K}$, so the probability that one test reads above a level c is the integral of the normal density from c to infinity,

$$\Pr(y_T > c \mid T) = \int_c^\infty \frac{1}{\sigma_T}\,\varphi\!\left(\frac{x - T}{\sigma_T}\right) \mathrm{d}x = 1 - \Phi\!\left(\frac{c - T}{\sigma_T}\right), \tag{3.1}$$

where x is the variable of integration and φ and Φ are the standard normal density and cumulative distribution function. A test read as exact turns its reading into a conductance, $G_{\text{test}} = Q/(y_T - T_c)$, and it calls the coil below the conductance $G^* = 30/(24 - 18) = 5.0\,\mathrm{kW/K}$ that holds 24°C at 30 kW exactly when $y_T > T_c + Q/G^* = 24^\circ\mathrm{C}$:

$$\begin{aligned}\Pr(G_{\text{test}} < 5.0\,\mathrm{kW/K} \mid T) &= \Pr(y_T > 24^\circ\mathrm{C} \mid T = 23.07^\circ\mathrm{C}) \\ &= 1 - \Phi\!\left(\frac{24 - 23.0701}{0.5}\right) = 1 - \Phi(1.8598) = 0.0315.\end{aligned}$$

With the coil perfectly known and unchanged there is still a probability 0.0315, one test in 32, that a test reads it below its requirement. This uncertainty arises from the variability that the test itself represents, and it is called *aleatory*. In conductance the scatter is wide: by the delta method on $G_{\text{test}} = Q/(y_T - T_c)$ one test scatters by $G^2\sigma_T/Q = 5.9171^2 \times 0.5/30 = 0.5835\,\text{kW/K}$, 9.9 % of G. A test reads the coil better than its design sheet, $G_{\text{test}} > 6\,\text{kW/K}$, when $y_T < 23°\text{C}$, with probability $\Phi\big((23 - 23.0701)/0.5\big) = \Phi(-0.1402) = 0.4442$: nearly one test in two calls a coil that has not changed better than design.

A unit's next corrective failure is a second aleatory quantity. Suppose the unit's corrective work orders arrive at a known constant rate $\lambda = 0.2\,\text{yr}^{-1}$. Let T_f be the time to the next corrective failure in years, measured from now, and let it be exponential given the rate, $T_f \mid \lambda \sim \text{Exponential}(\lambda)$, with density $\lambda\,\text{e}^{-\lambda t}$ for $t \geq 0$. The probability that the unit runs beyond t years without one is the integral of that density from t to infinity,

$$\Pr(T_f > t \mid \lambda) = \int_t^\infty \lambda\,\text{e}^{-\lambda x}\,\text{d}x = \text{e}^{-\lambda t}. \tag{3.2}$$

For the next year, $\Pr(T_f > 1\,\text{yr} \mid \lambda = 0.2\,\text{yr}^{-1}) = \text{e}^{-0.2} = 0.8187$: with a perfectly known rate there is still a probability $1 - 0.8187 = 0.1813$ that the unit needs a corrective work order within the year.

Figure 3.1(a) shows the scatter of the test. The horizontal axis numbers twelve later condition tests of the same unchanged coil, and the vertical axis is the conductance each test reports when its reading is taken as exact. The dots are $Q/(y - T_c)$ at the twelve quantiles of the normal distribution $\mathcal{N}(23.07, 0.5^2)$ of the reading, placed in an arbitrary order; the dashed line marks the true conductance and the red line the requirement G^*. Twelve tests of one coil report from 5.054 to 7.136 kW/K, a dashboard index from 0.842 to 1.189, none of them below 5.0 kW/K and five of them better than design: knowing the coil fixes the distribution of the test, not the test.

3.1.2 What is not known about the coil

Now suppose the conductance is not known. Before the commissioning test the handbook gives 7.5 kW/K, right to a factor of about 1.35. What is known about the conductance is then a probability distribution given the evidence, $p(G \mid \mathbf{y})$, or equivalently $p(g \mid \mathbf{y})$ for the log-conductance $g = \ln(G/G_{\text{ref}})$ with $G_{\text{ref}} = 6\,\text{kW/K}$ the design conductance. The handbook prior is $g \sim \mathcal{N}(\mu_g, 0.3^2)$ with $\mu_g = \ln(7.5/6) = 0.2231$, and $\ln 1.35 = 0.3001$ is the factor in the log. It makes the conductance plausibly $7.5/1.35 = 5.556$, 7.5 or $7.5 \times 1.35 = 10.125\,\text{kW/K}$, puts 95 % of its probability between 4.166 and 13.503 kW/K, and gives $\Pr(G < 5.0\,\text{kW/K}) = 0.0883$ before any test; as a health index against design the prior has the median 1.25 and the 95 % range 0.694 to 2.250. This is *epistemic* uncertainty: uncertainty caused by incomplete knowledge of the characteristics of the asset, which additional observations can reduce.

The commissioning test is the evidence that reduces it. Chapter 1 fits window 0 with three rows over the state $(T, g)^\mathsf{T}$ and reaches the commissioned value $\hat{g}_0 = -0.013916$ at $\sigma_p = 0.01\,\text{K}$ (-0.013926 in the limit of exact physics). The Laplace normal of g puts the median of the conductance at 5.917 kW/K and its 95 % band at 4.914 to 7.125 kW/K: a health index $h = 0.986$ (0.819 to 1.188) against design. It is the normal centered on the most probable state, not the posterior's own median: the exact posterior of the same rows has the mean -0.0013 for g, against the mode -0.0139. The posterior spread of g falls from 0.3 to 0.0948, 0.316 of the prior's, and later condition tests are read against this commissioned value. The calibration of a sensor is the other common case. A hall sensor that reads high by a fixed, unknown offset reads high in every test; the offset has one true value, and the evidence that reduces what is known about it is a reference instrument brought to an inspection, never more tests from the same sensor.

Figure 3.1(b) draws both distributions. The horizontal axis is the conductance in kW/K and the vertical axis the density per kW/K. The blue curve is the handbook prior, drawn as the lognormal density of G, and its three dots are the plausible values named above; the green curve is the posterior after the commissioning test, lognormal with the median 5.917 kW/K and the log-spread 0.0948; the red line is the requirement $G^* = 5.0$ kW/K. The prior's peak is 0.1855 per kW/K and the posterior's 0.7144. The conductance is a single number; the spread of each curve is the spread of knowledge about it, and one test has narrowed it to the band that the requirement line cuts in its lower tail.

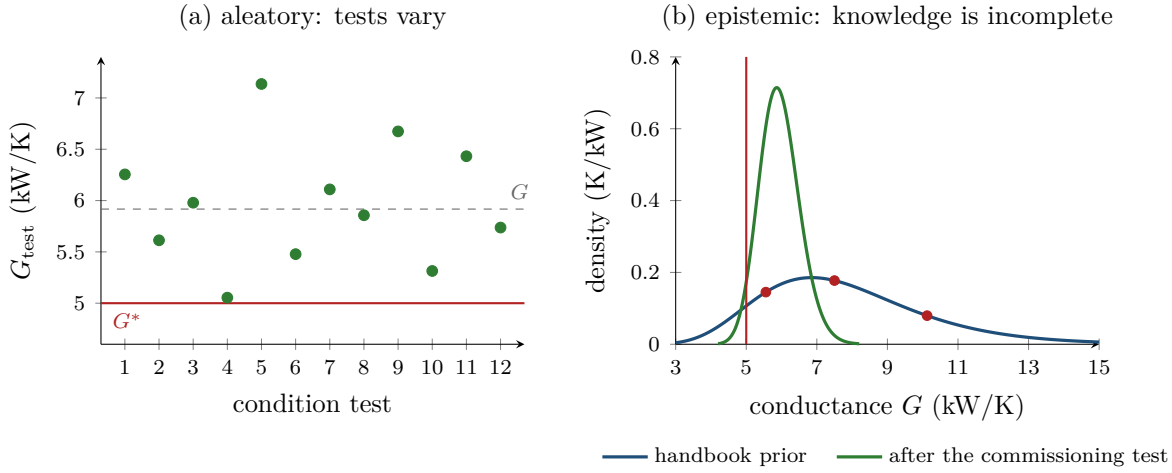


Figure 3.1: The two kinds of uncertainty. (a) Conductances reported by twelve condition tests of a coil known to be unchanged at $G = 5.917$ kW/K (dashed), each reading taken as exact, $G_{\text{test}} = Q/(y - T_c)$ with y the normal quantiles $T + 0.5\,\Phi^{-1}((i - 0.5)/12)$, $i = 1, \ldots, 12$, $T = 23.07$°C, in an arbitrary order; red: the requirement $G^* = 5.0$ kW/K. (b) The handbook prior of the conductance, $\ln(G/6) \sim \mathcal{N}(\ln 1.25, 0.3^2)$, density $\exp[-\ln^2(G/7.5)/0.18]/(0.3\sqrt{2\pi}\,G)$, with the plausible values 5.556, 7.5 and 10.125 kW/K (dots); and the posterior after the commissioning test, lognormal with median 5.917 kW/K and log-spread 0.0948; red: G^*.

3.1.3 The epistemic part and the scatter of the next test

In an asset record the distinction can depend on what information is available and on how the model is defined. The fouling a coil will gather over the next summer may be modeled as random, aleatory, while its conductance today is a fixed but poorly measured quantity, epistemic. The distinction is therefore a statement about the model, and it changes the calculation.

If the coil is known, the probability that the next test reads it below G^* is equation (3.1) at $c = 24$°C. If the coil is uncertain, the probability that the next reading y_{new} crosses c is the conditional probability averaged over what is known about the hall, which is the law of total probability derived in Section 3.4.2:

$$\Pr(y_{\text{new}} > c \mid \mathbf{y}) = \int \Pr(y_{\text{new}} > c \mid T)\, p(T \mid \mathbf{y})\, \mathrm{d}T = \int \left[1 - \Phi\left(\frac{c - T}{\sigma_T}\right)\right] p(T \mid \mathbf{y})\, \mathrm{d}T. \qquad (3.3)$$

For three plausible values the integral is a sum. Before any test, the three conductances of Section 3.1.2 put the hall air at $T = 18 + 30/G = 23.40$, 22.00 and 20.96°C. With the illustrative probabilities 0.25, 0.50 and 0.25 on them, the probability that the next test reads the coil below 5.0 kW/K is

$$\Pr(y_{\text{new}} > 24\text{°C}) = 0.25 \times 0.1151 + 0.50 \times (3.2 \times 10^{-5}) + 0.25 \times (6 \times 10^{-10}) = 0.0288,$$

against 1.5×10^{-5} at the mean conductance 7.670 kW/K, where $T = 21.91$°C: the average is 1953 times the value at the average, and almost all of it comes from the low conductance. For a Gaussian posterior $T \mid \mathbf{y} \sim \mathcal{N}(\mu, \Sigma_{TT})$ the integral has a closed form. The next reading is $y_{\text{new}} = T + \epsilon$ with the error $\epsilon \sim \mathcal{N}(0, \sigma_T^2)$ independent of T, and the sum of two independent normal quantities is normal with the sum of their variances, a standard result, so that

$$y_{\text{new}} \mid \mathbf{y} \sim \mathcal{N}(\mu,\ \Sigma_{TT} + \sigma_T^2), \qquad \Pr(y_{\text{new}} > c \mid \mathbf{y}) = 1 - \Phi\left(\frac{c - \mu}{\sqrt{\Sigma_{TT} + \sigma_T^2}}\right). \tag{3.4}$$

After the commissioning test, $\mu = 23.07$°C and $\sqrt{\Sigma_{TT}} = 0.481$ K, the predictive standard deviation is $\sqrt{0.4806^2 + 0.5^2} = 0.6935$ K, and the probability that the next test reads the coil below 5.0 kW/K is $1 - \Phi(1.341) = 0.0900$, against 0.0315 for the plug-in that takes the commissioned value as the truth. In the three-value sum the average of the tail exceeds the tail at the average because $1 - \Phi((c - T)/\sigma_T)$ is convex and increasing in T below the limit, where all three values lie, and $T = T_c + Q/G$ is itself convex in G: a convex increasing function of a convex function is convex, so both steps act in the same direction and the tail is convex in G. In the Gaussian case the reason is that the limit lies above the mean: widening the spread from 0.5 to 0.6935 K moves probability past it.

The scatter of the next test remains even when the uncertainty about the coil is removed. Table 3.1 repeats the commissioning fit of Chapter 1 with n readings at 30 kW in the commissioning window whose readings average 23.2°C, so that the state $(T, g)^{\mathsf{T}}$ carries one physics row, n reading rows and the handbook prior row, for a coil unchanged across the readings. The first column is the number of readings; the next four are the estimated conductance, the health index, the posterior spread of g and the 95 % band of G; and the last two are the probability that the coil is below its requirement and the probability that the next condition test reads it below. The spread of g halves each time the readings quadruple, from 0.0948 to 0.0122 at $n = 64$, and the index settles on 0.962, the dashboard's figure of the opening scene: that figure is the limit of many readings averaging 23.2°C, not the reading of one. The probability that the coil is below its requirement falls from 0.0265 to below 10^{-30}; the probability that the next test reads it below tends to $1 - \Phi(1.6) = 0.0548$, not to zero. The sensor's error is taken as independent from reading to reading, with no fixed offset; with a shared offset the readings stop adding information as they become correlated, and at a correlation of one a second test adds nothing (Section 3.5.7).

Table 3.1: The coil at 30 kW after n readings in the commissioning window whose readings average 23.2°C (the fit of Chapter 1 with n reading rows): the epistemic spread of the conductance shrinks, the scatter of the next test does not. The probability that $G < 5.0$ kW/K is read on the entry T, as $\Pr(T > 24\text{°C} \mid \mathbf{y})$.

n [–]	$\hat{G}$ [kW/K]	$\hat{h}$ [–]	$\sqrt{\Sigma_{gg}}$ [–]	95 % band of G [kW/K]	$\Pr(G < 5.0 \mid \mathbf{y})$ [–]	$\Pr(\text{next test} < 5.0 \mid \mathbf{y})$ [–]
1	5.917	0.986	0.0948	4.914–7.125	0.0265	0.0900
4	5.808	0.968	0.0480	5.286–6.380	3.8×10^{-4}	0.0674
16	5.779	0.963	0.0241	5.512–6.058	4.4×10^{-11}	0.0583
64	5.772	0.962	0.0122	5.636–5.911	4.7×10^{-38}	0.0557
∞	5.769	0.962	0	5.769	0	0.0548

The event $G < 5.0$ kW/K is the event $T > 24$°C: with the load and the supply air known inputs of the test, the coil is below its requirement exactly when the hall air is above its limit. The table reads it on the entry T, by the Laplace normal of T. The exact posterior of the same rows, with T integrated out and a quadrature along g, sets the exact value beside each. The probability that the coil is below its requirement is 0.0265 against the exact 0.0226 at $n = 1$,

3.76×10^{-4} against 3.30×10^{-4} at $n = 4$, and 4.43×10^{-11} against 4.45×10^{-11} at $n = 16$, relative gaps of 17.5, 14.1 and $-0.4\,\%$. The probability that the next test reads it below is 0.0900 against 0.0813, 0.0674 against 0.0648, and 0.0583 against 0.0576. The source is the curve of the physics row: the Laplace normal is centered on the mode, 23.0701°C at $n = 1$, and the exact posterior has the mean 23.0300°C, so the exact tail above the limit is the thinner one. The same event read on the Laplace normal of g instead gives 0.0378 at $n = 1$, farther from the exact value: once the physics row is curved, the normal of one entry is not the normal of the other, and Chapter 4 shows how the choice of entry enters.

Principle 3.1. *Collecting more condition tests can reduce the uncertainty about the coil without eliminating the scatter of the next test: evidence removes epistemic uncertainty, and the aleatory part remains as long as the reading is modeled as scattering from test to test.*

3.1.4 Distributions of what is uncertain

Three mathematical objects represent what is uncertain, and Figure 3.2 draws one of each.

A discrete unknown. Suppose the fan of a cooling unit (illustrative) is either sound or worn, and define the indicator $F = 1$ if it is worn and $F = 0$ if it is sound. The assignment $\Pr(F = 1) = 0.10$, $\Pr(F = 0) = 0.90$ is a discrete probability distribution: each value receives a probability, and the probabilities sum to one. Panel (a) of the figure has the two values on the horizontal axis and their probabilities as bars.

A continuous unknown. Suppose the health index $h = G/G_{\text{ref}}$ of a coil of this design inspected late in its life (illustrative) is unknown, and represent it by a density $p(h)$. A probability over a range is the integral of the density over that range,

$$\Pr(0.75 < h < 0.85) = \int_{0.75}^{0.85} p(h)\,\mathrm{d}h.$$

For a continuous variable the density itself is not a probability: it is a probability per unit of the variable, and it can exceed one. Panel (b) uses the illustrative density $\mathcal{N}(0.80, 0.04^2)$. The horizontal axis is the health index and the vertical axis the density, whose peak is $1/(0.04\sqrt{2\pi}) = 9.97$; the shaded area between 0.75 and 0.85 is $\Phi(1.25) - \Phi(-1.25) = 0.789$, the probability of that range.

A joint distribution. Suppose an illustrative hall with a weak air path has two unknown quantities, the conductance G of its air path and the supply-air temperature T_c, as it does when the discharge controller's own sensor is not trusted. Their *joint* distribution $p(G, T_c)$ describes the uncertainty about both together, including any dependence between them. The joint distribution is essential when a quantity depends on several uncertain inputs, as the hall air does,

$$T = T_c + \frac{Q}{G},$$

with the heat load Q in kW and G in kW/K. The individual distributions $p(G)$ and $p(T_c)$ are not always enough to determine the distribution of T; how G and T_c vary together is needed as well. Panel (c) draws two joint distributions with the same marginals, G with mean 3 kW/K and standard deviation 0.15 kW/K, T_c with mean 20°C and standard deviation 1 K, one with correlation $\mathrm{corr}(G, T_c) = 0$ and one with $\mathrm{corr}(G, T_c) = 0.8$. The horizontal axis is G, the vertical axis T_c, and each ellipse encloses 95 % of its distribution. Both ellipses span the same ranges of G and of T_c, so the marginals cannot tell them apart; Section 3.6.4 shows that at $Q = 30$ kW the hall air has a standard deviation of 1.12 K under the first and 0.67 K under the second.

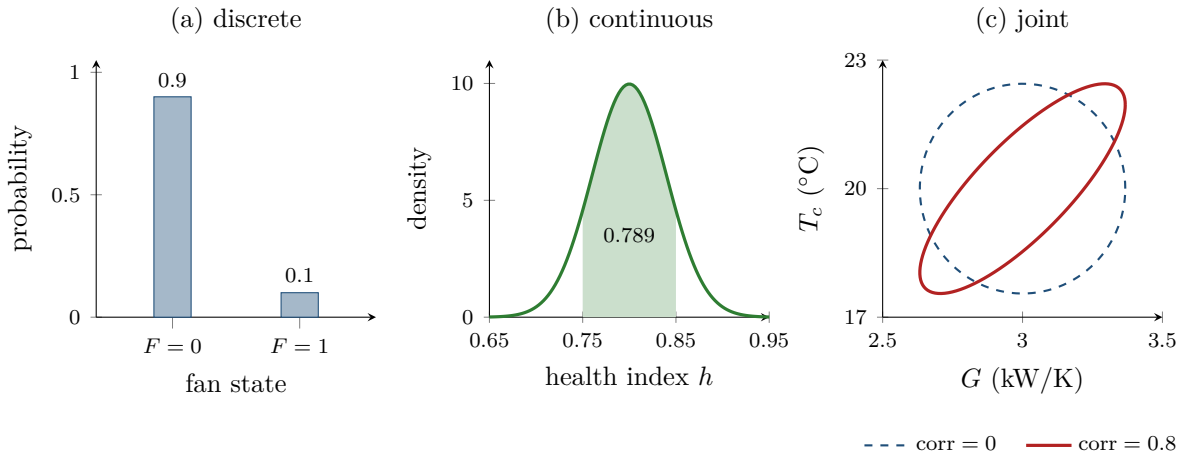


Figure 3.2: Distributions of what is uncertain (illustrative). (a) A discrete unknown: $\Pr(F = 1) = 0.10$ for a worn fan. (b) A continuous unknown: the density $\mathcal{N}(0.80, 0.04^2)$ of a coil's health index against design; shaded, $\Pr(0.75 < h < 0.85) = 0.789$. (c) Two joint distributions of G and T_c with the same normal marginals (means 3 kW/K and 20°C, standard deviations 0.15 kW/K and 1 K) and correlations 0 and 0.8; the ellipses enclose 95 % of each ($\chi^2_{2,0.95} = 5.991$).

3.2 Conditional probability and Bayes' theorem

Conditional probability is the foundation of every update in this book. An inspection record, read work order by work order, shows where it comes from, why the direction of the conditioning matters, and how Bayes' theorem reverses it.

3.2.1 Conditional probability from an inspection record

Let H be the event that the unit controller's approach-temperature alarm, a cruder signal than the dashboard's condition flag, fires on a coil in the month before its inspection, and F the event that the inspection finds the coil fouled. The alarm compares the approach temperature with a fixed set point, and the approach falls with the load: a fouled coil running at part load stays under the set point, which is why the alarm misses half the fouled coils. The question *given that the alarm fired, what is the probability that the coil is fouled* is written $\Pr(F \mid H)$. The phrase "given that" restricts attention to the situations in which H occurred.

Imagine 1000 inspections of coils of this design, each preceded by a month of the alarm, recorded in Table 3.2. The rows say whether the alarm fired, the columns whether the inspection found the coil fouled, and the margins are the totals: alarm and fouling together in 20 inspections, the alarm alone in 60, fouling alone in 20, and neither in 900.

Table 3.2: Outcomes of 1000 inspections of coils of this design, each preceded by a month of the unit controller's approach-temperature alarm (illustrative), in numbers of inspections.

[–]	coil fouled [–]	coil clean [–]	total [–]
alarm fired	20	60	80
no alarm	20	900	920
total	40	960	1000

From these counts,

$$\Pr(H) = \frac{80}{1000} = 0.08, \qquad \Pr(F) = \frac{40}{1000} = 0.04, \qquad \Pr(H \cap F) = \frac{20}{1000} = 0.02.$$

Among the 80 inspections that followed an alarm, the coil was fouled in 20. Therefore

$$\Pr(F \mid H) = \frac{20}{80} = 0.25.$$

The conditional probability is 25 %, the share of the alarm's work orders that find fouling, its *positive predictive value*, against an unconditional probability of 4 % that an inspection finds a fouled coil: in this record the alarm and the fouling are strongly associated, and still three work orders in four open a clean coil.

3.2.2 The definition

Dividing the numerator and the denominator of 20/80 by the total number of inspections turns the count ratio into a ratio of probabilities,

$$\Pr(F \mid H) = \frac{20/1000}{80/1000} = \frac{\Pr(H \cap F)}{\Pr(H)}.$$

For any two events A and B with $\Pr(B) > 0$ the definition is

$$\Pr(A \mid B) = \frac{\Pr(A \cap B)}{\Pr(B)}. \tag{3.5}$$

Conditioning on B restricts the sample space to the outcomes in B and renormalizes, so that the restricted probabilities again sum to one. The definition requires the conditioning event to have positive probability. Two events are *independent* when $\Pr(A \cap B) = \Pr(A)\,\Pr(B)$, which by equation (3.5) is the same as $\Pr(A \mid B) = \Pr(A)$: learning B does not change the probability of A.

Table 3.3 works the definition for four joint probabilities with the marginals of the record held fixed, $\Pr(H) = 0.08$ and $\Pr(F) = 0.04$. The first column is the probability that an inspection follows an alarm and finds fouling, the next two are the conditional probabilities in the two directions, and the last says what kind of alarm the row describes. The third row is the record of Table 3.2: $\Pr(F \mid H) = 0.02/0.08 = 0.25$ and $\Pr(H \mid F) = 0.02/0.04 = 0.50$. The second row is the independent case, in which $\Pr(F \mid H) = 0.0032/0.08 = 0.04 = \Pr(F)$: an alarm that fires as often on clean coils as on fouled ones carries no information. The joint probability cannot exceed the smaller marginal, 0.04; at that bound every fouled coil raises the alarm.

Table 3.3: Conditional probabilities from equation (3.5) for four joint probabilities of an alarm and a fouled coil, with $\Pr(H) = 0.08$ and $\Pr(F) = 0.04$ held fixed.

$\Pr(H \cap F)$ [–]	$\Pr(F \mid H)$ [–]	$\Pr(H \mid F)$ [–]	Kind of alarm [–]
0	0	0	never fires on a fouled coil
0.0032	0.04	0.08	independent of fouling: no information
0.02	0.25	0.50	the recorded alarm: positively associated
0.04	0.50	1.00	every fouled coil raises the alarm

3.2.3 The multiplication rule

Multiplying equation (3.5), written for $\Pr(B \mid A)$, by $\Pr(A)$ gives the multiplication rule,

$$\Pr(A \cap B) = \Pr(B \mid A)\,\Pr(A). \tag{3.6}$$

Conditioning in the other direction gives $\Pr(A \cap B) = \Pr(A \mid B)\ \Pr(B)$, and the two expressions for the same intersection are equal:

$$\Pr(B \mid A)\ \Pr(A) = \Pr(A \mid B)\ \Pr(B). \tag{3.7}$$

For the record, $0.25 \times 0.08 = 0.02 = 0.50 \times 0.04$. This relation is the starting point of Bayes' theorem.

Reversing the conditioning. The two conditional probabilities of the record differ: $\Pr(F \mid H) = 0.25$, but $\Pr(H \mid F) = 20/40 = 0.50$. The first asks about fouling among the inspections that followed an alarm, the share of the alarm's work orders that find something; the second asks about the alarm among the fouled coils, the alarm's detection rate. Equation (3.7) shows that they are equal only when $\Pr(H) = \Pr(F)$.

Common mistake

Treating $\Pr(A \mid B)$ and $\Pr(B \mid A)$ as the same number. Reversing the conditioning changes the question. In condition monitoring the vendor gives the probability of a flag given a fouled coil, $\Pr(\text{flag} \mid \text{fouled})$, from the design of the flag, and the planner needs the probability of a fouled coil given a flag, $\Pr(\text{fouled} \mid \text{flag})$; the second follows from the first only through Bayes' theorem and the base rate of fouling.

3.2.4 Bayes' theorem

Dividing both sides of equation (3.7) by $\Pr(B) > 0$ gives Bayes' theorem for discrete events,

$$\Pr(A \mid B) = \frac{\Pr(B \mid A)\ \Pr(A)}{\Pr(B)}. \tag{3.8}$$

For the record, the detection rate and the base rates give the share of the alarm's work orders that find fouling, $\Pr(F \mid H) = 0.50 \times 0.04/0.08 = 0.25$, the value read from the counts. In words,

$$\text{posterior} = \frac{\text{likelihood} \times \text{prior}}{\text{evidence}},$$

and the four terms have distinct roles when A is a hypothesis about the asset and B an observation. The *prior* $\Pr(A)$ is what was believed about the hypothesis before the observation. The *likelihood* $\Pr(B \mid A)$ is how probable the observation is when the hypothesis holds. The *evidence* $\Pr(B)$ is the overall probability of the observation, averaged over the hypotheses; it is the normalizing constant of the posterior and, for a whole model M, the quantity the notation writes $p(\mathbf{y} \mid M)$. The *posterior* $\Pr(A \mid B)$ is what is believed about the hypothesis after the observation is accounted for.

3.3 Bayesian diagnosis of fouling from a condition test

Bayes' theorem turns the known behavior of a flag into the probability of the fouling the flag is meant to detect; the answer to the opening scene's work order is far below the flag's detection probability, and a silent flag is evidence too.

3.3.1 The diagnosis problem

The hall's condition flag, the flag of the opening scene, is designed to detect a fouled coil at a condition test. Define the events

$$F = \{\text{the coil is fouled}\}, \qquad H = \{\text{the condition flag fires}\}.$$

The vendor's field data for coils of this design put a fouled coil at 5 % of commissioning tests, $\Pr(F) = 0.05$. The flag fires on 90 % of fouled coils, $\Pr(H \mid F) = 0.90$, its detection probability and the data sheet's nine in ten; it also fires falsely on 10 % of clean coils, $\Pr(H \mid \neg F) = 0.10$, its false-alarm probability. A flag is observed, and the question is the probability that the coil is in fact fouled, $\Pr(F \mid H)$.

The first step writes Bayes' theorem, equation (3.8), for the hypothesis F and the observation H:

$$\Pr(F \mid H) = \frac{\Pr(H \mid F)\ \Pr(F)}{\Pr(H)}.$$

The numerator is known, $\Pr(H \mid F)\ \Pr(F) = 0.90 \times 0.05 = 0.045$. The denominator, the probability of a flag, is not yet known.

3.3.2 The total probability of a flag

A flag can occur in two mutually exclusive situations: the coil is fouled and the flag fires, or the coil is clean and the flag fires falsely. In this simplified model the two exhaust the cases, so the flag event is the union of two disjoint events and its probability is their sum,

$$\Pr(H) = \Pr(H \cap F) + \Pr(H \cap \neg F).$$

The multiplication rule, equation (3.6), writes each intersection as a likelihood times a prior:

$$\begin{aligned} \Pr(H) &= \Pr(H \mid F)\ \Pr(F) + \Pr(H \mid \neg F)\ \Pr(\neg F) \\ &= 0.90 \times 0.05 + 0.10 \times 0.95 = 0.045 + 0.095 = 0.14. \end{aligned} \tag{3.9}$$

Flags occur at 14 % of tests; two thirds of them, 0.095 of the 0.14, at tests of a clean coil.

3.3.3 The posterior fouling probability

The third step divides,

$$\Pr(F \mid H) = \frac{0.045}{0.14} = 0.3214. \tag{3.10}$$

The cleaning work order of the opening scene therefore does not rest on a probability of 0.9: at 0.32 a clean coil is the likelier explanation, and whether the crew is worth sending before handover depends on what a cleaning costs against what handing over a fouled coil costs. Figure 3.3 compares the probability of fouling before and after the observation. Each row is one state of knowledge, the horizontal scale is a probability from 0 to 1, and the green part of each bar is the probability of fouling. The flag raises it from 5 % to 32.1 %; it increases suspicion substantially, but it does not establish that the coil is fouled. The third row is the case without a flag, derived in Section 3.3.4.

The base rate. The answer is not 90 %. The flag detects 90 % of fouled coils, but fouled coils are uncommon, and many flags occur while the coil is clean. Table 3.4 makes this concrete for 1000 commissioning tests. The rows are the actual condition of the coil, the columns whether the flag fired, and each entry is the expected number of tests: of the 50 tests of a fouled coil, 90 %, or 45, raise the flag; of the 950 tests of a clean coil, 10 %, or 95, raise it falsely. Of the 140 work orders the flag raises, 95 find a clean coil and only 45 a fouled one, so $\Pr(F \mid H) = 45/140 = 0.3214$. The table and Bayes' theorem give the same result, because the table is equation (3.9) multiplied by 1000.

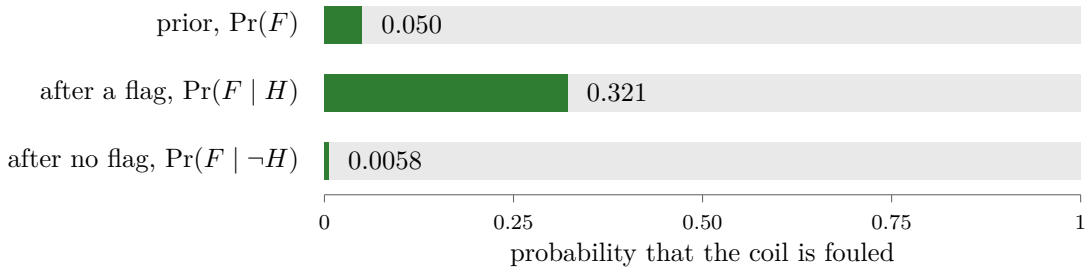


Figure 3.3: What a condition flag tells about fouling. Prior $\Pr(F) = 0.05$ per commissioning test; flag with detection probability $\Pr(H \mid F) = 0.90$ and false-alarm probability $\Pr(H \mid \neg F) = 0.10$. After a flag, $0.045/0.14 = 0.321$ (equation (3.10)); after no flag, $0.005/0.86 = 0.0058$ (Section 3.3.4).

Table 3.4: Expected outcomes of 1000 commissioning tests with $\Pr(F) = 0.05$, $\Pr(H \mid F) = 0.90$ and $\Pr(H \mid \neg F) = 0.10$, in numbers of tests.

Actual condition [–]	flag [–]	no flag [–]	total [–]
fouled	45	5	50
clean	95	855	950
total	140	860	1000

The odds form. Writing Bayes' theorem for F and for $\neg F$ with the same observation and dividing the two, the evidence $\Pr(H)$ cancels:

$$\frac{\Pr(F \mid H)}{\Pr(\neg F \mid H)} = \frac{\Pr(H \mid F)}{\Pr(H \mid \neg F)} \cdot \frac{\Pr(F)}{\Pr(\neg F)}. \tag{3.11}$$

Posterior odds equal the likelihood ratio times the prior odds. Here the likelihood ratio is $0.90/0.10 = 9$ and the prior odds are $0.05/0.95 = 0.05263$, so the posterior odds are 0.4737 and the posterior probability is $0.4737/(1 + 0.4737) = 0.3214$, as before. The flag multiplies the odds of fouling by nine whatever the base rate; the base rate decides where that multiplication starts.

3.3.4 The case without a flag

The same derivation applies to the observation that the flag does not fire. Bayes' theorem gives

$$\Pr(F \mid \neg H) = \frac{\Pr(\neg H \mid F)\ \Pr(F)}{\Pr(\neg H)},$$

with $\Pr(\neg H \mid F) = 1 - 0.90 = 0.10$, the probability that a fouled coil goes unflagged, and, by the same total probability as equation (3.9),

$$\Pr(\neg H) = 0.10 \times 0.05 + 0.90 \times 0.95 = 0.005 + 0.855 = 0.86.$$

Therefore

$$\Pr(F \mid \neg H) = \frac{0.005}{0.86} = 0.00581. \tag{3.12}$$

The absence of a flag reduces the fouling probability from 5 % to about 0.58 %, the third row of Figure 3.3; in odds, the silence multiplies the prior odds by $0.10/0.90 = 1/9$. In this example a flag substantially raises suspicion, and its absence is considerable evidence that the coil is clean.

Interpretation. Bayesian diagnosis does not only ask whether a reading crossed a threshold. It evaluates how strongly the observation supports competing explanations, weighted by their prior plausibility. The distinction is central in condition monitoring because a low health index is not uniquely associated with one physical cause: a test can read the coil low because it is fouled, because the load was above the logged one, because the supply air was warmer than logged, or because the sensor is biased. Each of these is a hypothesis with its own prior and its own likelihood for the same reading, and the inference becomes a comparison among them, the diagnosis category of Chapter 2; Chapters 6, 7 and 13 set these hypotheses against one another on the condition record of the two-unit hall.

3.4 Marginalization and dependence

The probability of a flag in Section 3.3.2 was computed without knowing whether the coil was fouled. The same operation, for discrete and continuous unknowns and for nuisance quantities, is what lets a shared filter make a coil and a fan dependent, and what an independence assumption costs.

3.4.1 Discrete marginalization over inspection grades

Suppose an inspection grades a coil of this design (illustrative) as clean, cleaning due or replace, and let S denote the grade present. Table 3.5 gives, for each grade, its probability and the probability that the flag H fires on a coil in it; the last two columns are computed below.

Table 3.5: Three inspection grades of a coil: probabilities, flag likelihoods, the joint probabilities of grade and flag, and the probabilities of the grades given a flag.

Grade s [–]	$\Pr(S = s)$ [–]	$\Pr(H \mid S = s)$ [–]	$\Pr(H \cap S = s)$ [–]	$\Pr(S = s \mid H)$ [–]
clean	0.85	0.05	0.0425	0.272
cleaning due	0.12	0.70	0.0840	0.538
replace	0.03	0.99	0.0297	0.190
total	1.00		0.1562	1.000

A flag can arise under any of the three grades. Because the grades are mutually exclusive and exhaustive, the events $H \cap \{S = s\}$ are disjoint and their union is H, so $\Pr(H) = \sum_s \Pr(H \cap S = s)$. The multiplication rule writes each term as a likelihood times a prior:

$$\Pr(H) = \sum_s \Pr(H \mid S = s)\, \Pr(S = s). \tag{3.13}$$

This is the *law of total probability*. Substituting,

$$\Pr(H) = 0.05 \times 0.85 + 0.70 \times 0.12 + 0.99 \times 0.03 = 0.0425 + 0.0840 + 0.0297 = 0.1562.$$

The probability of a flag has been computed without knowing which grade is present. That is what marginalization does: it accounts for the uncertainty in variables that are not part of the final question. The same sum is the denominator of Bayes' theorem over the three grades, and dividing each joint probability by it gives the last column of the table: given a flag, the coil is due for cleaning with probability $0.0840/0.1562 = 0.538$, clean with 0.272 and due for replacement with 0.190.

3.4.2 Continuous marginalization over a work-order log

When the unknown is continuous, the sum becomes an integral. Let the rate λ of a unit's corrective failures be uncertain and let $\mathbf{y}$ be the observed evidence; their joint distribution is $p(\lambda, \mathbf{y}) = p(\mathbf{y} \mid \lambda)\, p(\lambda)$. The probability density of the observations is obtained by integrating over every possible value of the rate,

$$p(\mathbf{y}) = \int p(\mathbf{y} \mid \lambda)\, p(\lambda)\, \mathrm{d}\lambda, \tag{3.14}$$

the *marginal likelihood* of the observations. Bayes' theorem for a continuous unknown then reads

$$p(\lambda \mid \mathbf{y}) = \frac{p(\mathbf{y} \mid \lambda)\, p(\lambda)}{\int p(\mathbf{y} \mid \lambda')\, p(\lambda')\, \mathrm{d}\lambda'}, \tag{3.15}$$

where λ' is a dummy variable of integration. The numerator evaluates the relative plausibility of one particular rate; the denominator ensures that the posterior integrates to one.

For the corrective work-order log of Chapter 2, $k = 2$ corrective failures in an exposure $T_{\mathrm{tot}} = 10$ unit-years of cooling units of this design, under the Poisson model $\Pr(k \mid \lambda) = (\lambda T_{\mathrm{tot}})^k \mathrm{e}^{-\lambda T_{\mathrm{tot}}}/k!$ and the prior $\mathrm{Gamma}(a, b) = \mathrm{Gamma}(2, 5\,\mathrm{yr})$, with density $b^a \lambda^{a-1} \mathrm{e}^{-b\lambda}/\Gamma(a)$, the standard gamma integral $\int_0^\infty \lambda^{a-1} \mathrm{e}^{-x\lambda}\, \mathrm{d}\lambda = \Gamma(a)/x^a$ with $x = b + T_{\mathrm{tot}}$ evaluates equation (3.14) in closed form:

$$\Pr(k \mid T_{\mathrm{tot}}) = \frac{T_{\mathrm{tot}}^k}{k!}\, \frac{b^a}{\Gamma(a)}\, \frac{\Gamma(a+k)}{(b + T_{\mathrm{tot}})^{a+k}} = \frac{10^2}{2} \cdot \frac{5^2}{1} \cdot \frac{3!}{15^4} = 0.1481.$$

Before the log, the model gave a probability of 0.148 to exactly two corrective failures in 10 unit-years; dividing the numerator of equation (3.15) by it gives the posterior $\mathrm{Gamma}(4, 15\,\mathrm{yr})$, with the mean rate $4/15 = 0.2667\,\mathrm{yr}^{-1}$.

The asset manager's question, whether the unit runs the next year without a corrective failure, is the same operation applied to the event $T_f > \tau$: the conditional survival of equation (3.2) averaged over the posterior of the rate,

$$R_{\mathrm{pred}}(\tau \mid \mathbf{y}) = \int_0^\infty \Pr(T_f > \tau \mid \lambda)\, p(\lambda \mid \mathbf{y})\, \mathrm{d}\lambda = \int_0^\infty \mathrm{e}^{-\lambda\tau}\, p(\lambda \mid \mathbf{y})\, \mathrm{d}\lambda. \tag{3.16}$$

Chapter 2 carries the rate in the state as the log-rate $\omega = \ln(\lambda/\lambda_{\mathrm{ref}})$, $\lambda_{\mathrm{ref}} = 0.1\,\mathrm{yr}^{-1}$, whose Newton fit gives $\hat{\omega} = \ln(8/3) = 0.98083$ with the information $\Lambda = a + k = 4$ and the posterior spread $\sigma_\omega = 1/\sqrt{\Lambda} = 0.50$. Equation (3.16) over that normal is the expectation of $\mathrm{e}^{-\lambda_{\mathrm{ref}}\, \mathrm{e}^{\omega}\tau}$ under $\omega \sim \mathcal{N}(\hat{\omega}, \sigma_\omega^2)$, the five-node Gauss–Hermite sum of equation (2.41); for $\tau = 1\,\mathrm{yr}$ it gives 0.7481, and 200 nodes print the same. This is the reported value. The gamma posterior gives the same integral in closed form. Combining the two exponentials and using the gamma integral with $x = b + \tau$, for the posterior $\mathrm{Gamma}(a, b)$,

$$R_{\mathrm{pred}}(\tau \mid \mathbf{y}) = \frac{b^a}{\Gamma(a)} \int_0^\infty \lambda^{a-1}\, \mathrm{e}^{-(b+\tau)\lambda}\, \mathrm{d}\lambda = \frac{b^a}{\Gamma(a)}\, \frac{\Gamma(a)}{(b+\tau)^a} = \left(\frac{b}{b+\tau}\right)^a, \tag{3.17}$$

which for $\mathrm{Gamma}(4, 15\,\mathrm{yr})$ and $\tau = 1\,\mathrm{yr}$ is $(15/16)^4 = 0.7725$, the exact value, from which the state route lies 0.0244 below. The probability of a corrective failure within the year is 0.2519 on the state route and 0.2275 exactly. The exact average exceeds the plug-in at the posterior mean rate, $\mathrm{e}^{-4/15} = \mathrm{e}^{-0.2667} = 0.7659$, because $\mathrm{e}^{-\lambda\tau}$ is convex in λ; the state route lies below both because the lognormal of the rate about $\hat{\lambda} = 0.2667\,\mathrm{yr}^{-1}$ has the heavier upper tail and the mean rate $0.3022\,\mathrm{yr}^{-1}$. The plug-in answers a different question, a unit whose rate is known. The predictive probability of equation (3.3) is the same operation applied to the next condition test: the conditional probability averaged over the posterior of the hall air.

3.4.3 The sensor's offset as a nuisance quantity

Suppose the log-conductance g of the coil is being estimated from its condition tests, but the model also contains the uncertain calibration offset b_T of the hall sensor, in K. The full posterior is $p(g, b_T \mid \mathbf{y})$, and the asset manager wants the distribution of g alone. It is obtained by integrating b_T out,

$$p(g \mid \mathbf{y}) = \int p(g, b_T \mid \mathbf{y}) \, \mathrm{d}b_T. \tag{3.18}$$

The offset does not have to be treated as known; its uncertainty enters the answer through the integration. This matters most in physics-constrained inference, where many unknown state entries influence the quantity studied: the marginal of one entry z_i is the posterior of the whole state $\mathbf{z}$ integrated over all its other entries.

Marginalization against optimization. These two operations are different and must not be confused. Marginalization, equation (3.18), accumulates probability across all values of b_T. Optimization over the nuisance quantity, $\max_{b_T} p(g, b_T \mid \mathbf{y})$, selects the best-fitting offset for each conductance. They generally give different results. A discrete illustration makes the difference plain. Let the coil be either fouled or clean and the offset take one of three values, with joint posterior probabilities 0.30, 0.05, 0.05 for a fouled coil and 0.20, 0.20, 0.20 for a clean one (illustrative). Marginalization gives $\Pr(\text{fouled}) = 0.40$ and $\Pr(\text{clean}) = 0.60$; maximization over b_T gives 0.30 against 0.20 and prefers the fouled coil. The fouled coil owes its single peak to one value of the offset; the clean coil is supported by every value. Maximizing over the offset over-calls fouling. The distinction returns when later chapters compare a maximum a posteriori estimate with the marginals of the posterior.

3.4.4 The coil and the fan behind one filter

Many maintenance plans assume that the components of a unit wear independently. Let F_A be the event that the year's inspection finds the coil fouled and F_B the event that it finds the fan worn. If each has probability 0.05, independence gives $\Pr(F_A \cap F_B) = 0.05^2 = 0.0025$ for the probability that both need a work order in the same year. The conclusion is invalid if the two share an important cause. The coil and the fan of a cooling unit sit behind one filter: a filter left in place too long loads the coil with dust and the fan with pressure drop. Let C be the event that a filter change is missed in a year, with $\Pr(C) = 0.05$. Figure 3.4 draws the dependency: the box at the top is the shared condition, the two boxes below are the events found at the inspection, and each arrow says that the probability of the event at its head depends on the condition at its tail. There is no arrow between the coil and the fan.

The probabilities of the two findings depend on whether the filter change was missed. Table 3.6 lists them: the rows are the two conditions and the columns the probability of each finding under that condition.

Table 3.6: Probabilities that the year's inspection finds the coil fouled and the fan worn, with and without a missed filter change (illustrative); $\Pr(C) = 0.05$.

Condition [–]	$\Pr(F_A \mid \cdot)$ [–]	$\Pr(F_B \mid \cdot)$ [–]
filter change missed, C	0.60	0.50
filter changed on time, $\neg C$	0.02	0.01

Assume that the two findings are *conditionally independent given* C:

$$\Pr(F_A \cap F_B \mid C) = \Pr(F_A \mid C)\,\Pr(F_B \mid C), \qquad \Pr(F_A \cap F_B \mid \neg C) = \Pr(F_A \mid \neg C)\,\Pr(F_B \mid \neg C).$$

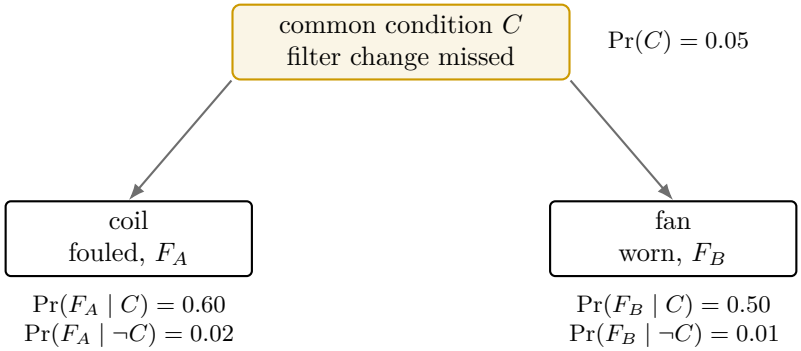


Figure 3.4: A probabilistic dependency graph for the coil and the fan of one cooling unit behind one filter. The shared condition C, a missed filter change, influences both events; the absence of an arrow between the coil and the fan states that they are conditionally independent given C. The probabilities are those of Table 3.6.

Once the history of the filter is known, the state of the coil says nothing more about the fan. This does not mean that the two are unconditionally independent, as the next subsection shows.

3.4.5 The joint probability and the independence error

The law of total probability, equation (3.13), applied to the event $F_A \cap F_B$ with the two conditions C and $\neg C$ as the states, gives

$$\Pr(F_A \cap F_B) = \Pr(F_A \cap F_B \mid C)\ \Pr(C) + \Pr(F_A \cap F_B \mid \neg C)\ \Pr(\neg C).$$

Under conditional independence each conditional joint probability is a product,

$$\Pr(F_A \cap F_B) = 0.60 \times 0.50 \times 0.05 + 0.02 \times 0.01 \times 0.95 = 0.015 + 0.00019,$$

so that

$$\Pr(F_A \cap F_B) = 0.01519. \tag{3.19}$$

The probability that the year's inspection finds both the coil fouled and the fan worn is about 1.52 %, and almost all of it, 0.015 of 0.01519, comes from the missed filter change.

The marginal probabilities follow from the same law,

$$\Pr(F_A) = 0.60 \times 0.05 + 0.02 \times 0.95 = 0.049, \qquad \Pr(F_B) = 0.50 \times 0.05 + 0.01 \times 0.95 = 0.0345.$$

If the two are incorrectly treated as independent, the joint probability is the product of the marginals,

$$\Pr(F_A \cap F_B) \overset{\text{incorrect}}{=} 0.049 \times 0.0345 = 0.00169,$$

or 0.169 %. The correct value, 1.519 %, is 8.99 times larger: ignoring the shared filter underestimates the probability that the coil and the fan need work in the same year by a factor of about nine in this illustrative unit, and the independent product is only 11 % of the correct value.

Common mistake

Multiplying the marginal probabilities of components that share a maintenance cause to obtain the probability that both need work together. The product is the joint probability only when the two are independent; a shared filter, a shared environment or a shared maintenance crew makes them dependent, and the year in which both need a work order is much more likely than the product suggests.

Principle 3.2. *Understanding what the components of an asset share can matter more than estimating the wear probability of each one with greater precision.*

3.4.6 Conditional independence given the filter

Two random variables X_1 and X_2 are conditionally independent given C when, for every value c of C,

$$p(x_1, x_2 \mid c) = p(x_1 \mid c)\, p(x_2 \mid c), \tag{3.20}$$

written $X_1 \perp X_2 \mid C$. Dividing both sides by $p(x_1 \mid c)$ and using the definition of conditional probability on the left gives the equivalent statement $p(x_2 \mid x_1, c) = p(x_2 \mid c)$: after conditioning on C, learning X_1 gives no additional information about X_2 under the assumed model.

When C is unknown, observing X_1 informs C, and that changes what is believed about X_2. If the coil is found fouled, the filter becomes suspect: by Bayes' theorem,

$$\Pr(C \mid F_A) = \frac{\Pr(F_A \mid C)\ \Pr(C)}{\Pr(F_A)} = \frac{0.60 \times 0.05}{0.049} = 0.612,$$

up from 0.05. The probability that the fan is worn too is then the law of total probability conditioned on F_A, with conditional independence used to drop F_A from $\Pr(F_B \mid C, F_A)$:

$$\begin{aligned}\Pr(F_B \mid F_A) &= \Pr(F_B \mid C)\ \Pr(C \mid F_A) + \Pr(F_B \mid \neg C)\ \Pr(\neg C \mid F_A) \\ &= 0.50 \times 0.612 + 0.01 \times 0.388 = 0.310,\end{aligned}$$

which equals $\Pr(F_A \cap F_B)/\Pr(F_A) = 0.01519/0.049$, against the unconditional $\Pr(F_B) = 0.0345$. A fouled coil raises the probability of a worn fan ninefold through the filter they share, so a fouled coil sends the crew to the fan. Conditional independence is therefore not unconditional independence: the variables are independent only once the common cause is known.

3.5 Conditional independence and a continuous update in a physical model

The same rules apply when the unknowns are physical quantities tied by a conservation relation. On the equipment hall they give the posterior of a state vector under Gaussian rows, the probability that the hall exceeds its limit after one test, and, in the two-entry hall of Chapter 1, a test reading that reaches the conductance through the physics row.

3.5.1 The hall as a probability model

The hall of the earlier chapters satisfies the steady heat balance

$$T = T_c + \frac{Q}{G}, \tag{3.21}$$

with the hall air T and the supply air T_c in °C, the heat load Q in kW and the conductance G of the air path in kW/K. A sensor reads the hall air with an additive error,

$$y_T = T + \epsilon, \qquad \epsilon \sim \mathcal{N}(0, \sigma_T^2),$$

where ϵ is the measurement error and σ_T its standard deviation in K. Figure 3.5 draws the system twice. Panel (a) is the plant graph: the source supplies the load Q to the hall node, which carries the free entry T; the green air path to the supply-air node carries the conductance, held in the state as the log-conductance $g = \ln(G/G_{\text{ref}})$, whose exponential is the health index; the supply-air node is pinned at T_c by the unit's discharge controller; and the dashed arrow is

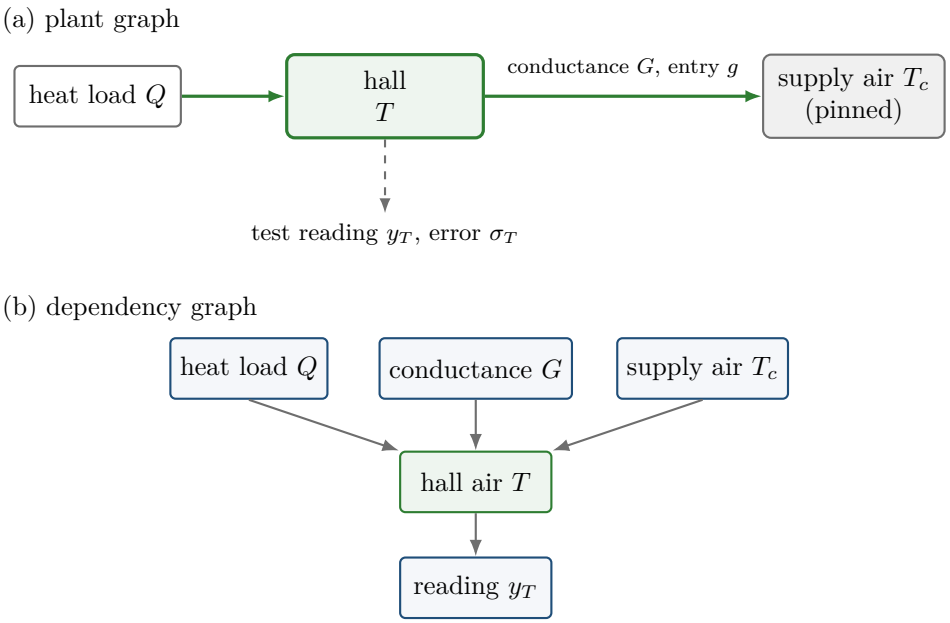


Figure 3.5: The hall in two drawings. (a) The plant graph: free entry T at the hall node and $g = \ln(G/G_{\text{ref}})$ on the air path; supply air T_c pinned by the discharge controller; source Q; the test reading y_T observed (dashed). (b) The dependency graph of the factorized probability model of this section, in which Q, G and T_c carry priors: arrows point from a quantity to what depends on it, and the reading depends on the inputs only through T.

the test reading y_T. Panel (b) is the dependency graph of the probability model of this section, in which Q, G and T_c are all treated as uncertain: the load, the conductance and the supply air determine the modeled hall air, and the reading depends on that temperature alone.

If the measurement error is independent of the physical inputs, the reading is conditionally independent of those inputs once the true temperature is known:

$$p(y_T \mid T, G, Q, T_c) = p(y_T \mid T). \tag{3.22}$$

This does not make the reading uninformative about the conductance. The reading informs T, and the physical relation (3.21) connects T to G: evidence propagates from the sensor reading to the temperature and from the temperature to the conductance. An observation of one quantity updates the beliefs about every quantity connected to it through shared physical constraints, and this is why a test of the hall air reads the condition of the coil.

The factorized probability model. Suppose G, Q and T_c are uncertain inputs with a joint prior $p(G, Q, T_c)$, which relaxes the pins of the plant graph and carries Q and T_c as free entries. The chain rule, the multiplication rule applied twice, and equation (3.22) give the joint model

$$p(G, Q, T_c, T, y_T) = p(G, Q, T_c)\, p(T \mid G, Q, T_c)\, p(y_T \mid T). \tag{3.23}$$

The physics factor is written as a Gaussian around the heat balance, $p(T \mid G, Q, T_c) = \mathcal{N}(T;\ T_c + Q/G,\ \sigma_p^2)$, with a model tolerance σ_p in K: a physical discrepancy is a finite σ_p, and the exact deterministic balance of equation (3.21) is the limit $\sigma_p \to 0$, in which the factor becomes the Dirac delta $\delta(T - T_c - Q/G)$. The joint distribution is the complete probability model, and the posterior wanted is $p(G, Q, T_c, T \mid y_T)$, over the whole state with T kept in it, which answers questions about the conductance, the temperature, the margins and every other quantity of the model. Each of its three factors is a row over that state: the prior, the physics row with tolerance σ_p and the reading row; the derivations of this section keep all of them as rows and

never remove an entry by substitution. In the limit $\sigma_p \to 0$, integrating T out against the delta gives the same posterior of the inputs as $p(G, Q, T_c \mid y_T) \propto p(G, Q, T_c)\, p(y_T \mid T = T_c + Q/G)$: every combination of inputs is weighted by how well the temperature it implies explains the reading.

3.5.2 Prior and reading for the hall air

The simplest continuous case has one unknown, the true hall air temperature T, so that the state is this one number,

$$T \in \mathbb{R}.$$

The general forms of the book write the state as a vector $\mathbf{z}$; here it has the single entry T, so every quantity of the fit below is a number: each row's residual and its slope in T, the objective's slope and curvature, the information and the posterior variance. On an illustrative hall running warmer than the series' hall, a thermal model at the planned load gives an uncertain prior estimate,

$$T \sim \mathcal{N}(\mu_0, \sigma_0^2) = \mathcal{N}(26, 2^2), \qquad \mu_0 = 26^\circ\mathrm{C}, \quad \sigma_0 = 2\,\mathrm{K}.$$

A condition test reports $y_T = 27^\circ$C, and the sensor's error is Gaussian with $\sigma_T = 0.5\,\mathrm{K}$, so that $y_T \mid T \sim \mathcal{N}(T, 0.5^2)$. The question is the posterior distribution of the true temperature after the test. Bayes' theorem for a continuous unknown, with the evidence dropped because it does not depend on the state, gives

$$p(T \mid y_T) \propto p(y_T \mid T)\, p(T),$$

with the Gaussian prior and the Gaussian likelihood

$$p(T) \propto \exp\left[-\frac{(T-\mu_0)^2}{2\sigma_0^2}\right], \qquad p(y_T \mid T) \propto \exp\left[-\frac{(y_T - T)^2}{2\sigma_T^2}\right].$$

Multiplying,

$$p(T \mid y_T) \propto \exp\left[-\frac{1}{2}\left(\frac{(T-\mu_0)^2}{\sigma_0^2} + \frac{(y_T - T)^2}{\sigma_T^2}\right)\right]. \tag{3.24}$$

Each term of the exponent is the square of a difference divided by its standard deviation. The two differences are the residuals of the problem, one prior row and one reading row, each already divided by its standard deviation,

$$r_1(T) = \frac{T - \mu_0}{\sigma_0}, \qquad r_2(T) = \frac{T - y_T}{\sigma_T}, \tag{3.25}$$

and equation (3.24) reads $p(T \mid y_T) \propto \exp[-\frac{1}{2}(r_1^2 + r_2^2)] = \mathrm{e}^{-\mathcal{L}(T)}$, with the objective $\mathcal{L}(T) = \frac{1}{2}(r_1^2 + r_2^2)$ of the notation chapter. Both rows are linear in T, each with a constant slope $j_\ell = \mathrm{d}r_\ell/\mathrm{d}T$: $j_1 = 1/\sigma_0$ and $j_2 = 1/\sigma_T$.

Gaussian rows over any state. The same step holds for a state of any size and any number of rows. The joint model factorizes by the chain rule, as equation (3.23) does for the hall: a prior factor over the inputs, a physics factor for each physical relation, and a reading factor for each reading, which by conditional independence, equation (3.22), depends on the state only through the quantity read. Row ℓ is one such factor written as a residual: the difference between the quantity computed from the state and what was observed, or what the prior or the physics expects, divided by its standard deviation σ_ℓ. A Gaussian factor is $\exp[-\frac{1}{2}r_\ell(\mathbf{z})^2]$ times a normalizer, $1/(\sqrt{2\pi}\,\sigma_\ell)$, that is free of the state. Multiplying the factors and dropping the normalizers and the evidence, Bayes' theorem gives

$$p(\mathbf{z} \mid \mathbf{y}) \propto \prod_\ell \exp\left[-\tfrac{1}{2}r_\ell(\mathbf{z})^2\right] = \exp\left[-\tfrac{1}{2}r(\mathbf{z})^{\mathsf{T}} r(\mathbf{z})\right] = \mathrm{e}^{-\mathcal{L}(\mathbf{z})}, \tag{3.26}$$

where $r(\mathbf{z})$ stacks the reading rows, the prior rows and the physics rows with their tolerances σ_p. For the one-entry hall the product has the two factors of equation (3.24) and the sum of squares is the number $r_1^2 + r_2^2$; the vector is needed when the state has several entries, as in Section 3.5.6 and in the hall of Chapter 4. A Gaussian prior over several entries, $\mathcal{N}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$ with correlated entries, enters as the rows $\mathbf{L}_0(\mathbf{z} - \boldsymbol{\mu}_0)$ with $\mathbf{L}_0^\mathsf{T}\mathbf{L}_0 = \boldsymbol{\Sigma}_0^{-1}$, since then $(\mathbf{z} - \boldsymbol{\mu}_0)^\mathsf{T}\boldsymbol{\Sigma}_0^{-1}(\mathbf{z} - \boldsymbol{\mu}_0)$ is the sum of their squares.

3.5.3 Completion of the square in the state

Suppose the rows are linear in the state, as both rows of equation (3.25) are. Then, for any point $\mathbf{z}_\mathrm{p}$,

$$r(\mathbf{z}) = r(\mathbf{z}_\mathrm{p}) + \mathbf{J}\,(\mathbf{z} - \mathbf{z}_\mathrm{p})$$

holds exactly, with the Jacobian $\mathbf{J} = \partial r/\partial \mathbf{z}$, a matrix of n_r rows and n_z columns for n_r rows and n_z entries. Write $\boldsymbol{\Delta} = \mathbf{z} - \mathbf{z}_\mathrm{p}$ and $r_0 = r(\mathbf{z}_\mathrm{p})$. The objective expands as

$$\mathcal{L}(\mathbf{z}) = \tfrac{1}{2}(r_0 + \mathbf{J}\boldsymbol{\Delta})^\mathsf{T}(r_0 + \mathbf{J}\boldsymbol{\Delta}) = \tfrac{1}{2}r_0^\mathsf{T}r_0 + \boldsymbol{\Delta}^\mathsf{T}\mathbf{J}^\mathsf{T}r_0 + \tfrac{1}{2}\boldsymbol{\Delta}^\mathsf{T}\mathbf{J}^\mathsf{T}\mathbf{J}\,\boldsymbol{\Delta}.$$

Its gradient, $\mathbf{J}^\mathsf{T}r_0 + \mathbf{J}^\mathsf{T}\mathbf{J}\,\boldsymbol{\Delta}$, vanishes at the step $\hat{\boldsymbol{\Delta}}$ that solves the *normal equations*,

$$(\mathbf{J}^\mathsf{T}\mathbf{J})\,\hat{\boldsymbol{\Delta}} = -\mathbf{J}^\mathsf{T}r(\mathbf{z}_\mathrm{p}), \qquad \hat{\mathbf{z}} = \mathbf{z}_\mathrm{p} + \hat{\boldsymbol{\Delta}}. \tag{3.27}$$

Writing $\mathbf{J}^\mathsf{T}r_0 = -\mathbf{J}^\mathsf{T}\mathbf{J}\,\hat{\boldsymbol{\Delta}}$ in the expansion completes the square,

$$\mathcal{L}(\mathbf{z}) = \tfrac{1}{2}(\boldsymbol{\Delta} - \hat{\boldsymbol{\Delta}})^\mathsf{T}\mathbf{J}^\mathsf{T}\mathbf{J}\,(\boldsymbol{\Delta} - \hat{\boldsymbol{\Delta}}) + \text{constant} = \tfrac{1}{2}(\mathbf{z} - \hat{\mathbf{z}})^\mathsf{T}\mathbf{J}^\mathsf{T}\mathbf{J}\,(\mathbf{z} - \hat{\mathbf{z}}) + \text{constant},$$

where the constant collects the terms free of $\mathbf{z}$. By equation (3.26) the posterior is the exponential of minus this quadratic form, the density of a multivariate Gaussian with mean $\hat{\mathbf{z}}$ and covariance

$$\boldsymbol{\Sigma} = (\mathbf{J}^\mathsf{T}\mathbf{J})^{-1}. \tag{3.28}$$

The matrix $\mathbf{J}^\mathsf{T}\mathbf{J}$ is the information matrix $\boldsymbol{\Lambda}$ of the normalized rows. When the rows are written unnormalized, each in its own unit with the weight $w_\ell = 1/\sigma_\ell^2$ and the weights collected in $\mathbf{W} = \mathrm{diag}(w_\ell)$, the same matrix is $\boldsymbol{\Lambda} = \mathbf{J}^\mathsf{T}\mathbf{W}\mathbf{J}$, with $\mathbf{J}$ there the Jacobian of the unnormalized rows. For rows linear in the state the normal equations reach $\hat{\mathbf{z}}$ in one step from any $\mathbf{z}_\mathrm{p}$, and equations (3.27) and (3.28) are the exact posterior mean and covariance. For nonlinear rows the same step is the Gauss–Newton step, repeated from the new point until it vanishes, and the result is a local Gaussian approximation, the Laplace approximation, whose derivation and curvature term Chapter 4 gives, with the three-entry hall as its example; Section 3.5.6 reads the two-entry hall of Chapter 1 in the terms of this section.

Precisions add. Split the rows into prior rows $\mathbf{L}_0(\mathbf{z} - \boldsymbol{\mu}_0)$, whose Jacobian $\mathbf{L}_0$ satisfies $\mathbf{L}_0^\mathsf{T}\mathbf{L}_0 = \boldsymbol{\Sigma}_0^{-1}$, and reading rows $\mathbf{V}_y^{-1/2}(\mathbf{A}\mathbf{z} - \mathbf{y})$, where the matrix $\mathbf{A}$ maps the state to the quantities read and the diagonal matrix $\mathbf{V}_y$ holds the variances of the readings. The Jacobian of the reading rows is $\mathbf{J}_y = \mathbf{V}_y^{-1/2}\mathbf{A}$; stacking the two blocks of rows stacks their Jacobians, and $\mathbf{J}^\mathsf{T}\mathbf{J}$ is the sum of the two blocks' products,

$$\boldsymbol{\Sigma}^{-1} = \boldsymbol{\Sigma}_0^{-1} + \mathbf{J}_y^\mathsf{T}\mathbf{J}_y = \boldsymbol{\Sigma}_0^{-1} + \mathbf{A}^\mathsf{T}\mathbf{V}_y^{-1}\mathbf{A}. \tag{3.29}$$

Starting the normal equations at $\mathbf{z}_\mathrm{p} = \boldsymbol{\mu}_0$, where the prior rows vanish, gives $\mathbf{J}^\mathsf{T}r(\boldsymbol{\mu}_0) = \mathbf{A}^\mathsf{T}\mathbf{V}_y^{-1}(\mathbf{A}\boldsymbol{\mu}_0 - \mathbf{y})$, and equation (3.27) gives

$$\hat{\mathbf{z}} = \boldsymbol{\mu}_0 + \boldsymbol{\Sigma}\,\mathbf{A}^\mathsf{T}\mathbf{V}_y^{-1}(\mathbf{y} - \mathbf{A}\boldsymbol{\mu}_0). \tag{3.30}$$

The posterior information is the prior information plus the information of the readings, and the posterior mean is moved from the prior mean toward the readings in proportion to their information.

The conditional of a joint Gaussian. The same posterior follows from the joint distribution of the state and the readings. Before the readings, $\mathbf{z} \sim \mathcal{N}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$ and $\mathbf{y} = \mathbf{A}\mathbf{z} + \boldsymbol{\epsilon}$ with $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{V}_y)$ independent of the state, so the stacked vector of $\mathbf{z}$ and $\mathbf{y}$ is Gaussian with means $\boldsymbol{\mu}_0$ and $\mathbf{A}\boldsymbol{\mu}_0$ and covariance blocks

$$\boldsymbol{\Sigma}_{zz} = \boldsymbol{\Sigma}_0, \qquad \boldsymbol{\Sigma}_{zy} = \boldsymbol{\Sigma}_0\mathbf{A}^\mathsf{T} = \boldsymbol{\Sigma}_{yz}^\mathsf{T}, \qquad \boldsymbol{\Sigma}_{yy} = \mathbf{A}\boldsymbol{\Sigma}_0\mathbf{A}^\mathsf{T} + \mathbf{V}_y.$$

Let $\mathbf{P}$ be the inverse of the joint covariance and $\mathbf{S} = \boldsymbol{\Sigma}_{zz} - \boldsymbol{\Sigma}_{zy}\boldsymbol{\Sigma}_{yy}^{-1}\boldsymbol{\Sigma}_{yz}$ the *Schur complement* of $\boldsymbol{\Sigma}_{yy}$. The block-inversion formula, checked by multiplying back, gives $\mathbf{P}_{zz} = \mathbf{S}^{-1}$ and $\mathbf{P}_{zy} = -\mathbf{S}^{-1}\boldsymbol{\Sigma}_{zy}\boldsymbol{\Sigma}_{yy}^{-1}$. With the readings fixed, the exponent of the joint density, times -1, is a quadratic in the state,

$$\tfrac{1}{2}(\mathbf{z}-\boldsymbol{\mu}_0)^\mathsf{T}\mathbf{P}_{zz}(\mathbf{z}-\boldsymbol{\mu}_0) + (\mathbf{z}-\boldsymbol{\mu}_0)^\mathsf{T}\mathbf{P}_{zy}(\mathbf{y}-\mathbf{A}\boldsymbol{\mu}_0) + \text{constant} = \tfrac{1}{2}(\mathbf{z}-\mathbf{m})^\mathsf{T}\mathbf{S}^{-1}(\mathbf{z}-\mathbf{m}) + \text{constant},$$

with $\mathbf{m} = \boldsymbol{\mu}_0 + \boldsymbol{\Sigma}_{zy}\boldsymbol{\Sigma}_{yy}^{-1}(\mathbf{y} - \mathbf{A}\boldsymbol{\mu}_0)$. The conditional density is the joint density divided by $p(\mathbf{y})$, which is free of the state, so

$$\mathbf{z} \mid \mathbf{y} \sim \mathcal{N}(\boldsymbol{\mu}_0 + \boldsymbol{\Upsilon}(\mathbf{y} - \mathbf{A}\boldsymbol{\mu}_0),\ \boldsymbol{\Sigma}_0 - \boldsymbol{\Upsilon}\mathbf{A}\boldsymbol{\Sigma}_0), \qquad \boldsymbol{\Upsilon} = \boldsymbol{\Sigma}_0\mathbf{A}^\mathsf{T}(\mathbf{A}\boldsymbol{\Sigma}_0\mathbf{A}^\mathsf{T} + \mathbf{V}_y)^{-1}, \tag{3.31}$$

since $\boldsymbol{\Sigma}_{zy}\boldsymbol{\Sigma}_{yy}^{-1}\boldsymbol{\Sigma}_{yz} = \boldsymbol{\Upsilon}\mathbf{A}\boldsymbol{\Sigma}_0$. This is the correction form: the prior mean is moved toward the readings by the gain $\boldsymbol{\Upsilon}$ times the innovation $\mathbf{y} - \mathbf{A}\boldsymbol{\mu}_0$, and the covariance is the Schur complement. The Kalman literature writes the gain $\mathbf{K}$; this book keeps $\mathbf{K}$ for the hat matrix and κ for an assurance level. It is the same posterior as equations (3.29) and (3.30), by the matrix inversion identity $(\boldsymbol{\Sigma}_0^{-1} + \mathbf{A}^\mathsf{T}\mathbf{V}_y^{-1}\mathbf{A})^{-1} = \boldsymbol{\Sigma}_0 - \boldsymbol{\Upsilon}\mathbf{A}\boldsymbol{\Sigma}_0$; the information form inverts a matrix of the size of the state, the correction form one of the size of the readings.

Remark 3.1. For one entry and one reading, $\mathbf{A} = 1$, $\boldsymbol{\Sigma}_0 = \sigma_0^2$ and $\mathbf{V}_y = \sigma_T^2$, and the three equations reduce to the scalar rules $1/\sigma_{\text{post}}^2 = 1/\sigma_0^2 + 1/\sigma_T^2$, $\mu_{\text{post}} = \sigma_{\text{post}}^2(\mu_0/\sigma_0^2 + y_T/\sigma_T^2)$ and $\mu_{\text{post}} = \mu_0 + \upsilon(y_T - \mu_0)$ with $\upsilon = \sigma_0^2/(\sigma_0^2 + \sigma_T^2)$: precisions add, and the posterior mean is the precision-weighted average of the prior mean and the reading.

3.5.4 The numbers and what changed

For the hall the state is the one number T, and the two rows of equation (3.25) have the slopes

$$j_1 = \frac{1}{\sigma_0} = 0.5\,\mathrm{K}^{-1}, \qquad j_2 = \frac{1}{\sigma_T} = 2\,\mathrm{K}^{-1}.$$

The row terms are $\mathcal{L}_\ell = \frac{1}{2}r_\ell^2$, with the slopes $\mathcal{L}'_\ell = j_\ell r_\ell$ and the curvatures $\mathcal{L}''_\ell = j_\ell^2$ in T, the primes denoting $\mathrm{d}/\mathrm{d}T$; slopes and curvatures add over the rows,

$$\mathcal{L}' = j_1 r_1 + j_2 r_2, \qquad \mathcal{L}'' = j_1^2 + j_2^2.$$

Because the rows are linear, the curvature has no further term and is the same at every T; it is the information

$$\Lambda = j_1^2 + j_2^2 = 0.5^2 + 2^2 = 0.25 + 4 = 4.25\,\mathrm{K}^{-2},$$

the prior precision $1/\sigma_0^2$ plus the reading precision $1/\sigma_T^2$, as equation (3.29) states in the scalar form of the remark above, with the prior variance $\sigma_0^2 = 4\,\mathrm{K}^2$ and the reading variance $\sigma_T^2 = 0.25\,\mathrm{K}^2$. It is the one-entry case of $\mathbf{J}^\mathsf{T}\mathbf{J}$, and equation (3.28) gives the posterior variance as its reciprocal,

$$\sigma_{\text{post}}^2 = \frac{1}{\Lambda} = \frac{1}{4.25} = 0.2353\,\mathrm{K}^2, \qquad \sigma_{\text{post}} = 0.4851\,\mathrm{K}.$$

With one entry the normal equations (3.27) are the one equation $\Lambda\,\Delta T = -\mathcal{L}'$, the Newton step $\Delta T = -\mathcal{L}'/\mathcal{L}''$. They start at the prior mean, $T_0 = 26°\mathrm{C}$, where the prior row is 0 and

the reading row $(26-27)/0.5=-2$, so that $\mathcal{L}'=0.5\times 0+2\times(-2)=-4\,\mathrm{K}^{-1}$, and the step is $\Delta T=4/4.25=0.9412\,\mathrm{K}$. The iterations are

l [–]	T [°C]	$\mathcal{L}(T)$ [–]	$\mathcal{L}'_1$ [K^{-1}]	$\mathcal{L}'_2$ [K^{-1}]	$\mathcal{L}'$ [K^{-1}]	$\mathcal{L}''$ [K^{-2}]	ΔT [K]
0	26	2.0000	0	−4	−4	4.25	0.9412
1	26.9412	0.1176	0.2353	−0.2353	0	4.25	0

At the second point the residuals are $r_1=0.4706$ and $r_2=-0.1176$, the slope is $\mathcal{L}'=0.5\times 0.4706+2\times(-0.1176)=0$, and the step vanishes: the rows are linear, and one step reaches the minimum. The correction form (3.31) confirms both moments: the variance of the reading before it arrives is $\sigma_0^2+\sigma_T^2=4+0.25=4.25\,\mathrm{K}^2$, $\upsilon=4/4.25=0.9412$, the mean $26+0.9412\times(27-26)=26.9412$°C and the Schur complement $4-0.9412\times 4=0.2353\,\mathrm{K}^2$. The answer is the minimum $\hat{T}=26.9412$°C, written μ_{post} below, with the standard deviation σ_{post}:

$$T\mid y_T\sim\mathcal{N}(26.9412,0.4851^2). \tag{3.32}$$

Figure 3.6(a) draws the update. The horizontal axis is the temperature and the vertical axis the probability density. The gray curve is the prior, the dashed blue curve the likelihood of the reading as a function of T (normalized to integrate to one), and the green curve the posterior. The posterior mean moved from 26 to 26.94°C, and the standard deviation fell from 2 to 0.49 K. The reading was more precise than the prior, so it received the greater weight, $\upsilon=0.94$; but the posterior mean is not exactly 27°C, because the prior also carries information.

Table 3.7 shows the trend for sensors of four precisions, all reading 27°C against the same prior. The first column is the sensor's standard deviation, the second the weight υ of the reading, and the remaining columns the posterior mean, the posterior standard deviation and the exceedance probability derived in Section 3.5.5. Each row is the same computation with the slopes $j_1=1/\sigma_0$ and $j_2=1/\sigma_T$, so that $\sigma_{\text{post}}^2=1/\Lambda=1/(0.25+1/\sigma_T^2)$. The second row is the case worked above. A sensor as uncertain as the prior, $\sigma_T=2\,\mathrm{K}$, receives half the weight and leaves the posterior standard deviation at $2/\sqrt{2}=1.41\,\mathrm{K}$.

Table 3.7: The hall-air update for four sensor precisions: prior $\mathcal{N}(26,2^2)$, reading $y_T=27$°C, equations (3.29) and (3.31), and the exceedance probability of $T_{\max}=28$°C derived in Section 3.5.5.

σ_T [K]	weight υ [–]	μ_{post} [°C]	σ_{post} [K]	$\Pr(T>T_{\max}\mid y_T)$ [–]
0.25	0.9846	26.9846	0.2481	2×10^{-5}
0.5	0.9412	26.9412	0.4851	0.0145
1.0	0.8000	26.8000	0.8944	0.0899
2.0	0.5000	26.5000	1.4142	0.1444

3.5.5 The requirement question

Suppose the hall-air limit of this illustration is $T_{\max}=28$°C, and the question is whether the hall meets it after the test, through the probability that it exceeds it, $\Pr(T>T_{\max}\mid y_T)$. Since the posterior is Gaussian, standardizing with the mean $\mu_{\text{post}}=\hat{T}$ and the standard deviation $\sigma_{\text{post}}=1/\sqrt{\Lambda}$ and using the standard normal cumulative distribution function Φ,

$$\begin{aligned}\Pr(T>T_{\max}\mid y_T)&=1-\Phi\left(\frac{T_{\max}-\mu_{\text{post}}}{\sigma_{\text{post}}}\right)\\&=1-\Phi\left(\frac{28-26.9412}{0.4851}\right)=1-\Phi(2.183)=0.0145,\end{aligned} \tag{3.33}$$

or about 1.45 %. Before the test the same calculation with the prior gives $\Pr(T > 28°\mathrm{C}) = 1 - \Phi((28-26)/2) = 1 - \Phi(1) = 0.1587$. The estimated exceedance probability fell from 15.9 % to 1.45 %, for the same fixed operating condition and the assumed Gaussian measurement model. The rows are linear, so this normal tail is the exact posterior's: a direct quadrature of the product (3.24) also gives 0.0145, and there is no gap to set beside it, as there is on the curved physics row of the next subsection. Figure 3.6(b) shades the two tails: the axes are those of panel (a), the dashed line is the limit, the gray area to its right under the prior is 0.159, and the green area under the posterior is 0.0145.

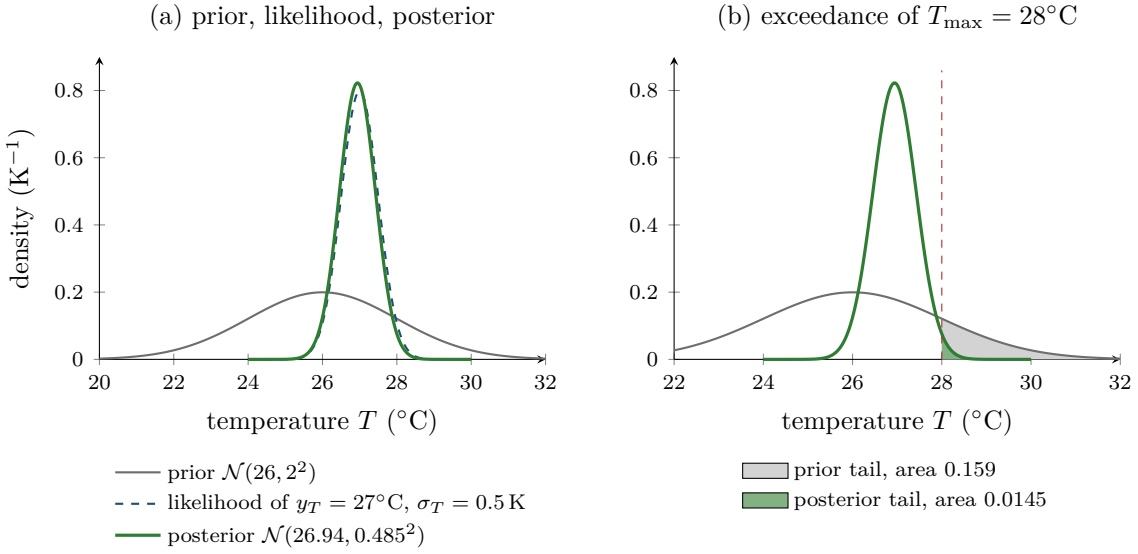


Figure 3.6: The hall-air update of equations (3.29)–(3.32). (a) Prior $\mathcal{N}(26, 2^2)$, likelihood of the test reading $y_T = 27°\mathrm{C}$ with $\sigma_T = 0.5\,\mathrm{K}$ drawn as the density $\mathcal{N}(27, 0.5^2)$ in T, and posterior $\mathcal{N}(26.9412, 0.4851^2)$. (b) The probability that the hall exceeds $T_{\max} = 28°\mathrm{C}$ (dashed): $1 - \Phi(1) = 0.159$ under the prior and $1 - \Phi(2.183) = 0.0145$ under the posterior (equation (3.33)).

The hall did not necessarily change when the test was taken; the information about it changed, and that is the mechanism of Bayesian updating. The calculation also walks the method of *How to read this book* on its smallest example. The plant graph of Figure 3.5(a) is step 0. The question $\Pr(T > T_{\max} \mid y_T)$, unknown on the left and reading on the right, is step 1. Attaching the prior and the reading as the two rows of equation (3.25) is step 2. The posterior, derived through the normal equations and the variance as the reciprocal of the information, equations (3.27) and (3.28) with one entry, and computed by the one Newton step $-\mathcal{L}'/\mathcal{L}''$ with $\Lambda = 4.25\,\mathrm{K}^{-2}$, is step 3. Reading the normal tail from it is step 4, and checking that tail by the correction form and its Schur complement is step 5. Step 6 is the report sentence: "after the condition test the hall exceeds 28°C with probability 1.45 %, posterior $26.94 \pm 0.49\,\mathrm{K}$". Step 7, the next measurement, reads Table 3.7: if 1.45 % were too high a residual doubt, a sensor with $\sigma_T = 0.25\,\mathrm{K}$ at the next condition test would, for a reading near the same value, reduce it to the order of 10^{-5}. The posterior is inferred first, and the asset manager's question is then answered from it:

$$\underbrace{p(T \mid y_T)}_{\text{inference}} \quad \longrightarrow \quad \underbrace{\Pr(T > T_{\max} \mid y_T)}_{\text{requirement question}}.$$

3.5.6 The two-entry hall and its physics row

The reading of the one-entry hall informed only the entry it measured; Section 3.5.1 claimed that a reading also reaches the conductance through the physics row. The commissioning fit of

Chapter 1 shows it. Its state is $\mathbf{z} = (T,\ g)^{\mathsf{T}}$, and its three rows, equation (1.11), are the physics row with tolerance $\sigma_p = 0.01\,\mathrm{K}$, the reading row of the commissioning test with $\sigma_T = 0.5\,\mathrm{K}$ and the handbook prior row on g with standard deviation 0.3; Chapter 1 solves their normal equations by Gauss–Newton from $(23°\mathrm{C},\ 0)$ (Table 1.4) and reaches $\hat{\mathbf{z}} = (23.070°\mathrm{C},\ -0.0139)^{\mathsf{T}}$, whose second entry is the commissioned value of the health index. Equation (3.29) holds row by row: with $\mathbf{J}_\ell$ the ℓ-th row of $\mathbf{J}$, $\mathbf{J}^{\mathsf{T}}\mathbf{J} = \sum_\ell \mathbf{J}_\ell^{\mathsf{T}}\mathbf{J}_\ell$, each row contributing the outer product of its own sensitivities. At $\hat{\mathbf{z}}$ the rows of $\mathbf{J}$ are $(100,\ 507.0)$, $(2,\ 0)$ and $(0,\ 3.333)$, and

$$\mathbf{J}^{\mathsf{T}}\mathbf{J} = \underbrace{\begin{bmatrix} 10\,000 & 50\,701 \\ 50\,701 & 257\,056 \end{bmatrix}}_{\text{physics}} + \underbrace{\begin{bmatrix} 4 & 0 \\ 0 & 0 \end{bmatrix}}_{\text{reading}} + \underbrace{\begin{bmatrix} 0 & 0 \\ 0 & 11.11 \end{bmatrix}}_{\text{prior}} = \begin{bmatrix} 10\,004 & 50\,701 \\ 50\,701 & 257\,067 \end{bmatrix}.$$

The reading's and the prior's terms are diagonal, because each of those rows carries one entry. The only off-diagonal entry, 50 701, comes from the physics row, the one row with sensitivities in both columns: this is equation (3.22) in vector form, the reading informing T and the physics row carrying that information to g. Without the reading row, the physics and prior terms alone invert to $\sqrt{\Sigma_{22}} = 0.3$, the prior spread; with it, the covariance of Chapter 1, equation (1.22), has $\sqrt{\Sigma_{11}} = 0.481\,\mathrm{K}$, $\sqrt{\Sigma_{22}} = 0.0948$ and correlation -0.9998. The test has cut the spread of the health index from 0.3 to 0.0948 although its own row has a zero in the column of g, and the posterior lies along the physics curve, on the short arc of the state plane of Figure 1.5 inside the reading band. Because the physics row is nonlinear, these spreads include the curvature term that Chapter 4 derives; $(\mathbf{J}^{\mathsf{T}}\mathbf{J})^{-1}$ alone gives $0.475\,\mathrm{K}$ and 0.094, and the exact posterior sits beside them as Chapter 1 shows.

3.5.7 Dependence inside and outside the physics graph

The common cause of Section 3.4.4 cost a factor of nine because it was left out of the product. A plant graph built on conservation relations does not leave out a shared cause that has a physical path: the cause is an entry of the state, and every row that depends on it carries it. The two-entry hall above shows this. The test reading and the health index share no row, yet the physics row couples them, and the posterior correlation of T and g, -0.9998, comes out of $\mathbf{J}^{\mathsf{T}}\mathbf{J}$ without being assumed. Had the filter been in the graph, with the unit's airflow as an entry and the coil's conductance and the fan's load each a row of it, the joint probability (3.19) would have followed from the rows, and the independent product of Section 3.4.5 could not have been formed. Conditional independence given the state is then the statement of Section 3.4.6 with the common cause known: it holds because the cause is in the state.

What the graph does not supply is the independence of the errors. Equation (3.22) assumes that a reading's error is independent of the inputs and, when there are several readings, of the other readings' errors, so that the posterior (3.26) is a product over rows. No conservation relation states that, and none checks it directly. Suppose the hall of Section 3.5.4 is tested twice by the same sensor, each reading with $\sigma_T = 0.5\,\mathrm{K}$, both reading 27°C, with errors correlated through the calibration offset they share, $\mathrm{corr}(\epsilon_1, \epsilon_2) = 0.5$. Treated as independent rows, the pair adds the precision $2/\sigma_T^2 = 8\,\mathrm{K}^{-2}$ to the prior's 0.25. With the correlation, equation (3.29) holds with $\mathbf{V}_y$ a full matrix, its off-diagonal entry $0.5\,\sigma_T^2$ and its rows whitened by $\mathbf{V}_y^{-1/2}$, and the pair adds only

$$\mathbf{A}^{\mathsf{T}}\mathbf{V}_y^{-1}\mathbf{A} = \frac{2}{\sigma_T^2\,(1 + 0.5)} = 5.33\,\mathrm{K}^{-2}, \qquad \mathbf{A} = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

the precision of the pair's average, whose error has the variance $\sigma_T^2\,(1 + 0.5)/2$. Table 3.8 compares the two treatments with the single test of Table 3.7. Its columns are the precision

Table 3.8: Two condition tests read by one sensor of $\sigma_T = 0.5\,\mathrm{K}$, both reading 27°C, against the prior $\mathcal{N}(26, 2^2)$ and the limit $T_{\max} = 28°\mathrm{C}$: the pair treated as independent, the pair with $\mathrm{corr}(\epsilon_1, \epsilon_2) = 0.5$, and one test, which is also the limit of a correlation of one.

Readings [–]	precision added [K^{-2}]	mean [°C]	σ_{post} [K]	$\Pr(T > T_{\max} \mid \mathbf{y})$ [–]
two tests, treated as independent	8.00	26.9697	0.3482	0.00154
two tests, correlation 0.5	5.33	26.9552	0.4232	0.00678
one test	4.00	26.9412	0.4851	0.0145

the readings add, the posterior mean, the posterior standard deviation and the probability of exceeding $T_{\max} = 28°\mathrm{C}$.

The independent treatment reports an exceedance probability of 0.00154 against the 0.00678 the correlated pair supports, 4.4 times too small, and a posterior standard deviation of 0.348 K against 0.423 K. The error has the direction of Section 3.4.5's: a posterior narrower than the evidence supports, and a limit that looks safer than it is. At a correlation of one the second test adds nothing, and the posterior is that of one test.

The shared offset is a common cause, and the remedy of Section 3.4.6 applies: put it in the state. A bias entry $b_T \sim \mathcal{N}(0,\ 0.5\,\sigma_T^2)$ carried by both reading rows, with independent remainders of variance $0.5\,\sigma_T^2$, makes the two rows conditionally independent given (T, b_T), and the state (T, b_T) gives the posterior of T of the correlated row of Table 3.8 to every printed digit. The posterior of b_T, however, has the standard deviation 0.348 K against its prior's 0.354 K, with $\mathrm{corr}(T, b_T \mid \mathbf{y}) = -0.81$. The readings barely inform it: a shared offset moves both readings together, as a change of T does, and only the prior on b_T bounds it.

The same error on the health index is the case an asset manager meets. Take the commissioning fit of Section 3.5.6 with two condition tests at 30 kW in the commissioning window, read by the one hall sensor and both reading 23.2°C: the fit behind Table 3.1, with $n = 2$. Treated as independent, the two reading rows give $\sqrt{\Sigma_{gg}} = 0.0676$, $\hat{G} = 5.845\,\mathrm{kW/K}$ with the band 5.120 to 6.673 kW/K, and, read on the entry T, a probability 0.0062 that the coil was handed over below its requirement. With the shared offset at the correlation 0.5, $\mathbf{V}_y$ full, they give $\sqrt{\Sigma_{gg}} = 0.0825$, $\hat{G} = 5.882\,\mathrm{kW/K}$ with the band 5.004 to 6.913 kW/K, and the probability 0.0162. The independent band of g is 82.0 % of the correlated one, 18.0 % too narrow, and its probability that the coil is below its requirement is 2.6 times too small; the exact posteriors give 0.0053 against 0.0139, the same factor of 2.6, and the correlated route on T lies 0.0023 above its exact value. The bias entry repairs it here too. The state $(T, g, b_T)^{\mathsf{T}}$, with $b_T \sim \mathcal{N}(0,\ 0.5\,\sigma_T^2)$ on both reading rows, gives $\sqrt{\Sigma_{gg}} = 0.0825$ and the probability 0.0162, the correlated row, while the posterior of b_T barely moves, standard deviation 0.347 K against the prior's 0.354 K, with $\mathrm{corr}(g, b_T \mid \mathbf{y}) = 0.81$: a sensor that reads high and a coil in better condition explain the same two readings, and the readings cannot tell them apart.

The offset follows the sensor through the condition record. The commissioning test and a later condition test read by the same sensor share it. The level of the health index at the later test inherits the offset in full; the change since commissioning, the gradient $(-1,\ 1)^{\mathsf{T}}$ on the two levels, cancels it when both tests run at the same load and supply air, and only partly otherwise, as the change paragraph of Section 3.6.4 shows. An offset that itself changes between the two tests reads as a change of the health index. Chapter 4 writes the trend rows that tie one window to the next, and Chapter 7 sets a bias entry against a derate entry.

The same assumption hides in four other places a physics graph does not reach. A prior shared by the units of one design, such as the handbook value every coil of a fleet starts from, couples the units; partial pooling (a hierarchical model) puts the shared mean in the state,

Chapter 9. A physics tolerance σ_p is treated as independent from window to window where the model's error is systematic, such as an airflow pinned wrong in every test; if the error moves with the season it reads as a seasonal swing of the health index, Chapter 13. Readings of one window are averaged as if their errors were independent when they are correlated in time, Chapter 24, in Volume II. And a heat path the graph omits reads as a change of the health index. Each makes the posterior narrower than the evidence supports. What the graph adds is redundancy. When more rows bear on the state than it has entries, a misfit of the rows taken together, a reading far from what the other rows predict, or a comparison of drawings with and without the suspected cause can reveal the dependence; Chapters 7 and 12 derive them on the condition record. No test separates an offset that moves the readings the way fouling does, as the shared offset moves them along T: only a reference instrument or a second, independent quantity does.

Principle 3.3. *A physics graph carries dependence through shared physical quantities as entries of the state, so the independent product cannot be formed for them. The independence it still assumes is that of the errors: it must be stated, tested where the rows are redundant, and, where it fails, removed by putting the shared cause in the state.*

3.6 Predictive variance and the covariance of a joint posterior

Two variances complete the chapter: the law of total variance, which separates the randomness of future work orders from the uncertainty of their rate, and the covariance of a posterior over a joint state, which decides how uncertain a derived quantity, such as the change of a health index, is, and why the shape of the posterior must be read rather than its marginals.

3.6.1 The law of total variance

Suppose the number K_τ of a unit's corrective work orders during a future period of length τ is Poisson given the rate, $K_\tau \mid \lambda \sim \text{Poisson}(\lambda\tau)$, and the rate itself is uncertain, with posterior $p(\lambda \mid \mathbf{y})$. The predictive variance of K_τ is wanted. Two facts give it. The law of total expectation, which is the law of total probability applied inside an expectation, states that $\mathbb{E}[K_\tau \mid \mathbf{y}] = \mathbb{E}\big[\mathbb{E}[K_\tau \mid \lambda] \mid \mathbf{y}\big]$, the conditional mean averaged over the rate; and the variance of any quantity is the mean of its square minus the square of its mean. Then

$$\text{Var}(K_\tau \mid \mathbf{y}) = \mathbb{E}\big[\mathbb{E}[K_\tau^2 \mid \lambda] \mid \mathbf{y}\big] - \big(\mathbb{E}\big[\mathbb{E}[K_\tau \mid \lambda] \mid \mathbf{y}\big]\big)^2,$$

and writing the conditional second moment as $\mathbb{E}[K_\tau^2 \mid \lambda] = \text{Var}(K_\tau \mid \lambda) + \mathbb{E}[K_\tau \mid \lambda]^2$,

$$\text{Var}(K_\tau \mid \mathbf{y}) = \mathbb{E}\big[\text{Var}(K_\tau \mid \lambda) \mid \mathbf{y}\big] + \Big\{\mathbb{E}\big[\mathbb{E}[K_\tau \mid \lambda]^2 \mid \mathbf{y}\big] - \big(\mathbb{E}\big[\mathbb{E}[K_\tau \mid \lambda] \mid \mathbf{y}\big]\big)^2\Big\}.$$

The brace is the variance of the conditional mean, which gives the *law of total variance*,

$$\text{Var}(K_\tau \mid \mathbf{y}) = \mathbb{E}\big[\text{Var}(K_\tau \mid \lambda) \mid \mathbf{y}\big] + \text{Var}\big(\mathbb{E}[K_\tau \mid \lambda] \mid \mathbf{y}\big). \tag{3.34}$$

The first term is the expected variability of the work-order count when the rate is specified, the aleatory part. The second is the additional predictive variability because the rate is uncertain, the epistemic part. For a Poisson count the mean and the variance are equal, $\mathbb{E}[K_\tau \mid \lambda] = \text{Var}(K_\tau \mid \lambda) = \lambda\tau$, so that $\text{Var}(K_\tau \mid \mathbf{y}) = \mathbb{E}[\lambda\tau \mid \mathbf{y}] + \text{Var}(\lambda\tau \mid \mathbf{y})$, and since τ is fixed,

$$\text{Var}(K_\tau \mid \mathbf{y}) = \tau\,\mathbb{E}[\lambda \mid \mathbf{y}] + \tau^2\,\text{Var}(\lambda \mid \mathbf{y}). \tag{3.35}$$

The predictive variance has been separated into a process term and a rate term.

3.6.2 Corrective work orders over a future period

Take the posterior of the unit's corrective-failure rate from the state, the normal of the log-rate of Section 3.4.2, with $\hat{\omega} = \ln(8/3)$ and $\sigma_\omega = 0.5$, so that λ is lognormal about $\hat{\lambda} = 4/15 = 0.2667\,\mathrm{yr}^{-1}$. The standard lognormal moments give

$$\mathbb{E}[\lambda \mid \mathbf{y}] = \hat{\lambda}\,\mathrm{e}^{\sigma_\omega^2/2} = 0.3022\,\mathrm{yr}^{-1}, \qquad \mathrm{Var}(\lambda \mid \mathbf{y}) = \hat{\lambda}^2\,\mathrm{e}^{\sigma_\omega^2}(\mathrm{e}^{\sigma_\omega^2} - 1) = 0.025934\,\mathrm{yr}^{-2}.$$

Over the next year, $\tau = 1\,\mathrm{yr}$, the expected count of corrective work orders is $\mathbb{E}[K_\tau \mid \mathbf{y}] = 0.3022$, and by equation (3.35) the predictive variance is

$$\mathrm{Var}(K_\tau \mid \mathbf{y}) = 0.3022 + 1^2 \times 0.0259 = 0.3281.$$

The rate term is a share of 0.0790 of the predictive variance, and the process term the rest. The count 0.3022 is a mean, not the probability of a work order: a year with two corrective failures counts twice in the mean, so the probability of at least one, 0.2519 from the same normal (Section 3.4.2), lies below it.

The exact posterior $\mathrm{Gamma}(a, b) = \mathrm{Gamma}(4, 15\,\mathrm{yr})$ stands beside these, with the standard gamma moments $\mathbb{E}[\lambda \mid \mathbf{y}] = a/b = 0.2667\,\mathrm{yr}^{-1}$, one corrective failure in 3.75 yr, and $\mathrm{Var}(\lambda \mid \mathbf{y}) = a/b^2 = 0.017778\,\mathrm{yr}^{-2}$. The same two terms are

$$\mathrm{Var}(K_\tau \mid \mathbf{y}) = 1\left(\frac{4}{15}\right) + 1^2\left(\frac{4}{15^2}\right) = 0.2667 + 0.0178 = 0.2844,$$

the mean 0.2667 against the reported 0.3022 and the variance 0.2844 against 0.3281; the exact probability of a corrective failure in the year, 0.2275, again lies below its mean. The share of the rate term under the gamma is

$$\frac{\tau^2 a/b^2}{\tau a/b + \tau^2 a/b^2} = \frac{\tau}{b + \tau},$$

which for $\tau = 1\,\mathrm{yr}$ and $b = 15\,\mathrm{yr}$ is $1/16 = 0.0625$, against the reported 0.0790. The gamma's result can be checked against its full predictive distribution. A Poisson count whose mean is averaged over a gamma rate has, by a standard result, the negative binomial distribution with shape a and success probability $\pi = b/(b + \tau)$, whose mean is $a(1 - \pi)/\pi = a\tau/b$ and whose variance is $a(1 - \pi)/\pi^2 = a\tau/b + a\tau^2/b^2$: exactly the two terms of equation (3.35). For $\pi = 15/16 = 0.9375$ they are 0.2667 and 0.2844. A quadrature of the state route at 1 yr returns its two moments, 0.3022 and 0.025934, to the printed digits.

Table 3.9 extends the calculation to longer periods. The first column is the period; the next four are the reported values from the state, the expected count, which is also the process term, the rate term, the total and the share of the rate term; and the last three the gamma's mean, variance and share. Over a year most of the predictive variance comes from the random failure process, and a longer work-order log, which shrinks only the rate term, would barely change the forecast of next year's work orders. The two terms are equal at 11.65 yr (15 yr under the gamma, as long as the exposure behind the posterior), and over 30 yr the rate term is 0.720 of the total (0.667 under the gamma). The gaps grow with the period, because the lognormal's mean rate lies above the gamma's and its upper tail is heavier. The 30 yr row exceeds the life of a cooling unit; it is kept for the trend of the share. Whether a longer log will meaningfully improve a work-order forecast depends on this split: more evidence shrinks the second term, and the first remains as long as the future count is modeled as a Poisson process.

3.6.3 One posterior over a joint state

The examples so far had one unknown at a time. A plant graph carries many at once. Suppose the state vector holds the air temperatures of two halls, the log-conductance of an air path and

Table 3.9: Predictive mean and variance of a unit's corrective work-order count over a future period τ, equation (3.35): reported from the normal of the log-rate, with the posterior Gamma$(4, 15\,\mathrm{yr})$ beside it; the first row is the next year.

τ [yr]	mean = process term [–]	rate term [–]	variance [–]	rate share [–]	gamma mean [–]	gamma variance [–]	gamma share [–]
1	0.3022	0.0259	0.3281	0.0790	0.2667	0.2844	0.0625
5	1.5109	0.6483	2.1592	0.3003	1.3333	1.7778	0.2500
15	4.5326	5.8351	10.3677	0.5628	4.0000	8.0000	0.5000
30	9.0652	23.3405	32.4057	0.7203	8.0000	24.0000	0.6667

the corrective-failure rates of two units,

$$\mathbf{z} = \begin{bmatrix} T_1 & T_2 & g & \lambda_A & \lambda_B \end{bmatrix}^{\mathsf{T}},$$

and the evidence $\mathbf{y}$ gathers the condition tests, the work-order logs and the inspection results. Everything known after the evidence is the joint posterior $p(\mathbf{z} \mid \mathbf{y}, M)$, where M names the model: the graph, its rows and its priors. Different questions are asked of this same distribution:

State estimation, $p(T_1 \mid \mathbf{y}, M)$: the marginal of one entry.

Threshold probability, $\Pr(T_2 > T_{\max} \mid \mathbf{y}, M)$: a tail of another marginal.

Parameter comparison, $\Pr(\lambda_A > \lambda_B \mid \mathbf{y}, M)$: a tail of the derived quantity $\lambda_A - \lambda_B$, which depends on the joint distribution of the two rates.

Expected derived quantity, $\mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}, M]$, such as the expected number of corrective work orders of the two units over the next year: the average of a function of the state.

Prediction, $\Pr(T_f > \tau \mid \mathbf{y}, M)$: a posterior predictive probability, equation (3.16) with the rate taken from the state.

The methods of computation differ, but every one of them depends on a consistent probability model of the physical asset and the evidence, and none needs a separate inference.

3.6.4 The posterior covariance and the variance budget

When the posterior can be approximated locally by a multivariate Gaussian, $\mathbf{z} \mid \mathbf{y}, M \approx \mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$, the covariance matrix $\boldsymbol{\Sigma}$ describes the uncertainty of each entry and how the entries vary together. The approximation takes the most probable state as the mean and the inverse curvature of the objective there as the covariance. For rows linear in the state it is exact, $\boldsymbol{\Sigma} = (\mathbf{J}^{\mathsf{T}}\mathbf{J})^{-1}$ by Section 3.5.3; for nonlinear rows it is the Laplace approximation, which Chapter 4 derives with its curvature term. For two entries,

$$\boldsymbol{\Sigma} = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12} & \Sigma_{22} \end{bmatrix}, \qquad \operatorname{corr}(z_1, z_2 \mid \mathbf{y}) = \frac{\Sigma_{12}}{\sqrt{\Sigma_{11}\Sigma_{22}}}.$$

The diagonal entries are the marginal variances, $\Sigma_{11} = \operatorname{Var}(z_1 \mid \mathbf{y})$; the off-diagonal entry is the covariance, and $\operatorname{corr}(z_1, z_2 \mid \mathbf{y})$ the correlation. The off-diagonal entries matter because physical relations and shared observations introduce posterior dependence between unknowns; they are not zero merely because the priors of the entries were independent. The hall of Section 3.5.6 is one case. Another carries the supply air as an entry: the state $\mathbf{z} = (T,\ T_c,\ g)^{\mathsf{T}}$ for the illustrative hall with a weak air path of Figure 3.2(c), with $Q = 30\,\mathrm{kW}$, the log-conductance taken against that hall's own reference, $g = \ln(G/G_0)$ with $G_0 = 3\,\mathrm{kW/K}$, independent priors

$T_c \sim \mathcal{N}(20, 1^2)$ and $g \sim \mathcal{N}(0, 0.05^2)$ (the 0.15 kW/K of the figure), a reading with $\sigma_T = 0.5\,\mathrm{K}$ equal to the prior prediction 30°C, and the physics row $(T - T_c - 10\,\mathrm{K}\,\mathrm{e}^{-g})/\sigma_p$ with $\sigma_p = 0.01\,\mathrm{K}$. At the prior means every residual is zero and the rows of $\mathbf{J}$ are $(1,\ -1,\ 10)/\sigma_p$ for the physics, $(1,\ 0,\ 0)/\sigma_T$ for the reading, and $(0,\ 1,\ 0)$ and $(0,\ 0,\ 1)/0.05$ for the priors; equation (3.28) gives

$$\boldsymbol{\Sigma} = (\mathbf{J}^\mathsf{T}\mathbf{J})^{-1} = \begin{bmatrix} 0.2083 & 0.1667 & -0.0042 \\ 0.1667 & 0.3334 & 0.0167 \\ -0.0042 & 0.0167 & 0.0021 \end{bmatrix},$$

and the correlation of T_c and g is $0.0167/\sqrt{0.3334 \times 0.0021} = 0.632$ (unrounded entries; 0.6325 at $\sigma_p = 0.001\,\mathrm{K}$). Without the reading row the same computation gives zero. A single reading leaves T_c and G positively correlated, since a warmer supply air and a larger conductance explain the same reading, and the correlation enters through the physics row, the only row with entries in the columns of both.

Suppose a derived quantity $f(\mathbf{z})$ has the gradient $\mathbf{g} = \partial f/\partial \mathbf{z}$ at $\hat{\mathbf{z}}$, a column with one entry per entry of the state. With $\boldsymbol{\delta} = \mathbf{z} - \hat{\mathbf{z}}$ the deviation of the state from its mean, the first-order change is $f - f(\hat{\mathbf{z}}) \approx \mathbf{g}^\mathsf{T}\boldsymbol{\delta}$, a scalar equal to its own transpose, and the covariance is $\boldsymbol{\Sigma} = \mathbb{E}[\boldsymbol{\delta}\boldsymbol{\delta}^\mathsf{T}]$, so that

$$\mathrm{Var}(f) \approx \mathbb{E}[\mathbf{g}^\mathsf{T}\boldsymbol{\delta}\,\boldsymbol{\delta}^\mathsf{T}\mathbf{g}] = \mathbf{g}^\mathsf{T}\,\mathbb{E}[\boldsymbol{\delta}\boldsymbol{\delta}^\mathsf{T}]\,\mathbf{g} = \mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}. \tag{3.36}$$

This is the *variance budget* of the output. Written out entry by entry, $\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}$ has one term per entry of the state, the squared sensitivity of the output to the entry times the variance of the entry, plus a cross term $2\,\nabla_i\nabla_j\,\Sigma_{ij}$, with $\nabla_i = \partial f/\partial z_i$, for every pair of entries that move together. For the sum $f = z_1 + z_2$ the gradient is $\mathbf{g} = (1,\ 1)^\mathsf{T}$, and the budget is exact because f is linear,

$$\mathrm{Var}(z_1 + z_2) = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12} & \Sigma_{22} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \Sigma_{11} + \Sigma_{22} + 2\Sigma_{12}. \tag{3.37}$$

With $\Sigma_{11} = \Sigma_{22} = 1\,\mathrm{K}^2$ (illustrative), the variance of the sum is $0.4\,\mathrm{K}^2$ for $\Sigma_{12} = -0.8\,\mathrm{K}^2$, $2\,\mathrm{K}^2$ for independent entries and $3.6\,\mathrm{K}^2$ for $\Sigma_{12} = 0.8\,\mathrm{K}^2$. A negative covariance makes the output more precisely known than the individual entries suggest; a positive one makes it less.

The same budget says how well the change since commissioning is known. Let $z_0 = g_0$ and $z_1 = g_t$ be the log-conductance at the commissioning test and at a later condition test, each known to the standard deviation 0.10, so that $\Sigma_{00} = \Sigma_{11} = 0.0100$. The change $f = z_1 - z_0$ has the gradient $\mathbf{g} = (-1,\ 1)^\mathsf{T}$, and the budget (3.36), exact because f is linear, gives

$$\mathrm{Var}(z_1 - z_0) = \begin{bmatrix} -1 & 1 \end{bmatrix} \begin{bmatrix} \Sigma_{00} & \Sigma_{01} \\ \Sigma_{01} & \Sigma_{11} \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \Sigma_{00} + \Sigma_{11} - 2\Sigma_{01}. \tag{3.38}$$

The cross term carries the sign opposite to the covariance, because the two entries of the gradient have opposite signs. For $\Sigma_{01} = +0.008$, 0 and -0.008 the variance of the change is 0.0040, 0.0200 and 0.0360, standard deviations of 0.0632, 0.1414 and 0.1897, while the sum takes them in the reverse order, 0.1897, 0.1414 and 0.0632. A positive covariance, such as the one a shared sensor offset puts between the two levels, makes the change better known than either level; a negative one makes it worse. Chapter 4's trend rows rest on this.

Table 3.10 works the budget for the hall air of equation (3.21), with the load known, $Q = 30\,\mathrm{kW}$, and the two unknowns of Figure 3.2(c) as the state $\mathbf{z} = (T_c,\ G)^\mathsf{T}$: T_c with standard deviation 1 K and G with mean 3 kW/K and standard deviation 0.15 kW/K. The gradient of T and the covariance are

$$\mathbf{g} = \begin{bmatrix} \partial T/\partial T_c \\ \partial T/\partial G \end{bmatrix} = \begin{bmatrix} 1 \\ -Q/G^2 \end{bmatrix} = \begin{bmatrix} 1 \\ -3.333 \end{bmatrix}, \qquad \boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0.15\,\varrho \\ 0.15\,\varrho & 0.0225 \end{bmatrix}, \qquad \varrho = \mathrm{corr}(T_c, G),$$

with $\mathbf{g}$ in K per unit of each entry and $\boldsymbol{\Sigma}$ in the products of the entries' units, so that $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g} = (1.00 + 0.25 - 1.00\,\varrho)\,\mathrm{K}^2$: the supply-air term $1^2 \times 1$, the conductance term $3.333^2 \times 0.0225$ and the cross term $2 \times 1 \times (-3.333) \times 0.15\,\varrho$. The columns of the table are the correlation, the three terms, the variance and the standard deviation of T. The two marginals are the same in every row, and the standard deviation of the hall air ranges from 0.67 to 1.43 K with the correlation alone. The budget is first order, since T is curved in G; a quadrature over the same joint normal gives 0.6701, 1.1203 and 1.4357 K against the budget's 0.6708, 1.1180 and 1.4318 K, gaps of at most 0.0039 K.

Table 3.10: Variance budget of the hall air $T = T_c + Q/G$, equation (3.36), with $Q = 30\,\mathrm{kW}$, G of mean $3\,\mathrm{kW/K}$ and standard deviation $0.15\,\mathrm{kW/K}$, T_c of standard deviation 1 K, and correlation $\varrho = \mathrm{corr}(T_c, G)$ (illustrative).

ϱ [–]	supply-air term [K^2]	conductance term [K^2]	cross term [K^2]	Var(T) [K^2]	sd [K]
−0.8	1.00	0.25	+0.80	2.05	1.43
0	1.00	0.25	0	1.25	1.12
0.8	1.00	0.25	−0.80	0.45	0.67

Two consequences follow from equation (3.36). An entry can be very uncertain and still not matter: if the output is insensitive to it, its whole term is near zero, and narrowing its variance buys nothing. And an entry can be poorly known while the output is well known: when the sensitivities and the covariance have opposite signs, the cross term subtracts from the diagonal terms. A test is worth scheduling when it shrinks a variance Σ_{kk} whose sensitivity is large, and worthless when the sensitivity is near zero.

3.6.5 Reading the shape of the posterior

The second consequence deserves a worked case, because it is where marginal spreads mislead. Carry the hall air, the load and the conductance as entries, the last two through the logarithms $\ell_Q = \ln(Q/Q_{\mathrm{ref}})$ and $g = \ln(G/G_0)$, with G_0 the reference of this illustrative hall, each with an independent prior of mean 0 and standard deviation 0.10, that is about 10 %, and with the supply air pinned. The references are chosen so that $Q_{\mathrm{ref}}/G_0 = 10\,\mathrm{K}$, a temperature rise of 10 K at the prior means, and a sensor with $\sigma_T = 0.2\,\mathrm{K}$ reads that temperature, $y_T = T_c + 10\,\mathrm{K}$. The state, its four rows and their 4×3 Jacobian are

$$\mathbf{z} = \begin{bmatrix} T \\ \ell_Q \\ g \end{bmatrix}, \qquad r(\mathbf{z}) = \begin{bmatrix} (T - y_T)/\sigma_T \\ \ell_Q/0.10 \\ g/0.10 \\ (T - T_c - 10\,\mathrm{K}\,\mathrm{e}^{\ell_Q - g})/\sigma_p \end{bmatrix},$$

$$\mathbf{J} = \begin{bmatrix} 1/\sigma_T & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 10 \\ 1/\sigma_p & -10\,\mathrm{e}^{\ell_Q - g}/\sigma_p & 10\,\mathrm{e}^{\ell_Q - g}/\sigma_p \end{bmatrix},$$

with the physics row of equation (3.21) carrying the tolerance $\sigma_p = 0.001\,\mathrm{K}$. At the prior means, $\mathbf{z}_{\mathrm{p}} = (T_c + 10\,\mathrm{K},\ 0,\ 0)^{\mathsf{T}}$, every residual is zero, the normal equations (3.27) give a zero step, and $\hat{\mathbf{z}} = \mathbf{z}_{\mathrm{p}}$; the curvature term that Chapter 4 adds for nonlinear rows is a sum of residuals

times second derivatives and vanishes with the residuals. There the Jacobian is

$$\mathbf{J} = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 10 \\ 1000 & -10\,000 & 10\,000 \end{bmatrix},$$

whose last two columns carry the relation between load and conductance: both are zero in the reading row and equal and opposite in the physics row, so the direction $(0,\ 1,\ 1)^{\mathsf{T}}$, load and conductance raised together, gives $\mathbf{J}\,(0,\ 1,\ 1)^{\mathsf{T}} = (0,\ 10,\ 10,\ 0)^{\mathsf{T}}$ and is seen by the prior rows alone. Multiplying by $\mathbf{J}^{\mathsf{T}}$,

$$\mathbf{J}^{\mathsf{T}}\mathbf{J}\,(0,\ 1,\ 1)^{\mathsf{T}} = 100\,(0,\ 1,\ 1)^{\mathsf{T}},$$

so $(0,\ 1,\ 1)^{\mathsf{T}}$ is an eigenvector of $\mathbf{J}^{\mathsf{T}}\mathbf{J}$ with the eigenvalue $100 = 1/0.10^2$, the prior precision alone, and of $\boldsymbol{\Sigma} = (\mathbf{J}^{\mathsf{T}}\mathbf{J})^{-1}$ with the eigenvalue 0.0100, the prior variance: the reading has left that direction untouched. Inverting $\mathbf{J}^{\mathsf{T}}\mathbf{J}$ numerically, equation (3.28) gives

$$\boldsymbol{\Sigma} = \begin{bmatrix} 0.039216 & 0.001961 & -0.001961 \\ 0.001961 & 0.005098 & 0.004902 \\ -0.001961 & 0.004902 & 0.005098 \end{bmatrix}.$$

For a Gaussian, the integral of equation (3.18) over T leaves the block of $\boldsymbol{\Sigma}$ in the rows and columns of ℓ_Q and g, the lower right 2×2 block. Its precision is not the corresponding block of $\mathbf{J}^{\mathsf{T}}\mathbf{J}$ but, by the block-inversion formula of Section 3.5.3 applied to $\boldsymbol{\Lambda} = \mathbf{J}^{\mathsf{T}}\mathbf{J}$, the Schur complement of the T block, $\boldsymbol{\Lambda}_{qq} - \boldsymbol{\Lambda}_{qT}\boldsymbol{\Lambda}_{TT}^{-1}\boldsymbol{\Lambda}_{Tq}$ with q the indices of ℓ_Q and g:

$$\boldsymbol{\Lambda}_{qq} - \boldsymbol{\Lambda}_{qT}\boldsymbol{\Lambda}_{TT}^{-1}\boldsymbol{\Lambda}_{Tq} = \begin{bmatrix} 2599.94 & -2499.94 \\ -2499.94 & 2599.94 \end{bmatrix} \longrightarrow \begin{bmatrix} 2600 & -2500 \\ -2500 & 2600 \end{bmatrix} \quad (\sigma_p \to 0),$$

against the block $\boldsymbol{\Lambda}_{qq}$ of $\mathbf{J}^{\mathsf{T}}\mathbf{J}$ itself, whose entries are $10^8 + 100$ on the diagonal and -10^8 off it. Each marginal standard deviation is $\sqrt{0.005098} = 0.0714$, down only from 0.100, and the correlation is $0.004902/0.005098 = 0.962$. The eigenvectors of the block are $(1,\ 1)^{\mathsf{T}}/\sqrt{2}$ and $(1,\ -1)^{\mathsf{T}}/\sqrt{2}$, with the eigenvalues 0.005098 ± 0.004902, that is 0.010000 and 0.000196: the posterior is as wide as the prior along the first and 7.1 times narrower along the second. The budget (3.36) with the gradients $\mathbf{g} = (0,\ 1,\ -1)^{\mathsf{T}}$ and $(0,\ 1,\ 1)^{\mathsf{T}}$ gives the variance of the difference, $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g} = 0.000392$, a standard deviation of 0.0198, and that of the sum, 0.0200, a standard deviation of 0.1414, exactly its prior value $\sqrt{2} \times 0.10$. Repeated at $\sigma_p = 0.01\,\mathrm{K}$, the block entries print the same, 0.005098 and 0.004902, and the correlation prints 0.961 against 0.962 at 0.001 K; the limit is reached at 0.001 K.

Remark 3.2. Eliminating T through the physics row turns the reading into the linearized row $(\ell_Q - g)/0.02$, since $\sigma_T/(T - T_c) = 0.02$, and the same block follows from a 3×2 Jacobian; along the difference alone, the prior precision 50 plus the reading's 2500 gives $1/2550 = 0.000392$.

Figure 3.7 draws the two distributions in the plane of the two entries. The horizontal axis is ℓ_Q and the vertical axis g; the dashed gray circle encloses 95 % of the prior, the green ellipse 95 % of the posterior, and the thin line is $\ell_Q - g = 0$, the value of the difference that the reading fixed. The posterior is a narrow ellipse lying along that line, its axes the two eigenvectors of the block: it is long in the direction the reading cannot see, $(1,\ 1)^{\mathsf{T}}$, where it reaches the prior circle because the reading added nothing there, and short in the direction it measured, $(1,\ -1)^{\mathsf{T}}$.

Read by its marginals, this posterior says that neither the load nor the conductance is well known, about 7 % each. Read by its shape, it says that their ratio, and with it the hall air at the present load, is known to 2 %. For the asset manager this is a condition test at an unmetered load: a test whose load is not a known input fixes only Q/G, so that the health index alone is known to 7 % while the ratio is known to 2 %. That is why the load must be a known input of

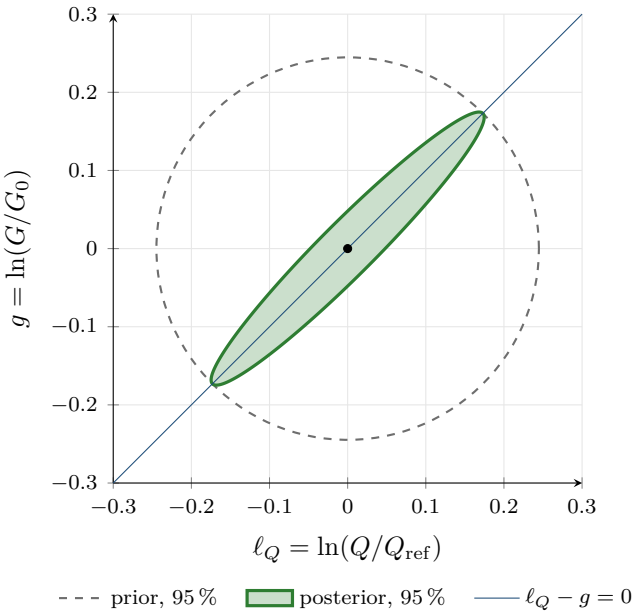


Figure 3.7: Prior and posterior of the log-load ℓ_Q and the log-conductance g after one test reading that fixes $\ell_Q - g$ within 0.02 (illustrative): the (ℓ_Q, g) block of the posterior of $\mathbf{z} = (T, \ell_Q, g)^{\mathsf{T}}$. Prior: independent, standard deviations 0.10. Posterior: standard deviations 0.0714, correlation 0.962; standard deviation of $\ell_Q - g$, 0.0198; of $\ell_Q + g$, 0.1414. Contours enclose 95 % ($\chi^2_{2,0.95} = 5.991$).

the physics rows before a test reads condition. A prediction at a different, specified load Q', $T' = T_c + Q'/G$, has a gradient with respect to (T, ℓ_Q, g) proportional to $(0,\ 0,\ 1)^{\mathsf{T}}$, so $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$ picks out Σ_{gg} and the prediction inherits its 7 % spread; a prediction that scales the present load, such as the hall at twice the unknown Q, has a gradient proportional to $(0,\ 1,\ -1)^{\mathsf{T}}$, the short axis of the ellipse, and its temperature rise is known to the same 2 % relative as the present one; the absolute spread doubles. Which statement holds depends on the gradient of the question, and the covariance, not the list of marginal variances, carries the answer.

Principle 3.4. *Read the posterior's shape, not its marginals. Wide marginal spreads with a strong correlation mean that the evidence determines a combination of the entries; a quantity that depends on that combination is well known, and the entries are not unidentified merely because each alone is uncertain.*

3.6.6 Mixtures over structure

A Gaussian approximation is most useful for continuous uncertainty near a well-identified posterior mode. Conditional probability is more general. If the asset may operate under either of two structures, $M \in \{M_A, M_B\}$ (two drawings of the plant), the law of total probability, applied over the models with $\mathbf{z}$ conditioned on the evidence throughout, gives

$$p(\mathbf{z} \mid \mathbf{y}) = \sum_j p(\mathbf{z}, M_j \mid \mathbf{y}) = \sum_j p(\mathbf{z} \mid \mathbf{y}, M_j)\ \Pr(M_j \mid \mathbf{y}), \tag{3.39}$$

where the second step is the multiplication rule. The weights follow from Bayes' theorem over the models, $\Pr(M_j \mid \mathbf{y}) = p(\mathbf{y} \mid M_j)\ \Pr(M_j) / \sum_i p(\mathbf{y} \mid M_i)\ \Pr(M_i)$, so that the evidence of each model, the normalizing constant of its own posterior, sets its share. The expression assumes

that the relevant state $\mathbf{z}$ has a common interpretation across the models. Such a posterior need not be Gaussian, even when each model's posterior is.

A two-structure illustration on the hall air shows what is lost. Let the drawing M_A, the air path as designed, give $T \mid \mathbf{y}, M_A \sim \mathcal{N}(26.9, 0.5^2)$, and a drawing M_B with a blocked filter give $\mathcal{N}(28.5, 0.5^2)$, with posterior model probabilities 0.7 and 0.3 (illustrative). By the law of total expectation the mixture mean is $0.7 \times 26.9 + 0.3 \times 28.5 = 27.38°\mathrm{C}$; by the law of total variance, equation (3.34) with the model in place of the rate, its variance is the mean of the within-model variances plus the variance of the model means,

$$\mathrm{Var}(T \mid \mathbf{y}) = 0.5^2 + 0.7 \times 0.3 \times (28.5 - 26.9)^2 = 0.25 + 0.5376 = 0.7876\,\mathrm{K}^2,$$

a standard deviation of 0.887 K. The probability of exceeding 28°C is the weighted sum of the two tails,

$$\Pr(T > 28°\mathrm{C} \mid \mathbf{y}) = 0.7\,(1 - \Phi(2.2)) + 0.3\,(1 - \Phi(-1)) = 0.7 \times 0.0139 + 0.3 \times 0.8413 = 0.262.$$

The most probable drawing alone would report 0.014, and a single Gaussian with the mixture's mean and variance would report $1 - \Phi(0.699) = 0.242$ and hide the two separate modes, 3.2 standard deviations apart. The law of total probability combines the uncertainty over the structure with the uncertainty over the numerical state, and this combination becomes central in the diagnosis of a low health index and in deciding which drawing of the asset its condition record supports. The next chapter assembles these pieces into the complete workflow of probabilistic inference, from the plant graph and its likelihood to the posterior, its marginals, its expectations and the probabilities of functions of the state.

3.7 What to retain

Definition 3.1 (Conditioning, marginalization and dependence)**.** A conditional probability restricts attention to the situations consistent with the evidence and renormalizes, $\Pr(A \mid B) = \Pr(A \cap B)/\Pr(B)$; Bayes' theorem reverses the conditioning through the prior and the evidence; marginalization sums or integrates over every unknown the question does not name; and the posterior covariance of a joint state, not its marginal variances, decides how uncertain a derived quantity is.

Four ideas carry forward. Asset-management questions contain two kinds of uncertainty: aleatory uncertainty is the scatter of the next condition test or the next corrective failure, and epistemic uncertainty is incomplete knowledge of the asset, its conductance or a sensor's calibration; more tests reduce the second and leave the first, and the law of total variance separates them in a prediction. One test read as exact is not the health index: the dashboard's 0.962 against design is what many readings averaging 23.2°C would give, while the fit of the one reading puts the commissioned value at 0.986, with a band from 0.819 to 1.188 that contains the design value. Conditional probability is how information is incorporated: Bayes' theorem updates a probability when an observation arrives, the base rate decides how far a flag moves it, and the share of a flag's work orders that find fouling is not the flag's detection rate. Asset quantities are often dependent: a shared filter, a conservation relation or a common supply air creates relations among components and unknown states, and conditional independence simplifies a calculation only where the model justifies it. A physics graph puts the shared physical causes in the state; the independence it still assumes is that of the errors, and a shared error, such as the offset of the sensor that reads every condition test, is removed by carrying it as an entry too. Marginalization answers a focused question without pretending that the other unknowns are known, whether they are nuisance quantities, competing grades of a coil or competing drawings of the asset.

Together these establish the mechanism of inference: the prior $p(\mathbf{z})$ and the observation model $p(\mathbf{y} \mid \mathbf{z})$ combine into the posterior $p(\mathbf{z} \mid \mathbf{y})$, the update multiplying the prior by the likelihood and normalizing the product. For Gaussian rows the product is $\exp(-\frac{1}{2} r^{\mathsf{T}} r)$ over the whole state, physics rows included; the normal equations give its most probable state, $\boldsymbol{\Sigma} = (\mathbf{J}^{\mathsf{T}}\mathbf{J})^{-1}$ its covariance, exact when the rows are linear, and the conditional of the joint Gaussian of state and readings, through the Schur complement, gives the same answer. Relations between quantities are read in that space: a reading informs an entry it does not measure through a row of $\mathbf{J}$ that carries both, and a combination of entries, such as the change of a health index since commissioning, is well known when it lies along a short eigenvector of $\boldsymbol{\Sigma}$, which $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$ measures. Once the posterior is available, the asset manager's question decides the next calculation: a marginal probability, an expectation, a prediction, a model comparison or a decision. Table 3.11 lists the chapter's terms with their expressions, units and the subsection that defines each.

Table 3.11: The terms of uncertainty and conditional probability and where each is defined.

Term	Expression	Units	Meaning	Defined in
Aleatory uncertainty	$\Pr(y_T > c \mid T) = 1 - \Phi((c - T)/\sigma_T)$	–	tests vary with the coil known	3.1.1
Epistemic uncertainty	$p(g \mid \mathbf{y})$	–	knowledge of a fixed quantity	3.1.2
Health index	$h = G/G_{\mathrm{ref}} = \mathrm{e}^{g}$	–	conductance against design; the commissioning fit is the commissioned value	3.1.2
Predictive exceedance	$\int \Pr(y_{\mathrm{new}} > c \mid T)\, p(T \mid \mathbf{y})\, \mathrm{d}T$	–	next test, averaged over the hall	3.1.3
Conditional probability	$\Pr(A \cap B)/\Pr(B)$	–	probability of A among cases of B	3.2.2
Multiplication rule	$\Pr(A \cap B) = \Pr(B \mid A)\Pr(A)$	–	joint from conditional	3.2.3
Bayes' theorem	$\Pr(B \mid A)\Pr(A)/\Pr(B)$	–	reverses the conditioning	3.2.4
Posterior odds	likelihood ratio $\times$ prior odds	–	Bayes without the evidence	3.3.3
Law of total probability	$\sum_s \Pr(H \mid s)\Pr(s)$	–	marginal of an observation	3.4.1
Marginal likelihood	$\int p(\mathbf{y} \mid \lambda) p(\lambda)\, \mathrm{d}\lambda$	–	evidence; normalizer	3.4.2
Predictive survival	$\int \mathrm{e}^{-\lambda\tau} p(\lambda \mid \mathbf{y})\, \mathrm{d}\lambda = 0.7481$	–	reported, over the log-rate normal; next year	3.4.2
Its exact value	$(b/(b+\tau))^{a} = 0.7725$	–	same average over the gamma; gap 0.0244	3.4.2
Nuisance marginal	$\int p(g, b_T \mid \mathbf{y})\, \mathrm{d}b_T$	–	integrate out, do not maximize	3.4.3
Conditional independence	$X_1 \perp X_2 \mid C$	–	joint given c factorizes	3.4.6
Common-cause joint	$\sum_c \Pr(F_A \mid c)\Pr(F_B \mid c)\Pr(c)$	–	coil and fan with a shared filter	3.4.5
Correlated reading errors	$\mathbf{A}^{\mathsf{T}}\mathbf{V}_y^{-1}\mathbf{A}$, $\mathbf{V}_y$ full; or a shared bias entry	K^{-2}	the precision the readings truly add	3.5.7
Gaussian rows	$p(\mathbf{z} \mid \mathbf{y}) \propto \exp(-\frac{1}{2} r^{\mathsf{T}} r)$	–	posterior of the whole state	3.5.2
Normal equations	$(\mathbf{J}^{\mathsf{T}}\mathbf{J})\,\boldsymbol{\Delta} = -\mathbf{J}^{\mathsf{T}} r$, $\boldsymbol{\Sigma} = (\mathbf{J}^{\mathsf{T}}\mathbf{J})^{-1}$	varies	most probable state, covariance	3.5.3
Precisions add	$\boldsymbol{\Sigma}^{-1} = \boldsymbol{\Sigma}_0^{-1} + \mathbf{J}_y^{\mathsf{T}}\mathbf{J}_y$	varies	prior plus reading information	3.5.3
Gaussian conditional	$\boldsymbol{\Sigma}_{zz} - \boldsymbol{\Sigma}_{zy}\boldsymbol{\Sigma}_{yy}^{-1}\boldsymbol{\Sigma}_{yz}$	varies	Schur complement	3.5.3
Exceedance probability	$1 - \Phi\big((T_{\max} - \mu_{\mathrm{post}})/\sigma_{\mathrm{post}}\big)$	–	tail of the posterior	3.5.5
Law of total variance	$\mathbb{E}[\mathrm{Var}(K_\tau \mid \lambda)] + \mathrm{Var}(\mathbb{E}[K_\tau \mid \lambda])$	–	process plus rate terms	3.6.1
Posterior covariance	$\boldsymbol{\Sigma}$, $\mathrm{corr}(z_1, z_2) = \Sigma_{12}/\sqrt{\Sigma_{11}\Sigma_{22}}$	varies	joint spread of the state	3.6.4
Variance budget	$\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$	varies	sensitivity^2 $\times$ variance, plus cross terms	3.6.4
Change since commissioning	$\Sigma_{00} + \Sigma_{11} - 2\Sigma_{01}$	–	a shared offset cancels in the change	3.6.4
Mixture posterior	$\sum_j p(\mathbf{z} \mid \mathbf{y}, M_j)\Pr(M_j \mid \mathbf{y})$	varies	uncertainty over structure	3.6.6

3.8 Exercises

Exercise 3.1 (Formulation)**.** (a) In 2000 inspections of units of this design the coil was found fouled in 100, the fan worn in 60, and both in 25. Compute $\Pr(F_A)$, $\Pr(F_B)$, $\Pr(F_A \cap F_B)$, $\Pr(F_B \mid F_A)$ and $\Pr(F_A \mid F_B)$, and state whether the two findings are independent. (0.05, 0.03, 0.0125, $25/100 = 0.25$ and $25/60 = 0.417$; not independent, since the independent product is $0.05 \times 0.03 = 0.0015$, not 0.0125.) (b) A cooling unit carries two fans and loses its cooling only when both fail; $\Pr(F_A) = 0.05$, $\Pr(F_B) = 0.04$ and $\Pr(F_A \cap F_B) = 0.015$ in a year. Compute the joint probability under an incorrect independence assumption, the probability that the unit keeps its cooling, and $\Pr(F_B \mid F_A)$, and state what the redundancy is worth. ($0.05 \times 0.04 = 0.002$; $1 - 0.015 = 0.985$, against 0.998 under independence; $0.015/0.05 = 0.30$; the two fans fail together far more often than independence predicts, by a factor of 7.5.)

Exercise 3.2 (Formulation)**.** A test at an unmetered load reads the hall $T = T_c + Q/G$ with T_c known, $G > 0$, and both Q and G uncertain. Starting from $\Pr(T > T_{\max} \mid \mathbf{y})$, write an equivalent event in Q and G, express the probability as an integral over their joint posterior without assuming independence, and give the requirement level G^* on the conductance at a known load. (Since $G > 0$, $T_c + Q/G > T_{\max} \iff Q > G(T_{\max} - T_c)$, so $\Pr(T > T_{\max} \mid \mathbf{y}) = \iint \mathbf{1}[Q > G(T_{\max} - T_c)]\, p(Q, G \mid \mathbf{y})\, \mathrm{d}Q\, \mathrm{d}G$, which keeps the correlation a reading of T creates between Q and G; at a known load the event is $G < G^* = Q/(T_{\max} - T_c)$, which is 5.0000 kW/K at 30 kW and 24°C, and at 60 kW and 27.5°C is 6.3158 kW/K.)

Exercise 3.3 (Derivation)**.** (a) A work-order trigger on a condition flag has a base rate of fouling $\Pr(F) = 0.02$, $\Pr(H \mid F) = 0.95$ and $\Pr(H \mid \neg F) = 0.05$. Derive the probability that the coil is fouled when the trigger fires, and explain why it is not 95 %. (0.95 is $\Pr(H \mid F)$, not $\Pr(F \mid H)$, and the low base rate makes most work orders false; $\Pr(H) = 0.95 \times 0.02 + 0.05 \times 0.98 = 0.068$, so $\Pr(F \mid H) = 0.019/0.068 = 0.279$.) (b) With prior $\Pr(F) = 0.10$ that the coil is fouled, the condition test's flag A has $\Pr(H_A \mid F) = 0.9$, $\Pr(H_A \mid \neg F) = 0.1$, and the inspection B has $\Pr(H_B \mid F) = 0.8$, $\Pr(H_B \mid \neg F) = 0.2$; the two are conditionally independent given the coil's state. Derive $\Pr(F \mid H_A, H_B)$ and explain why the assumption is needed. (The likelihoods multiply only under conditional independence given each state of the coil, and a cause that raises both findings whatever the coil's state would invalidate it; $\Pr(H_A, H_B \mid F) = 0.9 \times 0.8 = 0.72$ and $\Pr(H_A, H_B \mid \neg F) = 0.1 \times 0.2 = 0.02$; sequentially, the flag gives 0.5 and the inspection then 0.8; at once, $\Pr(F \mid H_A, H_B) = 0.072/(0.072 + 0.018) = 0.80$.)

Exercise 3.4 (Derivation)**.** A unit's future corrective work orders satisfy $K_\tau \mid \lambda \sim \text{Poisson}(\lambda\tau)$ with $\lambda \mid \mathbf{y} \sim \text{Gamma}(3, 10\,\text{yr})$ (shape and rate) and $\tau = 2\,\text{yr}$. Derive the expected count, the predictive variance and its two contributions, and state what happens to each when a much longer work-order log supporting a similar rate arrives. ($\mathbb{E}[K_\tau \mid \mathbf{y}] = 2 \times 0.3 = 0.6$; by equation (3.35), $\text{Var}(K_\tau \mid \mathbf{y}) = 0.6 + 2^2 \times 3/10^2 = 0.6 + 0.12 = 0.72$; with a and b growing in proportion, the rate term $\tau^2(a/b)/b$ tends to zero while the process term stays near 0.6; today the rate term's share is $2/12 = 0.167$.)

Exercise 3.5 (Application)**.** A unit runs through a normal year with probability 0.90 and a corrective-failure probability of 0.01, and through a hot summer, with the coil at its limit, with probability 0.10 and a corrective-failure probability of 0.15. Derive the unconditional probability of a corrective failure in a year, and the probability that the summer was hot given that a corrective failure occurred. ($\Pr(F) = 0.90 \times 0.01 + 0.10 \times 0.15 = 0.024$; $\Pr(\text{hot} \mid F) = 0.015/0.024 = 0.625$.)

Exercise 3.6 (Application)**.** (a) A hall air temperature has prior $\mathcal{N}(25, 1.5^2)$, and a condition test with $\sigma_T = 0.5\,\text{K}$ reads $y_T = 26°\text{C}$. For the state, the one number T, write the two residuals and their slopes j_ℓ, the information Λ, take the Newton step from the prior mean,

and compute σ^2_{post} and $\Pr(T > 27^\circ\text{C} \mid y_T)$. ($r_1 = (T-25)/1.5$ and $r_2 = (T-26)/0.5$, $j_1 = 0.6667$ and $j_2 = 2\,\text{K}^{-1}$, $\Lambda = 0.4444 + 4 = 4.4444\,\text{K}^{-2}$; at $T = 25$, $r_1 = 0$ and $r_2 = -2$, so $\mathcal{L}' = j_1 r_1 + j_2 r_2 = -4$, $\Delta T = 4/4.4444 = 0.9$ and $\hat{T} = 25.90^\circ\text{C}$; $\sigma^2_{\text{post}} = 1/4.4444 = 0.225\,\text{K}^2$, $\sigma_{\text{post}} = 0.4743\,\text{K}$; $1 - \Phi(1.10/0.4743) = 1 - \Phi(2.319) = 0.0102$.) (b) For the two-entry hall of equation (1.11), whose Jacobian at the built-to-design state (23°C, 0) has the rows (100, 500), (2, 0) and (0, 3.333), compute the three outer products $\mathbf{J}_\ell^{\mathsf{T}}\mathbf{J}_\ell$ and their sum, and state which row creates the off-diagonal entry and what the rank of its term is. (Physics $\begin{bmatrix} 10\,000 & 50\,000 \\ 50\,000 & 250\,000 \end{bmatrix}$, reading diag(4, 0), prior diag(0, 11.11); sum $\begin{bmatrix} 10\,004 & 50\,000 \\ 50\,000 & 250\,011 \end{bmatrix}$; only the physics row has entries in both columns, and its term has rank one, with the determinant $10^4 \times 2.5 \times 10^5 - (5 \times 10^4)^2 = 0$.)

Chapter 4

The Anatomy of Probabilistic Inference

4.1 From a plant graph to an acceptance probability

The equipment hall is being handed over, and the asset manager who takes it on has its commissioning record and nothing else: the hall is new, and this is window 0 of its life. The record holds one acceptance test, run with the cooling unit holding the supply air at 18°C, a test load of 30 kW, and one hall sensor reading 23.2°C; the record's pre-test estimate of the thermal resistance of the hall's air path; and the release the handover asks for, 60 kW, against the acceptance limit the record writes for full load, 28°C. Nothing has been run at full load. The pre-test estimate gives one quick answer: a resistance of 0.160 K/kW puts the hall at $18 + 60 \times 0.160 = 27.6$°C at full load, under the limit, so the hall can be released. The test gives a second: its rise of 5.2 K above the supply air, doubled with the load, puts the hall at 28.4°C, just above the limit, so the release is refused. These are the commissioning record's own entries: its pre-test estimate is neither Chapter 1's design sheet nor its handbook tile, and its 28°C acceptance limit is the commissioning criterion, not the 27.5°C the owner asks of the doubled load in Chapter 1.

The two quick answers disagree, and neither says how sure the asset manager can be: the pre-test estimate ignores the test, and the scaled test ignores the pre-test estimate and the error of the sensor. What neither settles is what one test, the pre-test estimate and the physics of the hall together say about a load that has never been run, and with what probability. The question of this chapter is

how a posterior distribution over the state of a plant graph is obtained from physical relations, readings and prior information, and how that posterior is turned into the probability that answers an asset manager's question.

Chapter 1 built the plant graph, Chapter 2 classified the questions and Chapter 3 derived the rules one at a time; here they are used together for the first time, on a state of three entries of which the reading touches only one.

Jargon: state estimation, maximum a posteriori (MAP) estimate, Hessian, Laplace approximation, delta method, change of variables, credible interval, prior and posterior predictive distributions, marginal likelihood, Bayes factor, rank-one update, reduced state, commissioning baseline, trend row, trend entry (degradation rate)

State estimation is the fit of the state of a plant to its readings and priors; its results are the MAP state and the posterior covariance defined next. The *maximum a posteriori (MAP) estimate* is the value of the unknowns at which the posterior density is highest; it is found by minimizing the negative logarithm of the posterior, the *objective*. The *Hessian* is the matrix of second derivatives of the objective; at the MAP estimate it measures the curvature of the objective, and for rows linear in the state and divided by their standard deviations it is $\mathbf{J}^\mathsf{T}\mathbf{J}$, the information matrix. The *Laplace approximation* replaces the posterior by the Gaussian whose mean is the MAP estimate and whose covariance is the inverse of the Hessian there. The *delta method* approximates the variance of a function of the unknowns by linearizing the function at the estimate. A *change of variables* carries a density from one coordinate to another, multiplying it by the derivative of the old coordinate with respect to the new so that each interval keeps its probability. A *credible interval* contains the unknown with a stated posterior probability. The *prior predictive* and *posterior predictive* distributions are the distributions of a reading not yet seen, averaged over the prior and over the posterior respectively. The *marginal likelihood*, or *evidence*, of a model is the probability density of the observed readings averaged over the prior; the *Bayes factor* is the ratio of the evidences of two models. A *rank-one update* adds one new reading to an existing Gaussian posterior by a closed-form correction of its mean and covariance, without solving the problem again. The *reduced state* is the vector of free entries that remain when the physical rows are taken as exact; every other entry of the state is then a function of it. The *commissioning baseline* is the fit of window 0, the commissioning test, against which every later condition test of the asset is read. A *trend row* ties a parameter's entry in one window to its entry in the next, and the *trend entry*, or *degradation rate*, is the entry of the state that carries the parameter's change per year.

Key idea

A condition or acceptance probability is computed in two operations. Inference, the state estimation of the plant, turns the plant graph, the readings and the priors into a posterior over the state: its mean is the MAP state and its covariance is the inverse curvature of the objective. Query evaluation turns that posterior into the asset manager's answer: an entry of the state and its variance, or a derived quantity $f(\mathbf{z})$ with variance $\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}$, and then a normal tail at the limit.

4.1.1 The complete inference workflow

The method fixed in Chapter 2 (Figure 2.7) is a progression: the plant graph (step 0), then the questions written over its state (step 1), the graph extended for them with the evidence attached (step 2), the posterior of the state that the rows give, derived and computed once (step 3), the answers read from that posterior (step 4), their checks (step 5), the report with its bands (step 6), and the next measurement when a band is too wide (step 7). Figure 4.1 draws the progression. The wide box at the top is the plant graph, built once; the seven boxes beneath are one pass that carries every question, read left to right along the top row and right to left along the bottom row; the second line of each box names what the step produces. The dashed arrow from step 7 back to step 3 is the loop taken when a new measurement is taken: its row is added and the posterior is rebuilt.

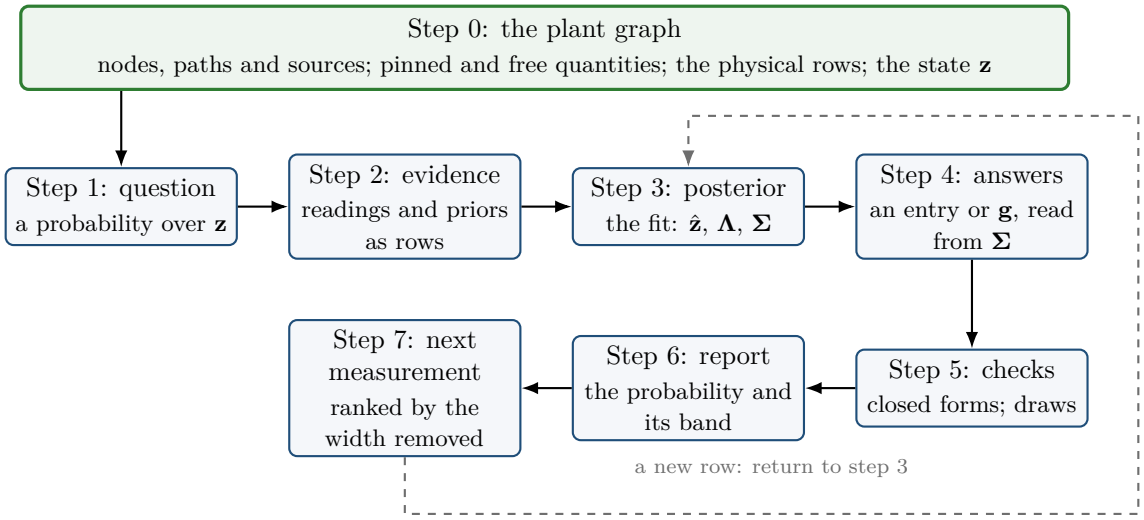


Figure 4.1: The inference workflow. Step 0 is built once; steps 1 to 7 are one pass, carrying every question of the plant. The second line of each box names what the step produces.

The distinction between steps 3 and 4 carries the method, and it is the distinction between the two operations of every inference. Step 3 is *inference*: the rows of step 2 are combined into one posterior of the state, its MAP state, information and covariance, derived and then computed for the plant; it does not depend on which question is asked. Step 4 is *query evaluation*: each question is reduced, from the asset manager's statement through its general and specialized forms, to quantities read from that posterior, in this chapter the posterior mean and variance of one state entry, the full-load temperature, and the standard normal distribution function evaluated at the limit. Keeping them apart is what lets one posterior serve every question, and a new measurement change step 3 alone.

4.1.2 The posterior and the questions put to it

The calculation divides into two operations of different kinds:

$$\underbrace{(\text{graph},\ \mathbf{y}) \longrightarrow p(\mathbf{z} \mid \mathbf{y})}_{\text{inference}}, \qquad \underbrace{p(\mathbf{z} \mid \mathbf{y}) \longrightarrow \text{condition or acceptance probability, or decision}}_{\text{query evaluation}}, \tag{4.1}$$

the second applied to the result of the first. Inference is done once for a given graph and a given body of evidence. Under a Gaussian posterior, exact or approximate, its result is the triple of the MAP state $\hat{\mathbf{z}}$, the posterior covariance $\boldsymbol{\Sigma}$, and the Jacobian of the model at $\hat{\mathbf{z}}$, from which the gradient $\mathbf{g}$ of any derived quantity is formed. Query evaluation is then a formula on the triple. An output that is itself an entry z_i of the state, such as a temperature at a node, has mean $\hat{z}_i$ and variance Σ_{ii}, read straight off the posterior. An output that is a function $f(\mathbf{z})$ of several entries needs its gradient and the quadratic form $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$. A probability that the output stays below a limit is a normal tail on either pair.

Two kinds of further work follow from this division and must be kept apart. A new question asked of the same evidence needs no new inference: a new function of the state is pushed through the existing $\boldsymbol{\Sigma}$. New evidence changes the posterior itself: either the inference is repeated with the new row, or the existing Gaussian is conditioned on the new row by a rank-one update, which Section 4.6.4 derives. The posterior represents what is known about the plant under the model; the queries represent what the asset manager wants to determine.

4.1.3 The inference in the terms of graphical models

The operations of this book are those of inference in a probabilistic graphical model, under the names an asset manager uses. The plant graph with its rows is a factor graph (Section 1.1.3): each row is a factor over the entries it touches, and the posterior $e^{-\mathcal{L}(\mathbf{z})}$ is the normalized product of the factors. Table 4.1 sets each operation beside its name in the literature on graphical models. Its first column is the operation as this book writes it, the second the standard name, and the third the chapters that derive it.

Table 4.1: The operations of the book and their names in the literature on probabilistic graphical models.

Operation in this book	Name in graphical models	Derived in
a row: prior, physics, reading or trend	a factor of a factor graph	Chapters 1, 3; the trend row, this chapter
the posterior $e^{-\mathcal{L}(\mathbf{z})}$	the normalized product of the factors	Chapter 3
the normal equations, $\mathbf{\Lambda} = \mathbf{J}^{\mathsf{T}}\mathbf{J}$, $\mathbf{\Sigma} = \mathbf{\Lambda}^{-1}$	exact inference in a linear-Gaussian factor graph, in information form	Chapter 3; this chapter
Gauss–Newton and the Laplace covariance	the Gaussian approximation at the mode of a nonlinear factor graph (factor-graph optimization)	this chapter; Chapter 5
the Schur complement; the marginal of a joint Gaussian	variable elimination of the entries not asked about	Chapter 3
the enumeration of the configurations of discrete entries; total probability	exact inference in a discrete Bayesian network by summation	Chapter 3
the conjugate updates of a probability or a rate	exact inference with conjugate factors	Chapter 3; Chapter 15, in Volume II
the evidence $p(\mathbf{y} \mid M)$	the normalizing constant (partition function) of the product	Chapter 2
the reweighted draws	importance sampling, with the Laplace normal as the proposal	Chapter 5, as the full-posterior route; Chapter 8, in the coordinates of the discharge lines, where a ridge curved in the state is straight
the linear entries eliminated exactly at fixed spreads, the spreads integrated on a grid, and composition draws of the mixture	exact marginalization of the conditionally linear-Gaussian entries, quadrature over the rest, and ancestral sampling	Chapter 9

The general algorithms of that literature are not needed for the plants of this book, and the table shows why: every posterior here is solved jointly, by factoring the information matrix of the whole state, which is exact for linear-Gaussian rows and costs little at a few dozen entries; the one exception, the fleet of Chapter 9, factors it at each node of a grid over the two entries its rows are not linear in. Message passing, belief propagation (the sum-product algorithm) and the junction tree compute the same marginals by local operations on the graph, and they earn their place on graphs too large to factor jointly or with discrete entries of many levels; on a Gaussian graph without loops, belief propagation returns exactly the Schur complements of the joint factorization. Factoring a large sparse information matrix is itself variable elimination, and its cost is set by the order in which the entries are eliminated: eliminated in time order, a

state indexed by time gives the Kalman filter and smoother. The record of 36 monthly windows that Chapter 10 fits is such a state, the state of the trend rows of Section 4.6.5. Markov chain Monte Carlo (Gibbs sampling and the Metropolis algorithm) draws from a posterior that the Laplace normal cannot approximate and reweighting cannot correct: one with several modes, or the funnel of a hierarchical model, in which the spread of a population of rates is itself unknown; the fleet of fifty coils of Chapter 9 has such a funnel and needs no chain on its record: its linear entries are eliminated exactly and its two spreads integrated on a grid. Variational inference is not used in this book.

These algorithms differ in what they give up. Variable elimination, the junction tree, belief propagation on a graph without loops and the Kalman smoother on linear-Gaussian rows are exact: they compute the posterior the joint factorization computes, in another order. The Laplace approximation, loopy belief propagation, the linearized Kalman filters and variational inference are approximate, and their error is a bias that more computation does not remove; loopy belief propagation and variational inference in particular tend to report variances too small, which for a tail probability against a limit is the error that matters most. Importance sampling and Markov chain Monte Carlo are exact in the limit of many draws, and their error is random and can be estimated, although a chain that never visits a second mode can look converged. The two kinds of error check each other, which is why the draws serve as the other route against the Laplace answer in step 5. The route also decides which objects come cheaply. The joint factorization gives $\hat{\mathbf{z}}$, $\boldsymbol{\Sigma}$, the gradient and the preposterior covariance $\boldsymbol{\Sigma}_a$ of a rank-one update in one pass, and the evidence in closed form under the Laplace approximation; draws give the posterior as a sample, from which a variance is an estimate, a preposterior needs the candidate's outcomes simulated and the posterior refitted or reweighted, and the evidence needs a further computation, since the draws do not carry the normalizing constant. The method is unchanged by the choice: the plant graph, the questions, the rows and the objective $\mathcal{L}(\mathbf{z})$ are the same for every algorithm, the general forms of Chapter 2 are expectations over the posterior whatever computes it, and only the computations of steps 3 and 4, and the checks of step 5 that test them, depend on the route.

4.1.4 The plant graph under two operating conditions (step 0)

The plant is an equipment hall with a heat-generating load and a single effective heat-removal path to the supply air, which the cooling unit's discharge controller holds at $T_c = 18°\mathrm{C}$. Two operating conditions are considered: a test condition with a heat load of $Q_1 = 30\,\mathrm{kW}$, and the full load of the release, $Q_2 = 60\,\mathrm{kW}$. At steady state the heat supplied to the hall leaves it through the path, and energy conservation at the hall node reads

$$Q = G\,(T - T_c), \tag{4.2}$$

where T is the hall air temperature in °C, Q the heat load in kW, and G the effective thermal conductance of the air path in kW/K. The supply air is pinned: the controller holds it at its set point, which is what makes T_c a boundary of the graph rather than an unknown, and it fails if the unit runs out of capacity and the discharge drifts up, which Section 4.6.7 takes up. The balance is taken at steady state through one effective path: this reduces the cooling of the hall to a single unknown, and it fails in the minutes after a load change, while heat stored in the hall and its equipment is still moving, which Section 4.6.7 and Exercise 4.6 take up. The two loads are pinned too, set by the test and by the release: this makes them known inputs of the rows, and it fails if the load is only known through a meter of finite accuracy, in which case it becomes an entry of the state. It is convenient to work with the thermal resistance of the path,

$$\rho = \frac{1}{G}, \tag{4.3}$$

in K/kW, with which equation (4.2) becomes

$$T = T_c + Q\rho. \tag{4.4}$$

Carrying the resistance in the state, rather than the conductance, is a choice: the temperature is then a linear function of the state, which makes every row of the example linear, the posterior exactly Gaussian and every result below exact rather than approximate, and it fails only where a Gaussian on the resistance puts real probability below zero, which Section 4.1.7 measures for the pre-test prior and Section 4.6.7 discusses. Chapter 1 carries the conductance in the state as the log-conductance $g = \ln(G/G_{\text{ref}})$, which keeps it positive but makes the temperature nonlinear in the state, and Section 4.6.5 returns to it.

At the two operating conditions equation (4.4) gives

$$T_1 = 18 + 30\rho, \tag{4.5}$$

$$T_2 = 18 + 60\rho, \tag{4.6}$$

with T_1 the hall temperature during the test and T_2 the temperature at full load, both in °C with ρ in K/kW. Figure 4.2 draws the graph. Each row is one operating condition: the box on the left is the heat source, the green box in the middle the hall node carrying a free temperature, and the gray box on the right the supply-air node, pinned at 18°C. The solid arrows are the physical paths, the heat entering the hall and the heat leaving it through the resistance. The green node between the rows is the resistance ρ, free and shared by both paths, as the dashed links show. The dashed arrows leave the graph for the evidence: the reading of the test temperature and the pre-test prior on ρ, attached in Section 4.1.7. The full-load node carries no reading; it is the scenario node of a what-if (contingency) analysis that the question of Section 4.1.6 requires, drawn here with the rest of the graph so that the figure shows the graph as extended for that question.

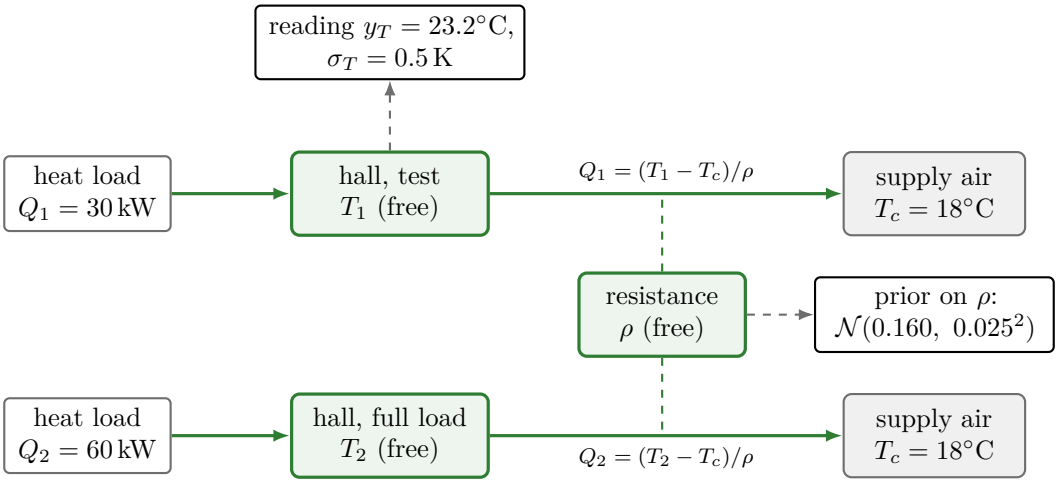


Figure 4.2: The plant graph of the equipment hall under its two operating conditions. Free: the resistance ρ and the temperatures T_1 and T_2 (green). Pinned: the supply-air temperature T_c, held by the cooling unit's discharge controller, and the loads Q_1, Q_2 (gray and white). Observed: T_1, by one reading (dashed arrow, top); ρ carries the pre-test prior (dashed arrow, right). The full-load node T_2 is the scenario node of the question and has no reading; it is linked to the test only through the shared resistance.

The resistance is taken to be the same at 60 kW as at 30 kW, one free entry shared by both rows: this is what lets a reading taken at the test condition carry information about the full-load condition, although the hall has never been operated at full load, and it fails if the

path changes with load, for example when the unit's fans or its coil change state, which one test reading cannot detect (Section 4.6.7) and a reading at full load would test (Section 4.6.3). The link between the two rows of Figure 4.2 is not a device of the calculation; it is the physical statement that both conditions are the same hall with the same heat-removal path, and the prediction at full load rests on it.

4.1.5 The state vector and the physical rows

The free quantities of the graph form the state vector,

$$\mathbf{z} = \begin{bmatrix} \rho \\ T_1 \\ T_2 \end{bmatrix}, \tag{4.7}$$

with entries in K/kW, °C and °C. The physical relations are written as rows that vanish when they hold:

$$r_1(\mathbf{z}) = T_1 - 18 - 30\rho = 0, \tag{4.8}$$

$$r_2(\mathbf{z}) = T_2 - 18 - 60\rho = 0. \tag{4.9}$$

These are the conservation rows of the graph, one for each hall node. The balance is taken to hold at both conditions up to a small model tolerance σ_p, a standard deviation that represents model discrepancy, and the answers are reported in the limit $\sigma_p \to 0$ of exact balances: the finite tolerance keeps every matrix of the calculation invertible while the reported numbers are those of an exact balance, and the limit understates the spread if the real discrepancy is finite, which Section 4.4.5 measures over three tolerances and Section 4.6.7 discusses. Section 4.2.3 puts the rows, with their tolerance, into one residual vector with the reading and the prior.

The three entries of the state are therefore not three independent unknowns. The two physical rows leave one direction of the state free, a change of ρ with both temperatures following it; in the limit of exact rows the state has three entries but one degree of freedom, and this fact governs the shape of the posterior covariance derived in Section 4.2.5 and the reduced state of Section 4.2.8.

4.1.6 The asset manager's questions (step 1)

The release raises the heat load from 30 kW to 60 kW. The maximum acceptable hall temperature, the acceptance limit of the commissioning record, is $T_{\max} = 28$°C. In the asset manager's words the question is: given the commissioning test and what is known of the air path, what is the probability that the hall stays below 28°C at full load? As a probability over the state, with the unknown on the left of the bar and everything treated as known on the right,

$$\Pr(T_2 < T_{\max} \mid \mathbf{y}, M), \tag{4.10}$$

where T_2 is the full-load entry of the state, $\mathbf{y}$ the evidence (the test reading and the pre-test prior), and M the assumed steady-state thermal model, equation (4.4) with the rows (4.8)–(4.9). The model is suppressed from the notation below when it is not in question. This is a question of the category *prediction under a threshold*: a tail probability of one marginal of the posterior.

Three estimation questions are asked of the same evidence, one for each entry of the state: the thermal resistance of the air path, the hall temperature at the test condition, and the hall temperature at full load. Table 4.2 lists the four questions: the first column is the question in the asset manager's words; the second the general form that every question of its category takes, written for a quantity $f(\mathbf{z})$ derived from the state and a threshold f^* on it; the third its mathematical form for the hall; the fourth its category; and the last the subsection where it is

answered. Which further measurement would narrow the answers most asks nothing new of the present evidence: it is the next measurement for these four questions, answered in Section 4.6. How fast the air path will lose conductance is not asked of this record, which has one window; it is the question a later window raises, and Section 4.6.5 asks and answers it there.

Table 4.2: The questions asked of the hall's commissioning test, the general form of their category, their mathematical form for the hall, their category and where each is answered. $\mathbf{y}$ is the test reading and the pre-test prior; $T_{\max} = 28°\mathrm{C}$.

Asset manager's question	General form	Mathematical form	Category	Answered in
What is the thermal resistance of the hall's air path at commissioning?	$\Pr(f(\mathbf{z}) \le f^* \mid \mathbf{y})$, moments	$p(\rho \mid \mathbf{y})$: $\hat{z}_1$, $\sqrt{\Sigma_{11}}$	estimation	4.3.6
What was the hall temperature during the commissioning test?	$\Pr(f(\mathbf{z}) \le f^* \mid \mathbf{y})$, moments	$p(T_1 \mid \mathbf{y})$: $\hat{z}_2$, $\sqrt{\Sigma_{22}}$	estimation	4.3.6
What will the hall temperature be at full load?	$\Pr(f(\mathbf{z}) \le f^* \mid \mathbf{y})$, moments, quantiles	$p(T_2 \mid \mathbf{y})$: $\hat{z}_3$, $\sqrt{\Sigma_{33}}$, credible interval	estimation	4.3.6
Can the hall be released for 60 kW within the 28°C acceptance limit?	$\Pr(f(\mathbf{z}) \ge f^* \mid \mathbf{y})$	$\Pr(T_2 < T_{\max} \mid \mathbf{y})$, with $f = -T_2$	prediction under a threshold	4.3.7

Because the full-load temperature is carried as an entry of the state, the threshold question needs exactly one marginal of the posterior: the mean $\hat{z}_3$ and the variance Σ_{33} of the third entry: the question is written over the state, and the position of T_2 in the state shows which part of the posterior it needs. Each estimation question decomposes in the same way, into one entry of $\hat{\mathbf{z}}$ and one diagonal entry of $\boldsymbol{\Sigma}$.

4.1.7 The evidence as rows (step 2)

The temperature reading. During the commissioning test at 30 kW the hall sensor reports $y_T = 23.2°\mathrm{C}$, with a measurement error of standard deviation $\sigma_T = 0.5\,\mathrm{K}$. The observation model writes the random reading Y as the true test temperature, the second entry of the state, plus an error,

$$Y = T_1 + \epsilon = \mathbf{h}^{\mathsf{T}}\mathbf{z} + \epsilon, \qquad \mathbf{h} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \qquad \epsilon \sim \mathcal{N}(0, \sigma_T^2), \tag{4.11}$$

where $\mathbf{h}$ selects the entry the sensor measures. The error is taken as Gaussian, unbiased, of known spread and independent of the state: this makes the reading one linear row of fixed weight $1/\sigma_T^2$, and it fails if the sensor is biased, which would shift every inferred quantity; the prior predictive check of Section 4.4.2 would flag a large bias, and the comparison of models of Section 4.2.2 is the test built for it. The reading lands on T_1 only; it is the physical row of the test hall, not a substitution, that will carry it to ρ.

Prior information. The commissioning record also carries its own pre-test estimate of the path's resistance, $0.160 \pm 0.025\,\mathrm{K/kW}$ ($6.25\,\mathrm{kW/K}$); it is neither the design sheet's $6\,\mathrm{kW/K}$ nor the handbook's $7.5\,\mathrm{kW/K}$ of Chapter 1. It is represented by the Gaussian pre-test prior

$$\rho \sim \mathcal{N}(\mu_0, \sigma_0^2), \qquad \mu_0 = 0.160\,\mathrm{K/kW}, \quad \sigma_0 = 0.025\,\mathrm{K/kW}. \tag{4.12}$$

The prior mean is the pre-test conductance $1/\mu_0 = 1/0.160 = 6.25\,\mathrm{kW/K}$. Taking the pre-test value as the prior for this hall is a choice: it is what places the full-load temperature at 27.6°C before any reading and what keeps the estimate from following the one reading alone, and it misleads if the estimate was made for air paths unlike this one, a conflict that the prior predictive check of Section 4.4.2 would show as a reading far from 22.8°C. A Gaussian prior on a resistance assigns some probability to negative, nonphysical values; here it is $\Phi(-0.160/0.025) = \Phi(-6.4) \approx 8 \times 10^{-11}$, negligible, and the Gaussian is kept for its analytical simplicity. A production model enforces positivity through a logarithmic parameterization, such as the log-conductance g of Chapter 1, or a prior with positive support.

The rows and the entries they act on. The physics and the evidence put four rows on the state: the physical rows of the test hall and of the full-load hall, equations (4.8)–(4.9), the reading row on T_1, and the prior row on ρ. Which row acts on which entry is a property of the graph, and it fixes where the Jacobian of the rows can be nonzero, marked $*$:

$$\mathbf{J} = \begin{bmatrix} * & * & 0 \\ * & 0 & * \\ 0 & * & 0 \\ * & 0 & 0 \end{bmatrix} \begin{matrix} \text{physics, test hall} \\ \text{physics, full-load hall} \\ \text{reading} \\ \text{prior} \end{matrix}, \qquad \text{columns } \rho, T_1, T_2. \tag{4.13}$$

A line of the pattern lists the entries one row sees, and a column lists the rows that see one entry. The third column has a single mark: the full-load temperature is seen by its own physical row and by nothing else. The second column shows that the reading and the test hall's balance are the two rows that see T_1, and the first that the two balances and the prior see ρ. No row joins the reading to T_2; the two physical rows, through the column ρ they share, are the only path from the one to the other.

Knowns and unknowns. Table 4.3 lists every quantity of the problem. The first column is the symbol, the second its value, and the third its role in the graph: pinned boundary condition, evidence, or free entry of the state. The graph and the probabilistic model are now specified completely enough for the inference to be carried out.

Table 4.3: Knowns and unknowns of the equipment-hall example.

Quantity	Value	Role
T_c [°C]	18	pinned: supply air, held by the discharge controller
Q_1 [kW]	30	pinned: test load
Q_2 [kW]	60	pinned: load of the release
y_T [°C]	23.2	evidence: the test reading
σ_T [K]	0.5	evidence: standard deviation of the reading error
μ_0 [K/kW]	0.160	evidence: prior mean of the resistance
σ_0 [K/kW]	0.025	evidence: prior standard deviation of the resistance
σ_p [K]	0.01, then $\to 0$	model: tolerance of the two physical rows
ρ [K/kW]	unknown	free: thermal resistance, the first entry of the state
T_1, T_2 [°C]	unknown	free: test and full-load temperatures, tied to ρ by the physical rows

Remark 4.1. Four spreads appear in what follows and must not be confused. The reading error has standard deviation σ_T (no index), a property of the sensor. The prior on the resistance has standard deviation σ_0. The physical rows have the model tolerance σ_p, a property of the

steady-state model. The posterior standard deviation of the i-th entry of the state is $\sqrt{\Sigma_{ii}}$, the square root of a diagonal entry of the posterior covariance, written $\mathrm{sd}(T_2 \mid \mathbf{y})$ when the entry is named.

4.2 The posterior of the state (step 3)

Every question of Table 4.2 is a property of one density, the posterior of the state given the reading and the pre-test prior. Each row of Section 4.1.7 contributes one factor to it, and the MAP state, the information matrix and the covariance are properties of that product. This section derives the posterior and computes it for the hall; Section 4.3 reads every answer from it.

4.2.1 The likelihood of the reading and the physical rows

The likelihood states how probable the observed reading would be under a proposed state. By equation (4.11), for a fixed state the reading is the test temperature T_1 plus a Gaussian error, so $Y \mid \mathbf{z} \sim \mathcal{N}(T_1,\ \sigma_T^2)$, and its density at the observed value is

$$p(y_T \mid \mathbf{z}) = \frac{1}{\sqrt{2\pi\sigma_T^2}} \exp\left[-\frac{(T_1 - y_T)^2}{2\sigma_T^2}\right]. \tag{4.14}$$

For the reading $y_T = 23.2\,^{\circ}\mathrm{C}$ and $\sigma_T = 0.5\,\mathrm{K}$, so that $2\sigma_T^2 = 0.5\,\mathrm{K}^2$,

$$p(y_T \mid \mathbf{z}) \propto \exp\left[-\frac{(T_1 - 23.2)^2}{0.5}\right]. \tag{4.15}$$

The likelihood is a function of the state with the reading held fixed; it is not a density in $\mathbf{z}$ and need not integrate to one over it. It depends on the first temperature only and is flat in ρ and T_2: the sensor measures a temperature and says nothing directly about the resistance or about the full-load hall.

What connects the three entries is the pair of physical rows. A steady-state balance that holds up to a model tolerance σ_p (a standard deviation in kelvin, the amount by which the balance of equation (4.4) is allowed to fail) is the density

$$p(T_1 \mid \rho) = \mathcal{N}(T_1;\ 18 + 30\rho,\ \sigma_p^2), \qquad p(T_2 \mid \rho) = \mathcal{N}(T_2;\ 18 + 60\rho,\ \sigma_p^2), \tag{4.16}$$

where $\mathcal{N}(x; \mu, \sigma^2)$ denotes the Gaussian density with mean μ and variance σ^2 evaluated at x. The exact rows (4.8)–(4.9) are the limit $\sigma_p \to 0$, in which each density concentrates on its row. The derivation below keeps σ_p finite, carries all three entries, and takes the limit at the end, where it is checked numerically; no row is used to eliminate an entry.

4.2.2 The prior density and Bayes' theorem

By equation (4.12) the prior density of the resistance is

$$p(\rho) = \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left[-\frac{(\rho - \mu_0)^2}{2\sigma_0^2}\right] \propto \exp\left[-\frac{(\rho - 0.160)^2}{2(0.025)^2}\right]. \tag{4.17}$$

The prior of the state is the product of this density with the two physical densities (4.16), $p(\mathbf{z}) = p(T_1 \mid \rho)\, p(T_2 \mid \rho)\, p(\rho)$, the factorization of Chapter 3 read along the graph: each hall

temperature depends on the resistance, and on nothing else. Bayes' theorem, derived in that chapter, gives the posterior of the state as

$$p(\mathbf{z} \mid y_T) = \frac{p(y_T \mid \mathbf{z})\, p(\mathbf{z})}{p(y_T)}, \qquad p(y_T) = \int p(y_T \mid \mathbf{u})\, p(\mathbf{u})\, \mathrm{d}\mathbf{u},$$

with $\mathbf{u}$ a variable of integration over the state. The denominator does not depend on $\mathbf{z}$, so

$$p(\mathbf{z} \mid y_T) \propto p(y_T \mid \mathbf{z})\, p(\mathbf{z}).$$

The evidence of the model. The denominator has a name of its own. Writing the model M explicitly, it is the evidence, or marginal likelihood, of the model,

$$p(y_T \mid M) = \int p(y_T \mid \mathbf{z}, M)\, p(\mathbf{z} \mid M)\, \mathrm{d}\mathbf{z}\,: \tag{4.18}$$

the density of the observed reading averaged over the prior of the state. For the linear-Gaussian model it is the prior predictive density of the reading evaluated at the observed value (Section 4.4.2). It also follows from the objective and its curvature, which is how it is computed when no closed form exists. The integrand is $\exp[-\mathcal{L}(\mathbf{z})]$, with the objective $\mathcal{L}$ half the sum of the four squared rows, each divided by its standard deviation (Section 4.2.3), times the normalizing constants $(2\pi\sigma_k^2)^{-1/2}$ of the four rows; when the objective is quadratic about its minimum $\hat{\mathbf{z}}$ with matrix $\mathbf{J}^{\mathsf{T}}\mathbf{J}$, whose inverse is the covariance $\boldsymbol{\Sigma}$ of Section 4.2.5, the Gaussian integral over the three entries gives

$$p(y_T \mid M) = \frac{(2\pi)^{3/2}\, \det(\boldsymbol{\Sigma})^{1/2}}{(2\pi)^2\, \sigma_p^2\, \sigma_T\, \sigma_0}\, \mathrm{e}^{-\mathcal{L}(\hat{\mathbf{z}})}. \tag{4.19}$$

The two directions of the state that the physical rows fix have posterior widths proportional to σ_p, and they cancel the two physical rows' normalizing constants: $\det(\boldsymbol{\Sigma})^{1/2}/\sigma_p^2$ tends to the width of the posterior along the free direction. The result is the volume of the posterior, $\det(\boldsymbol{\Sigma})^{1/2}$, times the height of the integrand at the MAP state. The formula is exact when the objective is quadratic, as it is here; in general it is the Laplace approximation to the evidence.

The evidence normalizes the posterior, and its second use is to compare models. Alternatives for the hall include M_{one}, a single thermal resistance; M_{two}, two distinct heat-removal paths; and M_{bias}, a model with a biased hall sensor, the drawing Chapter 7 sets against a loss of conductance. Each gives its own evidence for the same reading, and their ratio is a Bayes factor,

$$B_{\text{two,one}} = \frac{p(y_T \mid M_{\text{two}})}{p(y_T \mid M_{\text{one}})}, \tag{4.20}$$

which multiplies the prior odds of the two models to give their posterior odds, $\Pr(M_{\text{two}} \mid y_T)/\Pr(M_{\text{one}} \mid y_T) = B_{\text{two,one}} \cdot \Pr(M_{\text{two}})/\Pr(M_{\text{one}})$. That comparison is the category *hypothesis comparison*, and Chapter 7 develops it.

The posterior up to its constant. The data $\mathbf{y}$, the evidence in the sense of what is known about the hall, consists of the reading and the pre-test prior; the likelihood involves only the reading, and the prior enters through $p(\mathbf{z})$, so $p(\mathbf{z} \mid \mathbf{y})$ and $p(\mathbf{z} \mid y_T)$ name the same density. Substituting equations (4.14), (4.16) and (4.17) and collecting the four exponents,

$$p(\mathbf{z} \mid \mathbf{y}) \propto \exp\left[-\frac{1}{2}\left(\frac{(T_1 - 18 - 30\rho)^2}{\sigma_p^2} + \frac{(T_2 - 18 - 60\rho)^2}{\sigma_p^2} + \frac{(T_1 - y_T)^2}{\sigma_T^2} + \frac{(\rho - \mu_0)^2}{\sigma_0^2}\right)\right]. \tag{4.21}$$

This is the posterior of the state up to its normalizing constant. It combines the physical relations, the reading, the reading's uncertainty and the prior knowledge in one probability model, one term per row. Every later result of the chapter is a property of this expression.

4.2.3 The residual vector, the Jacobian and the information matrix

The maximum a posteriori estimate is the state at which the posterior density is greatest,

$$\hat{\mathbf{z}} = \arg\max_{\mathbf{z}} \; p(\mathbf{z} \mid \mathbf{y}). \tag{4.22}$$

The logarithm is increasing, so maximizing the posterior is equivalent to minimizing its negative logarithm. Dropping the terms that do not depend on $\mathbf{z}$ (the normalizing constants of the four densities and $\ln p(y_T)$) leaves half the sum of the squared exponents of (4.21). Each exponent is a row divided by its standard deviation, and the four rows stack into one residual vector,

$$\mathbf{r}(\mathbf{z}) = \begin{bmatrix} (T_1 - 18 - 30\rho)/\sigma_p \\ (T_2 - 18 - 60\rho)/\sigma_p \\ (T_1 - y_T)/\sigma_T \\ (\rho - \mu_0)/\sigma_0 \end{bmatrix} \begin{matrix} \text{physics, test hall} \\ \text{physics, full-load hall} \\ \text{reading} \\ \text{prior} \end{matrix} \tag{4.23}$$

whose entries are dimensionless, so that the objective is

$$\mathcal{L}(\mathbf{z}) = \tfrac{1}{2}\,\mathbf{r}(\mathbf{z})^{\mathsf{T}}\mathbf{r}(\mathbf{z}), \qquad p(\mathbf{z} \mid \mathbf{y}) \propto \exp\bigl[-\mathcal{L}(\mathbf{z})\bigr]. \tag{4.24}$$

In the notation of the book it is $\mathcal{L} = \frac{1}{2}\sum_k w_k r_k^2$ over the unnormalized rows r_1, r_2 of equations (4.8)–(4.9), $T_1 - y_T$ and $\rho - \mu_0$, with weights $w_k = 1/\sigma_k^2$; the normalized rows carry their weights inside them. The first two terms penalize a state that unbalances a hall, in units of the model tolerance; the third, a state that disagrees with the reading, in units of the sensor spread; the fourth, a resistance far from its pre-test value, in units of the prior spread. The weight of a physical row has a statistical meaning: $1/\sigma_p^2$ is the inverse variance of a model discrepancy, the amount by which the steady-state balance may fail.

The Jacobian. The derivatives of the four rows with respect to the three entries (ρ, T_1, T_2) form the 4×3 Jacobian $\mathbf{J} = \partial\mathbf{r}/\partial\mathbf{z}$, one line per row and one column per entry:

$$\mathbf{J} = \begin{bmatrix} -30/\sigma_p & 1/\sigma_p & 0 \\ -60/\sigma_p & 0 & 1/\sigma_p \\ 0 & 1/\sigma_T & 0 \\ 1/\sigma_0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} -3000 & 100 & 0 \\ -6000 & 0 & 100 \\ 0 & 2 & 0 \\ 40 & 0 & 0 \end{bmatrix} \quad (\sigma_p = 0.01\,\mathrm{K}), \tag{4.25}$$

with the first column in kW/K and the other two in K^{-1}. The entry J_{ki} is the change in row k, counted in units of that row's standard deviation, per unit change of entry i. Its nonzero entries sit where the pattern (4.13) of the graph places them. The entries $2 = 1/0.5$ and $40 = 1/0.025$ are the normalizing factors of the reading and the prior. Every row is linear in the state, so $\mathbf{J}$ is the same at every linearization point. The tolerance $\sigma_p = 0.01\,\mathrm{K}$, one-fiftieth of the sensor's spread, is tight enough that the answers no longer move (the limit is checked in Section 4.4.5) and loose enough, within two decades of σ_T, to keep the normal equations conditioned well enough for double precision.

The information matrix. For rows normalized by their standard deviations the information matrix is $\mathbf{\Lambda} = \mathbf{J}^{\mathsf{T}}\mathbf{J}$, equation (3.28). Written with the unnormalized rows, whose Jacobian is $\tilde{\mathbf{J}}$, and the diagonal weight matrix $\mathbf{W} = \mathrm{diag}(1/\sigma_p^2,\ 1/\sigma_p^2,\ 1/\sigma_T^2,\ 1/\sigma_0^2)$, it is the same matrix in the form used throughout the book,

$$\mathbf{\Lambda} = \mathbf{J}^{\mathsf{T}}\mathbf{J} = \tilde{\mathbf{J}}^{\mathsf{T}}\mathbf{W}\tilde{\mathbf{J}} = \mathbf{J}_{\mathrm{d}}^{\mathsf{T}}\mathbf{W}_{\mathrm{d}}\mathbf{J}_{\mathrm{d}} + \mathbf{\Lambda}_{\mathrm{prior}}, \tag{4.26}$$

where the last form keeps the prior rows apart: $\mathbf{J}_{\mathrm{d}}$ and $\mathbf{W}_{\mathrm{d}}$ hold the physical and reading rows, and $\mathbf{\Lambda}_{\mathrm{prior}}$, here $\mathrm{diag}(1/\sigma_0^2, 0, 0)$, is the prior precision. For rows that are nonlinear in the state,

equation (4.26) is the Gauss–Newton approximation to the Hessian of $\mathcal{L}$: it drops the curvature term $\sum_k r_k\, \partial^2 r_k / \partial \mathbf{z}\, \partial \mathbf{z}^\mathsf{T}$, the second derivatives of the rows weighted by their residuals, which vanishes for linear rows. Writing $\mathbf{j}_k$ for the k-th line of $\mathbf{J}$, the product is a sum of one outer product per row, so the information of independent rows adds, as equation (3.29) states for priors and readings; with the numbers of equation (4.25),

$$\mathbf{J}^\mathsf{T}\mathbf{J} = \sum_{k=1}^{4} \mathbf{j}_k^\mathsf{T}\mathbf{j}_k = \begin{bmatrix} 45\,001\,600 & -300\,000 & -600\,000 \\ -300\,000 & 10\,004 & 0 \\ -600\,000 & 0 & 10\,000 \end{bmatrix}. \tag{4.27}$$

The entry in the corner is $3000^2 + 6000^2 + 40^2$: the two physical rows and the prior. The reading contributes only the $4 = 1/\sigma_T^2$ of the (T_1, T_1) entry. The reading and prior rows alone would give $\operatorname{diag}(1600, 4, 0)$: no information at all would reach T_2, and none would pass from the reading to ρ. It is the off-diagonal entries of the physical rows that carry the reading at T_1 to ρ and from ρ to T_2. The matrix does not contain the reading's value y_T, only its standard deviation: the information is fixed by the design of the measurement and by the prior, not by the value observed. Its eigenvalues are 1.155, 1.0003×10^4 and 4.5012×10^7, all positive, with a condition number of 3.9×10^7 that grows as $1/\sigma_p^2$.

4.2.4 The normal equations and the MAP state

The stationary condition. By the chain rule the gradient of the objective (4.24) is

$$\nabla \mathcal{L}(\mathbf{z}) = \mathbf{J}^\mathsf{T}\mathbf{r}(\mathbf{z}), \tag{4.28}$$

and the MAP state is the point where it vanishes, $\mathbf{J}^\mathsf{T}\mathbf{r}(\hat{\mathbf{z}}) = \mathbf{0}$: three equations, one per entry.

The Gauss–Newton step. The step that sets the gradient of the linearized rows to zero solves the normal equations of equation (3.27),

$$(\mathbf{J}^\mathsf{T}\mathbf{J})\, \Delta\mathbf{z} = -\mathbf{J}^\mathsf{T}\mathbf{r}(\mathbf{z}), \qquad \mathbf{z}^{(l+1)} = \mathbf{z}^{(l)} + \Delta\mathbf{z}, \tag{4.29}$$

the Gauss–Newton iteration. The rows of this example are linear, so the linearization is exact and one step lands on $\hat{\mathbf{z}}$ from any starting state.

The free direction. The two physical lines of $\mathbf{J}$ annihilate one direction of the state,

$$\mathbf{J}_{\text{phys}}\, \mathbf{v} = \mathbf{0}, \qquad \mathbf{J}_{\text{phys}} = \begin{bmatrix} -3000 & 100 & 0 \\ -6000 & 0 & 100 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} 1 \\ 30 \\ 60 \end{bmatrix}, \tag{4.30}$$

the free direction: a change of ρ with T_1 and T_2 following by $Q_1 = 30\,\text{kW}$ and $Q_2 = 60\,\text{kW}$ times that change, which unbalances neither hall. A pull of the reading on T_1 is answered by a move along $\mathbf{v}$, because any move off it would cost the physical rows far more than it gains on the reading; from a state that balances both halls, the step of the normal equations therefore lies, to within the tolerance, along $\mathbf{v}$.

4.2.5 The posterior covariance from the information matrix

The point estimate does not answer the acceptance question, which depends on how uncertain the state remains.

The covariance. The rows are linear and Gaussian, so by equation (3.28) the posterior is exactly $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ with

$$\boldsymbol{\Sigma} = (\mathbf{J}^{\mathsf{T}}\mathbf{J})^{-1},$$

the inverse of the matrix (4.27), with rows and columns ordered (ρ, T_1, T_2), so that its entries carry $(\mathrm{K/kW})^2$, K^2, and $\mathrm{K{\cdot}K/kW}$ for the products. A larger information means a narrower posterior: the covariance is the inverse of the steepness of the objective.

The marginals. The estimation questions ask for entries. The marginals of a Gaussian are read off its mean and covariance,

$$\rho \mid \mathbf{y} \sim \mathcal{N}(\hat{z}_1,\ \Sigma_{11}), \qquad T_1 \mid \mathbf{y} \sim \mathcal{N}(\hat{z}_2,\ \Sigma_{22}), \qquad T_2 \mid \mathbf{y} \sim \mathcal{N}(\hat{z}_3,\ \Sigma_{33}), \tag{4.31}$$

so the resistance is $\hat{\rho} = \hat{z}_1$ with standard deviation $\sqrt{\Sigma_{11}}$, the test temperature $\hat{z}_2$ with $\sqrt{\Sigma_{22}}$, and the full-load temperature, the entry the threshold question asks about, $\hat{T}_2 = \hat{z}_3$ with $\sqrt{\Sigma_{33}}$. How the entries move together is read off the correlations $\Sigma_{ij}/\sqrt{\Sigma_{ii}\Sigma_{jj}}$, which the physical rows drive toward one.

Coordinates along and across the physical rows. The shape of $\boldsymbol{\Sigma}$, and its limit as the tolerance shrinks, are read by measuring the state along the free direction and across the two physical rows. Write

$$\mathbf{z} = \mathbf{c} + \mathbf{B}\mathbf{u}, \qquad \mathbf{u} = \begin{bmatrix} \rho \\ d_1 \\ d_2 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 0 & 0 \\ 30 & 1 & 0 \\ 60 & 0 & 1 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} 0 \\ 18 \\ 18 \end{bmatrix}, \tag{4.32}$$

where $d_1 = T_1 - 18 - 30\rho$ and $d_2 = T_2 - 18 - 60\rho$ are the two hall imbalances in kelvin and the first column of $\mathbf{B}$ is $\mathbf{v}$. The rows are linear in $\mathbf{u}$ with Jacobian $\mathbf{JB}$, whose columns are $\mathbf{Jv}$, $\mathbf{J}\mathbf{e}_2$ and $\mathbf{J}\mathbf{e}_3$,

$$\mathbf{JB} = \begin{bmatrix} 0 & 1/\sigma_p & 0 \\ 0 & 0 & 1/\sigma_p \\ 30/\sigma_T & 1/\sigma_T & 0 \\ 1/\sigma_0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 100 & 0 \\ 0 & 0 & 100 \\ 60 & 2 & 0 \\ 40 & 0 & 0 \end{bmatrix}.$$

The first column, $\mathbf{Jv} = (0,\ 0,\ 60,\ 40)^{\mathsf{T}}$, is seen only by the reading (with slope $2 \times 30 = 60$) and the prior (slope 40); a move along $\mathbf{v}$ costs the information $\|\mathbf{Jv}\|^2 = 60^2 + 40^2 = 5200\,(\mathrm{kW/K})^2$. The other two columns carry the entries $1/\sigma_p$, so every direction that unbalances a hall costs about $1/\sigma_p^2 = 10^4\,\mathrm{K}^{-2}$: a tiny tolerance leaves the state free to move cheaply only along $\mathbf{v}$, a narrow ridge in the objective. The third column is orthogonal to the other two, so the information matrix in these coordinates is block diagonal,

$$(\mathbf{JB})^{\mathsf{T}}(\mathbf{JB}) = \begin{bmatrix} 1/\sigma_0^2 + 30^2/\sigma_T^2 & 30/\sigma_T^2 & 0 \\ 30/\sigma_T^2 & 1/\sigma_p^2 + 1/\sigma_T^2 & 0 \\ 0 & 0 & 1/\sigma_p^2 \end{bmatrix} = \begin{bmatrix} 5200 & 120 & 0 \\ 120 & 10\,004 & 0 \\ 0 & 0 & 10\,000 \end{bmatrix},$$

and its inverse, the covariance $\boldsymbol{\Sigma}_{\mathbf{u}}$ of $\mathbf{u}$, is

$$\boldsymbol{\Sigma}_{\mathbf{u}} = \begin{bmatrix} 1.9236 \times 10^{-4} & -2.3074 \times 10^{-6} & 0 \\ -2.3074 \times 10^{-6} & 9.9988 \times 10^{-5} & 0 \\ 0 & 0 & 1.0000 \times 10^{-4} \end{bmatrix} \quad (\sigma_p = 0.01\,\mathrm{K}).$$

Three results follow. First, the full-load imbalance d_2 is seen by its own row only: $(\boldsymbol{\Sigma}_{\mathbf{u}})_{33} = \sigma_p^2$, the variance that row allows, and it is uncorrelated with ρ and d_1. Second, by the block-inversion

formula of Section 3.5.3, the first diagonal entry of $\mathbf{\Sigma_u}$ is the inverse of the Schur complement of the (d_1, d_1) entry in the upper 2×2 block,

$$\Sigma_{11} = (\mathbf{\Sigma_u})_{11} = \left[\frac{1}{\sigma_0^2} + \frac{30^2}{\sigma_T^2} - \frac{(30/\sigma_T^2)^2}{1/\sigma_p^2 + 1/\sigma_T^2} \right]^{-1} = \left[\frac{1}{\sigma_0^2} + \frac{30^2}{\sigma_T^2 + \sigma_p^2} \right]^{-1}, \tag{4.33}$$

the first row of $\mathbf{B}$ being $\mathbf{e}_1^\mathsf{T}$; a tolerance acts on the reading as extra sensor variance and widens the spread of ρ. Third, the covariance of the state is $\mathbf{\Sigma} = \mathbf{B}\mathbf{\Sigma_u}\mathbf{B}^\mathsf{T}$; its third diagonal entry, with $T_2 = 18 + 60\rho + d_2$ and d_2 uncorrelated with ρ, is

$$\mathrm{Var}(T_2 \mid \mathbf{y}) = \Sigma_{33} = \mathbf{g}^\mathsf{T}\mathbf{\Sigma}\mathbf{g} + \sigma_p^2 = 60^2\,\Sigma_{11} + \sigma_p^2, \qquad \mathbf{g} = (60,\ 0,\ 0)^\mathsf{T}, \tag{4.34}$$

with $\mathbf{g}$ the gradient of the prediction $18 + 60\rho$. A model discrepancy therefore enters the answer twice: it widens Σ_{11}, and it adds its own variance at the node that has no reading.

The limit of exact rows. As $\sigma_p \to 0$, $(\mathbf{\Sigma_u})_{22}$ and $(\mathbf{\Sigma_u})_{33}$, both at most σ_p^2, vanish and $(\mathbf{\Sigma_u})_{11}$ tends to $1/5200$ by equation (4.33), so $\mathbf{\Sigma_u} \to \mathrm{diag}(1/5200,\ 0,\ 0)$ and

$$\mathbf{\Sigma} = \mathbf{B}\mathbf{\Sigma_u}\mathbf{B}^\mathsf{T} \ \longrightarrow\ \frac{\mathbf{B}\mathbf{e}_1\mathbf{e}_1^\mathsf{T}\mathbf{B}^\mathsf{T}}{5200} = \frac{\mathbf{v}\mathbf{v}^\mathsf{T}}{5200} = \frac{1}{5200}\begin{bmatrix} 1 & 30 & 60 \\ 30 & 900 & 1800 \\ 60 & 1800 & 3600 \end{bmatrix}, \tag{4.35}$$

of rank one. The matrix has two zero eigenvalues and no ordinary inverse, and the posterior correlation of T_1 and T_2 is exactly one. This is not a failure of the inference; it is the consequence of exact physical rows, which leave the state one degree of freedom. The state-vector form needs no special treatment of it: the normal equations are solved at a finite tolerance, where $\mathbf{J}^\mathsf{T}\mathbf{J}$ is invertible, and the limit is derived in the coordinates of equation (4.32) and read from the sequence of tolerances. The distinction between exact constraints and uncertain physical relations matters whenever a Hessian and a covariance are constructed over the full state, and the tolerance is the device that keeps both defined.

The objective along the free direction. On the line $\mathbf{z} = \hat{\mathbf{z}} + t\,\mathbf{v}$ through the MAP state, in the limit of exact rows, the physical rows keep their values at $\hat{\mathbf{z}}$ (zero in the limit) and only the reading and prior rows change, as $r_3(\hat{\mathbf{z}}) + 60t$ and $r_4(\hat{\mathbf{z}}) + 40t$, the slopes being the entries of $\mathbf{J}\mathbf{v} = (0,\ 0,\ 60,\ 40)^\mathsf{T}$. The objective on the line is the reading term $\frac{1}{2}(r_3(\hat{\mathbf{z}}) + 60t)^2$ plus the prior term $\frac{1}{2}(r_4(\hat{\mathbf{z}}) + 40t)^2$. At the minimum the slopes of the two terms cancel, $\mathbf{v}^\mathsf{T}\mathbf{J}^\mathsf{T}\mathbf{r}(\hat{\mathbf{z}}) = 60\,r_3 + 40\,r_4 = 0$: the reading pulls with weight 60 and the prior with weight 40. The gradient vanishes at $\hat{\mathbf{z}}$, so the objective on the line is

$$\mathcal{L}(\hat{\mathbf{z}}+t\mathbf{v}) = \mathcal{L}(\hat{\mathbf{z}}) + \tfrac{1}{2}\,t^2\,\|\mathbf{J}\mathbf{v}\|^2, \qquad \frac{\mathrm{d}^2}{\mathrm{d}t^2}\mathcal{L}(\hat{\mathbf{z}}+t\mathbf{v}) = \mathbf{v}^\mathsf{T}\mathbf{J}^\mathsf{T}\mathbf{J}\,\mathbf{v} = 60^2 + 40^2 = 5200\,(\mathrm{kW/K})^2, \tag{4.36}$$

a parabola whose curvature is the information along $\mathbf{v}$, equal to $1/\Sigma_{11}$ in the limit; the rise of $\frac{1}{2}$ is reached at $t = \pm 1/\sqrt{5200} = \pm\sqrt{\Sigma_{11}}$.

Remark 4.2. Substituting the exact rows $T_1 = 18 + 30\rho$ and $T_2 = 18 + 60\rho$ into the objective gives the same curve as a function of the resistance alone,

$$\mathcal{L}(\rho) = \frac{1}{2}\left[\frac{(\rho - \mu_0)^2}{\sigma_0^2} + \frac{(y_T - 18 - 30\rho)^2}{\sigma_T^2} \right], \tag{4.37}$$

whose second derivative is the 5200 of equation (4.36) and whose minimum is $\hat{\rho} = 880/5200$.

4.2.6 The Gaussian form of the posterior

In this example the prior is Gaussian, the reading error and the model tolerance are Gaussian, and every row is linear in the state. The negative log posterior is therefore exactly quadratic in $\mathbf{z}$, the curvature term $\sum_k r_k\,\partial^2 r_k/\partial\mathbf{z}\,\partial\mathbf{z}^\mathsf{T}$ is zero, the Gauss–Newton matrix $\mathbf{J}^\mathsf{T}\mathbf{J}$ is the exact Hessian, and its inverse is the exact posterior covariance under the stated model. Nothing in equations (4.29)–(4.31) is an approximation; the only limit taken, $\sigma_p \to 0$, is derived in equation (4.35).

For a nonlinear physical model, for instance one written in the log-conductance g, the negative log posterior is not quadratic. The same construction is then made locally: the objective is expanded to second order about the MAP state,

$$\mathcal{L}(\mathbf{z}) \approx \mathcal{L}(\hat{\mathbf{z}}) + \tfrac{1}{2}(\mathbf{z}-\hat{\mathbf{z}})^\mathsf{T}\mathbf{H}(\mathbf{z}-\hat{\mathbf{z}}), \qquad \mathbf{H} = \left.\frac{\partial^2\mathcal{L}}{\partial\mathbf{z}\,\partial\mathbf{z}^\mathsf{T}}\right|_{\hat{\mathbf{z}}} = \mathbf{J}^\mathsf{T}\mathbf{J} + \sum_k r_k \left.\frac{\partial^2 r_k}{\partial\mathbf{z}\,\partial\mathbf{z}^\mathsf{T}}\right|_{\hat{\mathbf{z}}},$$

with no first-order term because the gradient vanishes at the minimum, and the posterior is replaced by the Gaussian that this expansion defines, the Laplace approximation:

$$\mathbf{z} \mid \mathbf{y} \approx \mathcal{N}\big(\hat{\mathbf{z}},\ \mathbf{H}^{-1}\big). \tag{4.38}$$

It holds when the posterior is close to Gaussian over its own width and when $\mathbf{H}$ is positive definite in the chosen free coordinates. The distinction matters throughout the book: the calculation of this chapter is exact, while a Laplace approximation in general is not, and its adequacy is one of the checks of step 5.

4.2.7 Rows that are not Gaussian

Every row of the hall is Gaussian, and its term in the objective is half its squared residual. Chapter 2's log of corrective work orders already has a row of another kind, a count of work orders in an exposure of unit-years. The construction above carries over to it unchanged once the objective is defined by the posterior rather than by a sum of squares.

The objective. The objective is the negative logarithm of the posterior up to a constant. Each row, a likelihood factor or a prior factor, contributes one term, its negative log-probability:

$$\mathcal{L}(\mathbf{z}) = \sum_k \mathcal{L}_k(\mathbf{z}), \qquad \mathcal{L}_k(\mathbf{z}) = -\ln p_k(\cdot \mid \mathbf{z}), \qquad p(\mathbf{z} \mid \mathbf{y}) \propto \exp\big[-\mathcal{L}(\mathbf{z})\big]. \tag{4.39}$$

A Gaussian row is the case $\mathcal{L}_k = \frac{1}{2}r_k^2$, and the objective (4.24) of the hall is this sum for its four rows.

What the fit uses. The MAP state, the Newton step and the Laplace covariance use the objective only through its gradient and its Hessian, and both are sums over the rows:

$$\nabla\mathcal{L} = \sum_k \nabla\mathcal{L}_k, \qquad \mathbf{H} = \sum_k \nabla^2\mathcal{L}_k, \qquad \mathbf{H}\,\Delta\mathbf{z} = -\nabla\mathcal{L}, \qquad \mathbf{\Sigma} = \mathbf{H}(\hat{\mathbf{z}})^{-1}. \tag{4.40}$$

The value of $\mathcal{L}$ enters only the evidence, equation (4.19). When the state has one entry z, the gradient is the slope $\mathcal{L}'(z)$, the Hessian the curvature $\mathcal{L}''(z)$, the step $\Delta z = -\mathcal{L}'/\mathcal{L}''$ and the variance $1/\mathcal{L}''(\hat{z})$; nothing else changes.

The Gaussian row. For $\mathcal{L}_k = \frac{1}{2}r_k^2$ with Jacobian line $\mathbf{j}_k = \partial r_k/\partial\mathbf{z}^\mathsf{T}$, the chain rule gives

$$\nabla\mathcal{L}_k = \mathbf{j}_k^\mathsf{T} r_k, \qquad \nabla^2\mathcal{L}_k = \mathbf{j}_k^\mathsf{T}\mathbf{j}_k + r_k\,\frac{\partial^2 r_k}{\partial\mathbf{z}\,\partial\mathbf{z}^\mathsf{T}}. \tag{4.41}$$

Summed over the rows these are $\mathbf{J}^\mathsf{T}\mathbf{r}$ and $\mathbf{J}^\mathsf{T}\mathbf{J}$ plus the curvature term of Section 4.2.6.

The count row. A record of n_k events (the count of Chapter 2) whose expected number is $\mu_k(\mathbf{z})$, Poisson in its exposure, has the term $\mathcal{L}_k = \mu_k - n_k \ln \mu_k$. Its gradient and Hessian are

$$\nabla \mathcal{L}_k = (\mu_k - n_k)\, \nabla \ln \mu_k, \qquad \nabla^2 \mathcal{L}_k = \mu_k\, \nabla \ln \mu_k\, \nabla \ln \mu_k^{\mathsf{T}} + (\mu_k - n_k)\, \nabla^2 \ln \mu_k. \tag{4.42}$$

The first term of the Hessian is the product of one line with itself, so it is $\mathbf{j}_k^{\mathsf{T}} \mathbf{j}_k$ for the line and the residual

$$\mathbf{j}_k = \sqrt{\mu_k}\, \nabla \ln \mu_k^{\mathsf{T}}, \qquad r_k = \frac{\mu_k - n_k}{\sqrt{\mu_k}}, \tag{4.43}$$

which also give $\mathbf{j}_k^{\mathsf{T}} r_k = \nabla \mathcal{L}_k$. This pair, the *working residual* and the *working Jacobian line*, is how a count row joins the Jacobian beside Gaussian rows: one more line of $\mathbf{J}$, one more entry of $\mathbf{r}$, and the same products (4.40). The remaining term $(\mu_k - n_k)\, \nabla^2 \ln \mu_k$ plays the part of $r_k\, \partial^2 r_k$ in a Gaussian row. It vanishes when $\ln \mu_k$ is linear in the state, which is the case when the state carries the log-rate, the canonical parameter of the Poisson distribution.

The working pair reproduces the row's gradient and the Gauss–Newton part of its Hessian, and nothing more. Its half square is not the row's term, $\frac{1}{2} r_k^2 \neq \mathcal{L}_k$, and $\mathbf{j}_k$ is not the derivative of r_k except where $r_k = 0$; the fit never uses either, because it uses only the products (4.40). For a Gaussian row that is linear in the state, the working pair is the row's own residual and Jacobian line, which is why the two forms are the same in the hall.

The information and its curvature term. With every row in this form, the Hessian is

$$\mathbf{H} = \mathbf{J}^{\mathsf{T}} \mathbf{J} + \mathbf{C}, \qquad \mathbf{\Lambda} = \mathbf{J}^{\mathsf{T}} \mathbf{J}, \tag{4.44}$$

with $\mathbf{C}$ the sum of the rows' remaining terms. The matrix $\mathbf{\Lambda}$ is the information; it is the average of $\mathbf{H}$ over the records the model could have produced, the *expected information*. The two are equal when $\mathbf{C} = \mathbf{0}$: for linear Gaussian rows, and for count rows carried on their canonical parameter. Where $\mathbf{C}$ is not zero, the chapters name it, keep it in the Newton step and report the covariance from $\mathbf{H}$. Once $\hat{\mathbf{z}}$ and $\mathbf{\Sigma}$ are in hand, the kind of row no longer matters: every answer of step 4 is read from them in the same way.

4.2.8 The reduced state

The full state carries every free quantity of the graph, and in the limit of exact rows the physical rows make some of its entries functions of the others; in the hall, two of the three entries follow from the third. The same posterior can then be written over the entries that remain free, the reduced state, as a projection of the rows onto the null space of the physical rows; the two-unit hall of Chapter 5 and every plant after it are fitted in this form, and so are the windows of a condition record, so the derivation below is written for any plant, with the hall as its instance.

The null space of the physical rows. Let the full state $\mathbf{z}$ have n_z entries and the graph n_d physical rows, with Jacobian $\mathbf{G} = \mathbf{J}_\mathrm{phys}$, an $n_\mathrm{d} \times n_\mathrm{z}$ matrix. Split the full state into $n_\mathrm{f} = n_\mathrm{z} - n_\mathrm{d}$ free entries $\mathbf{z}_\mathrm{f}$ and n_d dependent entries $\mathbf{z}_\mathrm{d}$, and $\mathbf{G}$ into the matching columns, $\mathbf{G} = [\, \mathbf{G}_\mathrm{free} \;\; \mathbf{G}_\mathrm{dep} \,]$, an $n_\mathrm{d} \times n_\mathrm{f}$ and an $n_\mathrm{d} \times n_\mathrm{d}$ block (the entries ordered free first). The split is a choice: any split serves for which $\mathbf{G}_\mathrm{dep}$ is square and invertible, that is, for which the physical rows can be solved for the dependent entries at given free ones. A move $\Delta\mathbf{z}$ of the full state keeps the linearized balances when $\mathbf{G}\, \Delta\mathbf{z} = \mathbf{0}$, that is $\mathbf{G}_\mathrm{free}\, \Delta\mathbf{z}_\mathrm{f} + \mathbf{G}_\mathrm{dep}\, \Delta\mathbf{z}_\mathrm{d} = \mathbf{0}$, so

$$\Delta\mathbf{z}_\mathrm{d} = \mathbf{N}_\mathrm{b}\, \Delta\mathbf{z}_\mathrm{f}, \qquad \mathbf{N}_\mathrm{b} = -\mathbf{G}_\mathrm{dep}^{-1} \mathbf{G}_\mathrm{free}, \qquad \mathbf{N} = \begin{bmatrix} \mathbf{I}_{n_\mathrm{f}} \\ \mathbf{N}_\mathrm{b} \end{bmatrix}, \qquad \mathbf{G}\mathbf{N} = \mathbf{0}. \tag{4.45}$$

Each physical row's tolerance divides both of its blocks and cancels in $\mathbf{N}_\mathrm{b}$. The n_f columns of the $n_\mathrm{z} \times n_\mathrm{f}$ matrix $\mathbf{N}$ are independent, through its identity block, and $\mathbf{G}$ has rank n_d, so they

span the null space of the physical rows, the tangent of the set on which the balances hold; each column moves one free entry by one unit and the dependent entries as the linearized balances require. In the hall $n_z = 3$ and $n_d = 2$; with $\mathbf{z}_f = \rho$ and $\mathbf{z}_d = (T_1, T_2)^\mathsf{T}$,

$$\mathbf{G}_{\text{free}} = \begin{bmatrix} -Q_1/\sigma_p \\ -Q_2/\sigma_p \end{bmatrix}, \qquad \mathbf{G}_{\text{dep}} = \begin{bmatrix} 1/\sigma_p & 0 \\ 0 & 1/\sigma_p \end{bmatrix}, \qquad \mathbf{N}_b = \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix}, \qquad \mathbf{N} = \begin{bmatrix} 1 \\ 30 \\ 60 \end{bmatrix},$$

the entries of $\mathbf{N}_b$ the loads in kW: a unit change of resistance raises each hall temperature by its load. The single column of $\mathbf{N}$ is the free direction $\mathbf{v}$ of equation (4.30). Taking T_1 as the free entry would serve as well, since the block of (ρ, T_2) is invertible too; the resistance is chosen because the prior acts on it and each hall temperature follows from it by its row of $\mathbf{N}_b$.

The rows of the reduced state. With exact balances the full state lies on the set where the physical rows vanish, and the reduced state $\mathbf{z}_f$ names a point of that set; the point itself is $\mathbf{z}_{\text{full}}(\mathbf{z}_f) = (\mathbf{z}_f,\ \mathbf{z}_d(\mathbf{z}_f))$, with $\mathbf{z}_d(\mathbf{z}_f)$ the solution of the physical rows at the given free entries, and its derivative with respect to $\mathbf{z}_f$ is $\mathbf{N}$. Write $\mathbf{r}_{\text{rows}}$ for the evidence rows, the readings and the priors, as functions of the full state. The rows of the reduced state are the evidence rows at the point of the set,

$$\mathbf{r}(\mathbf{z}_f) = \mathbf{r}_{\text{rows}}\big(\mathbf{z}_{\text{full}}(\mathbf{z}_f)\big), \tag{4.46}$$

and the objective on the reduced state is the full objective (4.24) on the set, where its physical rows contribute nothing. In the hall, with $\mathbf{z}_{\text{full}}(\rho) = (\rho,\ T_c + Q_1\rho,\ T_c + Q_2\rho)^\mathsf{T}$, the rows are

$$\mathbf{r}(\rho) = \begin{bmatrix} (T_c + Q_1\rho - y_T)/\sigma_T \\ (\rho - \mu_0)/\sigma_0 \end{bmatrix} \begin{matrix} \text{reading} \\ \text{prior} \end{matrix},$$

and the objective is the curve (4.37), reached here as a projection and a limit rather than by a substitution made by hand.

The reduced Jacobian. A move $\Delta\mathbf{z}_f$ of the reduced state moves the point of the set by $\mathbf{N}\,\Delta\mathbf{z}_f$, so the chain rule gives the Jacobian of the reduced rows as the Jacobian of the evidence rows projected onto the null space,

$$\mathbf{J}_{\text{red}} = \mathbf{J}_{\text{rows}}\mathbf{N}, \tag{4.47}$$

where $\mathbf{J}_{\text{rows}} = \partial\mathbf{r}_{\text{rows}}/\partial\mathbf{z}$ is the Jacobian of the evidence rows (readings and priors) with respect to the full state, the lines of $\mathbf{J}$ that are not physical rows. A reading whose row touches only dependent entries sees the free entries through $\mathbf{N}_b$. In the hall $\mathbf{J}_{\text{rows}}$ is the reading and prior lines of equation (4.25),

$$\mathbf{J}_{\text{rows}} = \begin{bmatrix} 0 & 1/\sigma_T & 0 \\ 1/\sigma_0 & 0 & 0 \end{bmatrix}, \qquad \mathbf{J}_{\text{red}} = \begin{bmatrix} Q_1/\sigma_T \\ 1/\sigma_0 \end{bmatrix}.$$

The reading, whose row in the full state touches only T_1, sees the resistance on the reduced state with the slope Q_1/σ_T: its entry is nonzero, so the reading alone determines the one free entry (the hall is observable from its one sensor). The information of the reduced state is $\mathbf{J}_{\text{red}}^\mathsf{T}\mathbf{J}_{\text{red}} = Q_1^2/\sigma_T^2 + 1/\sigma_0^2$, the reading's term and the prior's, equal to $\|\mathbf{J}\mathbf{v}\|^2$ of equation (4.36).

Nonlinear physical rows. When the physical rows are nonlinear in the state, as they are when a conductance is carried as a log-derate or a recirculation fraction multiplies a temperature difference, the same objects are evaluated point by point. $\mathbf{z}_{\text{full}}(\mathbf{z}_f)$ is the solution of the physical rows at the given free entries, found by Newton's method on those rows alone; $\mathbf{N}(\mathbf{z}_f)$, from the blocks of $\mathbf{G}$ at that solution, and $\mathbf{J}_{\text{red}}(\mathbf{z}_f) = \mathbf{J}_{\text{rows}}\,\mathbf{N}(\mathbf{z}_f)$ are evaluated there and

change with the free entries. The Gauss–Newton iteration (4.29) runs on the reduced rows, $(\mathbf{J}_{\mathrm{red}}^{\mathsf{T}}\mathbf{J}_{\mathrm{red}})\,\Delta\mathbf{z}_{\mathrm{f}} = -\mathbf{J}_{\mathrm{red}}^{\mathsf{T}}\mathbf{r}(\mathbf{z}_{\mathrm{f}})$, each iterate solving the physical rows again. At the MAP state, $\boldsymbol{\Sigma}_{\mathrm{red}} = (\mathbf{J}_{\mathrm{red}}^{\mathsf{T}}\mathbf{J}_{\mathrm{red}})^{-1}$ is the Gauss–Newton form of the Laplace covariance (4.38); the term it drops holds the second derivatives of the reduced rows, which now include the curvature of the set, and each chapter checks it numerically. The hall's rows are linear, $\mathbf{N}$ is constant, one step is exact and $\boldsymbol{\Sigma}_{\mathrm{red}}$ is the exact covariance.

The reduced covariance as the limit. The covariance of the reduced state is $\boldsymbol{\Sigma}_{\mathrm{red}} = (\mathbf{J}_{\mathrm{red}}^{\mathsf{T}}\mathbf{J}_{\mathrm{red}})^{-1}$, and it is the free-entry block $\boldsymbol{\Sigma}_{\mathrm{ff}}$ of the full $\boldsymbol{\Sigma}$ in the limit of exact rows. Write a move of the full state in the coordinates $\Delta\mathbf{z} = \mathbf{B}\,\Delta\mathbf{u}$ with $\mathbf{B} = [\,\mathbf{N}\;\;\mathbf{B}_{\mathrm{d}}\,]$, where $\mathbf{B}_{\mathrm{d}} = [\,\mathbf{0}\;\;\mathbf{I}_{n_{\mathrm{d}}}\,]^{\mathsf{T}}$ is the $n_{\mathrm{z}} \times n_{\mathrm{d}}$ matrix that moves only the dependent entries. The square matrix $\mathbf{B}$ is invertible and its first block row is $[\,\mathbf{I}_{n_{\mathrm{f}}}\;\;\mathbf{0}\,]$, so the first n_{f} coordinates are the move of the free entries; in the hall $\mathbf{B}$ is the matrix of equation (4.32). Since $\mathbf{G}\mathbf{N} = \mathbf{0}$ and $\mathbf{G}\mathbf{B}_{\mathrm{d}} = \mathbf{G}_{\mathrm{dep}}$, the Jacobian in these coordinates, physical rows first, and its information matrix are

$$\mathbf{J}\,[\,\mathbf{N}\;\;\mathbf{B}_{\mathrm{d}}\,] = \begin{bmatrix} \mathbf{0} & \mathbf{G}_{\mathrm{dep}} \\ \mathbf{J}_{\mathrm{red}} & \mathbf{J}_{\mathrm{rows}}\mathbf{B}_{\mathrm{d}} \end{bmatrix}, \qquad \begin{bmatrix} \mathbf{J}_{\mathrm{red}}^{\mathsf{T}}\mathbf{J}_{\mathrm{red}} & \mathbf{Y} \\ \mathbf{Y}^{\mathsf{T}} & \mathbf{Z} + \mathbf{G}_{\mathrm{dep}}^{\mathsf{T}}\mathbf{G}_{\mathrm{dep}} \end{bmatrix},$$

with $\mathbf{Y} = \mathbf{J}_{\mathrm{red}}^{\mathsf{T}}\mathbf{J}_{\mathrm{rows}}\mathbf{B}_{\mathrm{d}}$ and $\mathbf{Z} = \mathbf{B}_{\mathrm{d}}^{\mathsf{T}}\mathbf{J}_{\mathrm{rows}}^{\mathsf{T}}\mathbf{J}_{\mathrm{rows}}\mathbf{B}_{\mathrm{d}}$. By the block-inversion formula of Section 3.5.3 the free-entry block of the covariance in these coordinates, which is $\boldsymbol{\Sigma}_{\mathrm{ff}}$ because the first block row of $\mathbf{B}$ is $[\,\mathbf{I}_{n_{\mathrm{f}}}\;\;\mathbf{0}\,]$, is the inverse of the Schur complement of the lower block, and that block grows without bound as the tolerances shrink: each line of $\mathbf{G}_{\mathrm{dep}}$ carries its row's $1/\sigma_p$, so with a common tolerance $\mathbf{G}_{\mathrm{dep}}^{\mathsf{T}}\mathbf{G}_{\mathrm{dep}} = \sigma_p^{-2}\,\tilde{\mathbf{G}}_{\mathrm{dep}}^{\mathsf{T}}\tilde{\mathbf{G}}_{\mathrm{dep}}$, with $\tilde{\mathbf{G}}_{\mathrm{dep}}$ the invertible block of the unnormalized rows:

$$\boldsymbol{\Sigma}_{\mathrm{ff}} = \Big[\mathbf{J}_{\mathrm{red}}^{\mathsf{T}}\mathbf{J}_{\mathrm{red}} - \mathbf{Y}\big(\mathbf{Z} + \mathbf{G}_{\mathrm{dep}}^{\mathsf{T}}\mathbf{G}_{\mathrm{dep}}\big)^{-1}\mathbf{Y}^{\mathsf{T}}\Big]^{-1} \;\longrightarrow\; \big(\mathbf{J}_{\mathrm{red}}^{\mathsf{T}}\mathbf{J}_{\mathrm{red}}\big)^{-1} = \boldsymbol{\Sigma}_{\mathrm{red}} \qquad (\sigma_p \to 0). \quad (4.48)$$

The correction the bracket subtracts is of order σ_p^2; for nonlinear rows the equality holds with every block evaluated at the point of linearization, the MAP state. In the hall $\boldsymbol{\Sigma}_{\mathrm{ff}}$ is the scalar Σ_{11}, $\mathbf{Y} = (Q_1/\sigma_T^2,\ 0)$, $\mathbf{Z} = \mathrm{diag}(1/\sigma_T^2,\ 0)$ and $\mathbf{G}_{\mathrm{dep}}^{\mathsf{T}}\mathbf{G}_{\mathrm{dep}} = \sigma_p^{-2}\,\mathbf{I}_2$, and the bracket is the Schur complement of equation (4.33). The same coordinates give the covariance of the dependent entries in the limit, $\mathbf{N}_{\mathrm{b}}\boldsymbol{\Sigma}_{\mathrm{red}}\mathbf{N}_{\mathrm{b}}^{\mathsf{T}}$, in the hall the lower block of the limit (4.35): the dependent entries are known through the free ones, each by its row of $\mathbf{N}_{\mathrm{b}}$. The exact balance is this limit, never a substitution made by hand.

4.2.9 The fit of the hall state

The fit is run at the tolerance $\sigma_p = 0.01\,\mathrm{K}$ of Table 4.3, and the limit of exact rows is read off equation (4.35).

The step from the pre-test state. The iteration starts at the forward calculation with the pre-test resistance, $\mathbf{z}^{(0)} = (0.160,\ 22.8,\ 27.6)^{\mathsf{T}}$, where both physical rows and the prior row vanish and the reading row is $(22.8 - 23.2)/0.5 = -0.8$, so that $\mathcal{L}(\mathbf{z}^{(0)}) = 0.32$ and $\mathbf{J}^{\mathsf{T}}\mathbf{r}(\mathbf{z}^{(0)}) = (0,\ -1.6,\ 0)^{\mathsf{T}}$: only the reading is out of balance, and it pulls on T_1 alone. The normal equations (4.29) are

$$\begin{bmatrix} 45\,001\,600 & -300\,000 & -600\,000 \\ -300\,000 & 10\,004 & 0 \\ -600\,000 & 0 & 10\,000 \end{bmatrix} \Delta\mathbf{z} = \begin{bmatrix} 0 \\ 1.6 \\ 0 \end{bmatrix}, \qquad \Delta\mathbf{z} = \begin{bmatrix} 0.0092296 \\ 0.27694 \\ 0.55378 \end{bmatrix}.$$

Divided by its first entry, the step is $\Delta\mathbf{z} = 0.0092296 \times (1,\ 30.005,\ 60.000)^{\mathsf{T}}$: it lies, to within the tolerance, along the free direction $\mathbf{v}$ of equation (4.30). Table 4.4 lists the iteration. Each

line is one iterate: its three entries, the objective, and the largest entry of the gradient (4.28). After the first step the gradient is at round-off and the second step moves no entry by more than 10^{-13}.

Table 4.4: Gauss–Newton iterations for the hall state from the pre-test state (0.160, 22.8, 27.6), with $\sigma_p = 0.01\,\mathrm{K}$.

Iterate l	ρ [K/kW]	T_1 [°C]	T_2 [°C]	$\mathcal{L}$ [–]	$\max_i \lvert(\mathbf{J}^\mathsf{T}\mathbf{r})_i\rvert$ [kW/K or K^{-1}]
0	0.160000	22.8000	27.6000	0.32000	1.6
1	0.169230	23.0769	28.1538	0.09845	$< 10^{-8}$
2	0.169230	23.0769	28.1538	0.09845	$< 10^{-8}$

The MAP state. The fit returns

$$\hat{\mathbf{z}} = \begin{bmatrix} \hat{\rho} \\ \hat{T}_1 \\ \hat{T}_2 \end{bmatrix} = \begin{bmatrix} 0.16923 \\ 23.0769 \\ 28.1538 \end{bmatrix} : \tag{4.49}$$

a resistance of 0.16923 K/kW, a test temperature of 23.077°C and a full-load temperature of 28.154°C, each read straight off its entry, with the effective conductance $\hat{G} = 1/\hat{\rho} = 5.909\,\mathrm{kW/K}$ evaluated at the estimated resistance.

4.2.10 The covariance and the resistance

The covariance. Inverting the matrix (4.27),

$$\boldsymbol{\Sigma} = \left(\mathbf{J}^\mathsf{T}\mathbf{J}\right)^{-1} = \begin{bmatrix} 1.9236 \times 10^{-4} & 5.7685 \times 10^{-3} & 1.1542 \times 10^{-2} \\ 5.7685 \times 10^{-3} & 0.17309 & 0.34611 \\ 1.1542 \times 10^{-2} & 0.34611 & 0.69260 \end{bmatrix} \quad (\sigma_p = 0.01\,\mathrm{K}). \tag{4.50}$$

Its eigenvalues are 0.866 along (nearly) $\mathbf{v}$, and 1.0×10^{-4} and 2.2×10^{-8} in the two directions the physics fixes, both shrinking as σ_p^2.

The entries read off. The resistance has

$$\hat{\rho} = \hat{z}_1 = 0.16923\,\mathrm{K/kW}, \qquad \sqrt{\Sigma_{11}} = \sqrt{1.9236 \times 10^{-4}} = 0.01387\,\mathrm{K/kW},$$

the test temperature $\sqrt{\Sigma_{22}} = 0.416\,\mathrm{K}$, and the full-load temperature $\hat{T}_2 = \hat{z}_3 = 28.154$°C with $\sqrt{\Sigma_{33}} = \sqrt{0.69260} = 0.8322\,\mathrm{K}$. The correlations $\Sigma_{ij}/\sqrt{\Sigma_{ii}\Sigma_{jj}}$ are 0.99971 between ρ and T_1, 0.99993 between ρ and T_2, and 0.99964 between T_1 and T_2.

The limit of exact rows. The limit covariance (4.35) gives $\Sigma_{11} = 1/5200$, so that

$$\sqrt{\Sigma_{11}} = 0.013868\,\mathrm{K/kW}, \qquad \hat{\rho} = 880/5200 = 0.169231\,\mathrm{K/kW}, \qquad \hat{T}_2 = 18{+}60\hat{\rho} = 28.1538^\circ\mathrm{C},$$

the resistance being the minimum of equation (4.37); the rest of the chapter carries these limit values.

The objective along the free direction. Figure 4.3 draws the objective (4.36) on the line $\mathbf{z} = \hat{\mathbf{z}} + t\,\mathbf{v}$ in the limit of exact rows, with t read on the horizontal axis as the resistance $\rho = \hat{z}_1 + t$ in K/kW and the dimensionless objective on the vertical axis: the solid curve is the objective, the dashed curve the prior term and the dotted curve the reading term; the dot marks the minimum and the double arrow joins the two points where the objective has risen by $\frac{1}{2}$. At the minimum $r_3(\hat{\mathbf{z}}) = -0.2462$ and $r_4(\hat{\mathbf{z}}) = 0.3692$, and the slopes cancel, $60 \times (-0.2462) + 40 \times 0.3692 = 0$.

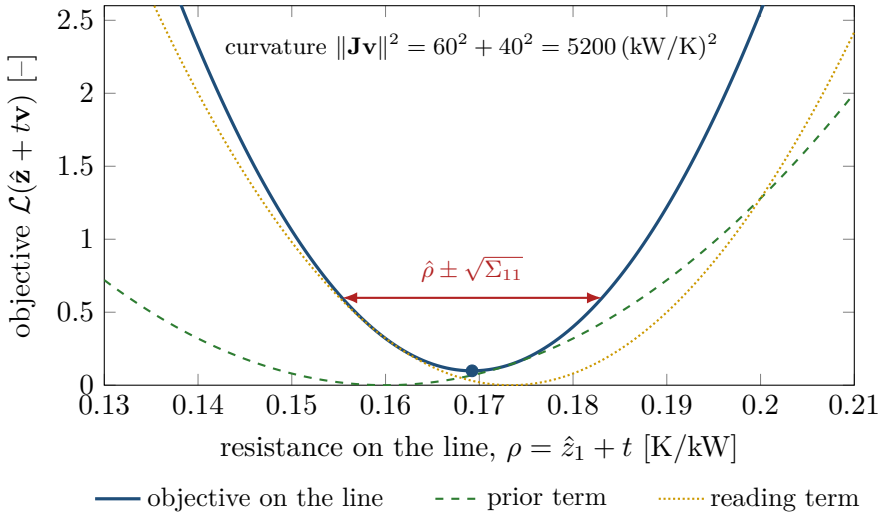


Figure 4.3: The objective on the line $\mathbf{z} = \hat{\mathbf{z}} + t\,\mathbf{v}$ along the free direction $\mathbf{v} = (1, 30, 60)^{\mathsf{T}}$, in the limit of exact physical rows, with its prior term $\frac{1}{2}(r_4(\hat{\mathbf{z}}) + 40t)^2$ and reading term $\frac{1}{2}(r_3(\hat{\mathbf{z}}) + 60t)^2$; the horizontal axis is the resistance $\hat{z}_1 + t$. Dot: the minimum at $\hat{\rho} = 0.16923\,\mathrm{K/kW}$, $\mathcal{L} = 0.0985$. Double arrow: the points $\hat{\rho} \pm \sqrt{\Sigma_{11}}$, $\sqrt{\Sigma_{11}} = 0.01387\,\mathrm{K/kW}$, where the objective has risen by $\frac{1}{2}$.

The posterior of the resistance. By the first marginal of equation (4.31),

$$\rho \mid \mathbf{y} \sim \mathcal{N}(\hat{z}_1,\ \Sigma_{11}) = \mathcal{N}(0.169231,\ 0.013868^2), \tag{4.51}$$

a standard deviation of $0.013868/0.169231 = 0.082$ of the estimate. Figure 4.4 compares it with the prior. The horizontal axis is the resistance and the vertical axis the probability density in kW/K; the dashed curve is the prior of equation (4.12) and the solid curve the posterior of equation (4.51). The reading has shifted the resistance upward, from 0.160 to 0.169, and narrowed it, from a standard deviation of 0.025 to 0.0139.

4.2.11 The reduced state of the hall

At the tolerance $\sigma_p = 0.01\,\mathrm{K}$ of equation (4.25), $\mathbf{G}_{\mathrm{free}} = (-3000,\ -6000)^{\mathsf{T}}$ and $\mathbf{G}_{\mathrm{dep}} = \mathrm{diag}(100,\ 100)$. From the pre-test resistance 0.160 the reduced rows (4.46) are $(-0.8,\ 0)$, $\mathbf{J}_{\mathrm{red}}^{\mathsf{T}}\mathbf{r} = -48$, and one Gauss–Newton step $48/5200 = 0.0092308$ lands on $\hat{\rho} = 0.169231\,\mathrm{K/kW}$, which a second step does not move. The carried entries have $\mathbf{N}_{\mathrm{b}}\mathbf{\Sigma}_{\mathrm{red}}\mathbf{N}_{\mathrm{b}}^{\mathsf{T}}$ with diagonal 0.17308 and $0.69231\,\mathrm{K}^2$ and off-diagonal $0.34615\,\mathrm{K}^2$.

Table 4.5 sets the reduced state against the full one. Its first two rows give $\mathbf{N}$ and $\mathbf{J}_{\mathrm{red}}$, which do not depend on the tolerance; the remaining rows give, for the full state at three tolerances and for the reduced state in the last column, the (ρ, ρ) entry of the covariance, its square root, the MAP resistance, and the relative difference of the full entry from $\mathbf{\Sigma}_{\mathrm{red}}$. The difference shrinks a hundredfold per decade of σ_p, as the order σ_p^2 of the correction in equation (4.48) requires, and at $\sigma_p = 0.001\,\mathrm{K}$ the two agree to every printed digit.

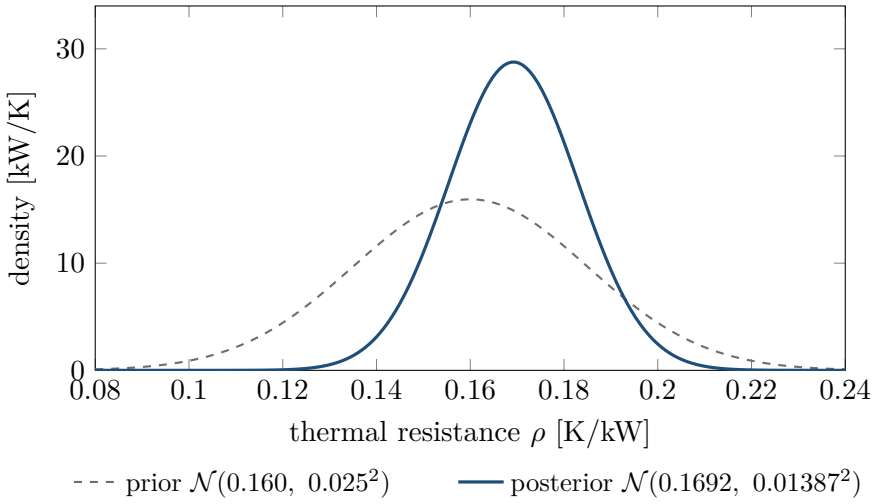


Figure 4.4: Prior and posterior densities of the thermal resistance, the first entry of the state: $\mathcal{N}(0.160, 0.025^2)$ and $\mathcal{N}(0.169231, 0.013868^2)$. The reading moves the resistance up and reduces its spread.

Table 4.5: The reduced state of the hall against the free-entry marginal of the full state. $\mathbf{N}$ is in [–, kW, kW]; the last column is the reduced state, exact in the limit of exact rows.

Quantity	$\sigma_p = 0.1\,\mathrm{K}$	$0.01\,\mathrm{K}$	$0.001\,\mathrm{K}$	reduced state
null-space basis $\mathbf{N}$	$(1,\ 30,\ 60)^\mathsf{T}$ at every tolerance			
$\mathbf{J}_\mathrm{red} = \mathbf{J}_\mathrm{rows}\mathbf{N}$ [kW/K]	$(60,\ 40)^\mathsf{T}$, information $5200\,(\mathrm{kW/K})^2$			
Σ_{11}, $\mathbf{\Sigma}_\mathrm{red}$ $[(\mathrm{K/kW})^2]$	1.97568×10^{-4}	1.92361×10^{-4}	1.92308×10^{-4}	1.92308×10^{-4}
standard deviation [K/kW]	0.014056	0.013869	0.013868	0.013868
$\hat{\rho}$ [K/kW]	0.169119	0.169230	0.169231	0.169231
relative difference from $\mathbf{\Sigma}_\mathrm{red}$ [–]	2.7×10^{-2}	2.8×10^{-4}	2.8×10^{-6}	–

4.3 The answers (step 4)

Each question of Table 4.2 is read from the posterior of Section 4.2.

4.3.1 Uncertainty propagation and the delta method

The rows of the hall are linear, so the entries of $\boldsymbol{\Sigma}$ are exact variances. In most plant models an output is a nonlinear function of the state, $f(\mathbf{z})$, and the posterior is Gaussian at best approximately, $\mathbf{z} \mid \mathbf{y} \approx \mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$.

Linearization of the output. The first-order Taylor expansion of f about $\hat{\mathbf{z}}$ is

$$f(\mathbf{z}) \approx f(\hat{\mathbf{z}}) + \nabla f(\hat{\mathbf{z}})^{\mathsf{T}}(\mathbf{z} - \hat{\mathbf{z}}).$$

Writing $\mathbf{g} = \nabla f(\hat{\mathbf{z}})$ for the column gradient,

$$f(\mathbf{z}) \approx f(\hat{\mathbf{z}}) + \mathbf{g}^{\mathsf{T}}(\mathbf{z} - \hat{\mathbf{z}}). \tag{4.52}$$

Taking the posterior expectation, with $\mathbb{E}[\mathbf{z} - \hat{\mathbf{z}} \mid \mathbf{y}] = \mathbf{0}$ because $\hat{\mathbf{z}}$ is the mean of the Gaussian,

$$\mu_f = \mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}] \approx f(\hat{\mathbf{z}}).$$

Taking the variance, the constant $f(\hat{\mathbf{z}})$ drops out, and the variance of the linear combination $\mathbf{g}^{\mathsf{T}}(\mathbf{z} - \hat{\mathbf{z}}) = \sum_i [\mathbf{g}]_i (z_i - \hat{z}_i)$ is the double sum of its covariances,

$$\operatorname{Var}(f(\mathbf{z}) \mid \mathbf{y}) \approx \operatorname{Var}\Big(\sum_i [\mathbf{g}]_i z_i \;\Big|\; \mathbf{y}\Big) = \sum_i \sum_j [\mathbf{g}]_i \, [\mathbf{g}]_j \, \operatorname{Cov}(z_i, z_j \mid \mathbf{y}),$$

which in matrix form is

$$\operatorname{Var}(f(\mathbf{z}) \mid \mathbf{y}) \approx \mathbf{g}^{\mathsf{T}} \boldsymbol{\Sigma}\, \mathbf{g}. \tag{4.53}$$

This is the first-order delta method, one of the most reused formulas in probabilistic plant modeling. Its only approximation is the linearization (4.52); when f is linear and the posterior Gaussian, it is exact.

Application to the thermal system. The full-load temperature is the third entry of the state, so its gradient is the third unit vector and the quadratic form picks one diagonal entry,

$$\mathbf{g} = \frac{\partial T_2}{\partial \mathbf{z}} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \qquad \mathbf{g}^{\mathsf{T}} \boldsymbol{\Sigma}\, \mathbf{g} = \Sigma_{33}.$$

The same temperature can also be predicted as a derived quantity of the resistance alone, $f(\mathbf{z}) = T_c + Q_2 \rho = 18 + 60\rho$, with

$$\mathbf{g} = \frac{\partial f}{\partial \mathbf{z}} = \begin{bmatrix} 60 \\ 0 \\ 0 \end{bmatrix}, \qquad \mathbf{g}^{\mathsf{T}} \boldsymbol{\Sigma}\, \mathbf{g} = 60^2\, \Sigma_{11}.$$

In the limit the two are equal, since the limit covariance (4.35) has $\Sigma_{33} = 3600/5200 = 60^2\, \Sigma_{11}$; at a finite tolerance they differ by σ_p^2, equation (4.34), because the entry carries the discrepancy of the full-load balance and the prediction from ρ does not. The two calculations, reading the variance off the covariance and propagating the resistance's uncertainty through a gradient, agree in the limit because they are the same transformation of the same uncertainty: an output that is an entry of the state needs no propagation, and an output that is not needs its gradient.

4.3.2 The gradient and the variance budget

Physical meaning of the gradient. The first entry of the gradient $\mathbf{g} = (60,\ 0,\ 0)^{\mathsf{T}}$ of the prediction $18 + 60\rho$, 60 kW, means that a change of 0.01 K/kW in the resistance changes the full-load temperature by $\Delta T_2 = 60 \times 0.01 = 0.6\,\mathrm{K}$. The gradient is the link between physical uncertainty and its consequence for the asset. A state entry can have a large posterior variance and still contribute little to an output that is insensitive to it, with $g_i \approx 0$; conversely, a modest uncertainty in a highly influential entry can dominate the uncertainty of a prediction.

Several uncertain entries. Suppose the supply-air temperature is uncertain as well, the controller holding it only near its set point, so that the full-load temperature $T = T_c + Q\rho$ depends on the sub-vector $\boldsymbol{\theta} = (T_c, \rho)^{\mathsf{T}}$ of the state, with posterior covariance $\boldsymbol{\Sigma}_\theta$. The gradient is $\mathbf{g} = (\partial T/\partial T_c,\ \partial T/\partial \rho)^{\mathsf{T}} = (1,\ Q)^{\mathsf{T}}$, and equation (4.53) gives

$$\mathrm{Var}(T) = \mathbf{g}^{\mathsf{T}} \boldsymbol{\Sigma}_\theta\, \mathbf{g} = \begin{bmatrix} 1 & Q \end{bmatrix} \begin{bmatrix} \mathrm{Var}(T_c) & \mathrm{Cov}(T_c, \rho) \\ \mathrm{Cov}(T_c, \rho) & \mathrm{Var}(\rho) \end{bmatrix} \begin{bmatrix} 1 \\ Q \end{bmatrix}.$$

Multiplying out,

$$\mathrm{Var}(T) = \mathrm{Var}(T_c) + Q^2\ \mathrm{Var}(\rho) + 2Q\ \mathrm{Cov}(T_c, \rho). \tag{4.54}$$

This is the variance budget of the output: one term per entry, each its sensitivity squared times its variance, plus a cross term from their covariance. The first term is the spread the supply air alone puts on the temperature; the second, the spread the resistance alone puts on it, $Q^2\ \mathrm{Var}(\rho)$; the third corrects for the two moving together in the posterior. If the readings cannot separate the two entries, the posterior leaves them correlated, and the cross term either adds to the diagonal terms or cancels part of them. With a negative correlation and same-signed sensitivities it is negative, and the output can be well known although each entry is poorly known: the readings then fix the combination $T_c + Q\rho$ that the output depends on, not the entries separately. Adding the two individual variances without the cross term is wrong in either direction, which is why the joint posterior, not its marginals, must be kept; Exercise 4.2 works a case in which ignoring the cross term overstates the spread by more than half.

4.3.3 The density of a transformed quantity

The delta method gives a derived quantity a mean and a variance. Some questions need more: the whole density of the quantity, or a tail probability far from the center, where the shape matters and a Gaussian with the right variance can still be wrong. The tool for these is the *change of variables*, and it is used throughout the book whenever a quantity is computed in one coordinate and reported in another.

Probability is conserved; density is not. Let x have the density $p_x(x)$, and let $w = h(x)$ be a monotone function with inverse $x = h^{-1}(w)$. A short interval $[x, x + \mathrm{d}x]$ is mapped onto a short interval $[w, w + \mathrm{d}w]$, and the two carry the same probability, because they hold the same outcomes:

$$p_w(w)\,|\mathrm{d}w| = p_x(x)\,|\mathrm{d}x|.$$

Dividing by $|\mathrm{d}w|$ gives the density of the transformed quantity,

$$p_w(w) = p_x\big(h^{-1}(w)\big)\ \left|\frac{\mathrm{d}x}{\mathrm{d}w}\right|. \tag{4.55}$$

The first factor evaluates the old density at the point that maps to w. The second, the *change-of-variable factor*, records how much the map stretches or squeezes the axis there: where a small change of w corresponds to a large change of x, the probability of the long x-interval is packed

into a short w-interval, and the density of w must be larger to hold it. Without the factor, $p_x(h^{-1}(w))$ is not a density of w at all: it does not in general integrate to one. A monotone map also carries quantiles over unchanged: if h is increasing, $\Pr(w \le h(x^*)) = \Pr(x \le x^*)$, so the median of w is h of the median of x, while the mode, which depends on the factor, moves.

The conductance. For the conductance $G = 1/\rho$ of the hall, the inverse map is $\rho = 1/G$ and $|\mathrm{d}\rho/\mathrm{d}G| = 1/G^2$, so equation (4.55) gives

$$p_G(G) = \frac{1}{G^2}\, p(\rho = 1/G \mid \mathbf{y}).$$

It is skewed to the right of $\hat{G}$ because the factor $1/G^2$ is larger at small G, which squeezes the left of the curve and stretches its right.

Two routes to a derived quantity. The delta method and the change of variables answer the same question with different approximations. For a scalar entry with posterior $x \mid \mathbf{y} \approx \mathcal{N}(\hat{x}, \sigma^2)$ and a monotone output $w = h(x)$:

- the *delta method* linearizes h at $\hat{x}$, with gradient $g = \mathrm{d}w/\mathrm{d}x$ evaluated there, and reports $w \mid \mathbf{y} \approx \mathcal{N}(h(\hat{x}),\ g^2\sigma^2)$, a Gaussian in w;
- the *change of variables* keeps h exact and reports $p_w(w) = p_x(h^{-1}(w))\,|\mathrm{d}x/\mathrm{d}w|$, with the Gaussian of x as its only approximation.

The two factors are reciprocals: the delta method's gradient $\mathrm{d}w/\mathrm{d}x$ is the inverse of the change-of-variable factor $\mathrm{d}x/\mathrm{d}w$, but the delta method evaluates it once, at $\hat{x}$, while the change of variables evaluates it at every w. The delta method is the change of variables with the map replaced by its tangent, which is why the two agree on the spread near the center and differ in the tails. When h is linear the factor is constant and the two coincide. When h bends across the posterior, the delta method puts a symmetric Gaussian on a quantity whose density is skewed, and can put probability where the quantity cannot go: below zero for a conductance or a rate, outside $(0, 1)$ for a probability such as a probability of detection.

> **Principle 4.1.** *The gradient carries a spread; the change-of-variable factor carries a density. Report $\sqrt{\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}}$ when a standard deviation is the answer, and transform the density when the answer is a shape, a tail probability or a bound near the edge of what the quantity can be.*

4.3.4 Credible intervals

The asset manager often needs an interval for an uncertain quantity as well as its mean. For the Gaussian marginal $T_2 \mid \mathbf{y} \sim \mathcal{N}(\hat{z}_3, \Sigma_{33})$ of equation (4.31) the central 95 % credible interval is the interval that leaves 2.5 % of the posterior in each tail. Since $\Phi(1.960) = 0.975$,

$$\left[\hat{z}_3 - 1.960\,\sqrt{\Sigma_{33}},\ \hat{z}_3 + 1.960\,\sqrt{\Sigma_{33}}\right]. \tag{4.56}$$

The same form serves any entry, with $\hat{z}_i$ and $\sqrt{\Sigma_{ii}}$, and a derived quantity, with $f(\hat{\mathbf{z}})$ and $\sqrt{\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}}$ from the delta method (4.53).

The interval and a threshold probability carry different information. The interval describes the range that holds most of the posterior probability; a threshold probability answers directly whether the plant meets a stated limit. For an acceptance decision the threshold probability is usually the more actionable of the two. Neither is meaningful apart from the model assumptions, the quality of the measurement and the sources of uncertainty that the model includes; Section 4.6.7 lists those it leaves out.

4.3.5 The temperature-threshold probability

The question of equation (4.10) needs one marginal of the posterior. The marginal of the full-load temperature is read off the third entry of the state, equation (4.31): $T_2 \mid \mathbf{y} \sim \mathcal{N}(\hat{z}_3,\ \Sigma_{33})$.

The standardized temperature. For a Gaussian variable the standardized variable $(T_2 - \hat{z}_3)/\sqrt{\Sigma_{33}}$ has the standard normal distribution, and subtracting a constant and dividing by a positive constant preserve the inequality, so

$$\Pr(T_2 < T_{\max} \mid \mathbf{y}) = \Pr\!\left(\frac{T_2 - \hat{z}_3}{\sqrt{\Sigma_{33}}} < \frac{T_{\max} - \hat{z}_3}{\sqrt{\Sigma_{33}}} \;\Big|\; \mathbf{y}\right) = \Phi\!\left(\frac{T_{\max} - \hat{z}_3}{\sqrt{\Sigma_{33}}}\right). \tag{4.57}$$

The probability of exceeding the limit is its complement, $1 - \Pr(T_2 < T_{\max} \mid \mathbf{y})$.

The prior marginal. Before the test reading the state had the prior formed by the physical and prior rows alone. Its Jacobian $\mathbf{J}_\mathrm{p}$ is $\mathbf{J}$ of equation (4.25) without the reading's line, its mean is the pre-test state $\mathbf{z}^{(0)} = (0.160,\ 22.8,\ 27.6)^\mathsf{T}$, and its covariance $\mathbf{\Sigma}_\mathrm{p} = (\mathbf{J}_\mathrm{p}^\mathsf{T}\mathbf{J}_\mathrm{p})^{-1}$ tends, by the argument of equation (4.35) with $\mathbf{J}_\mathrm{p}\mathbf{v} = (0,\ 0,\ 40)^\mathsf{T}$ in place of $\mathbf{J}\mathbf{v}$, to $\mathbf{v}\mathbf{v}^\mathsf{T}/40^2 = \sigma_0^2\,\mathbf{v}\mathbf{v}^\mathsf{T}$. Its third diagonal entry is $(\mathbf{\Sigma}_\mathrm{p})_{33} = 60^2 \times 0.025^2 = 2.25\,\mathrm{K}^2$, so the full-load temperature had the prior marginal $\mathcal{N}(27.6,\ 1.5^2)$, the third entry of $\mathbf{z}^{(0)}$ with $\sqrt{(\mathbf{\Sigma}_\mathrm{p})_{33}} = 1.5\,\mathrm{K}$.

The question through the resistance. In the limit of exact rows the posterior lies on the line $\mathbf{z} = \hat{\mathbf{z}} + t\,\mathbf{v}$ through the MAP state along the free direction, $\mathbf{\Sigma} = \mathbf{v}\mathbf{v}^\mathsf{T}/5200$ by equation (4.35). On that line the third entry is $\hat{z}_3 + 60t$ and the first $\hat{z}_1 + t$, so the event $\{T_2 < T_{\max}\}$ is the event $\{t < t^*\}$ with $t^* = (T_{\max} - \hat{z}_3)/60$, which is the event that the resistance stays below

$$\rho^* = \hat{z}_1 + t^* = \hat{z}_1 + \frac{T_{\max} - \hat{z}_3}{60}. \tag{4.58}$$

With the marginal (4.31) of the first entry,

$$\Pr(\rho < \rho^* \mid \mathbf{y}) = \Phi\!\left(\frac{\rho^* - \hat{z}_1}{\sqrt{\Sigma_{11}}}\right) = \Phi\!\left(\frac{t^*}{\sqrt{\Sigma_{11}}}\right),$$

which is the probability (4.57): both standardized distances are t^* divided by the spread of t, since the third entry moves 60 times as far as the first along $\mathbf{v}$ and $\sqrt{\Sigma_{33}} = 60\sqrt{\Sigma_{11}}$ in the limit. Reading the third entry and propagating the first one along the free direction are the same calculation; the full-load question is a question about the resistance.

The dependence on the value read. The information matrix (4.27), and with it $\mathbf{\Sigma}$, does not contain the value observed, so a different reading leaves every standard deviation where it is and moves only the mean, along one vector. The reading enters the normal equations only through the part $-y_T/\sigma_T$ of the third entry of the residual vector that depends on y_T; that row's line of the Jacobian is $\mathbf{h}^\mathsf{T}/\sigma_T$, so differentiating $\hat{\mathbf{z}} = \mathbf{z} - \mathbf{\Sigma}\mathbf{J}^\mathsf{T}\mathbf{r}(\mathbf{z})$ with respect to y_T gives the gain vector

$$\frac{\partial \hat{\mathbf{z}}}{\partial y_T} = \frac{\mathbf{\Sigma}\mathbf{h}}{\sigma_T^2}, \tag{4.59}$$

the second column of $\mathbf{\Sigma}$ divided by σ_T^2. The threshold probability (4.57) is computed with a spread fixed before the reading and a mean set by its value, so the reading can move it in either direction.

4.3.6 The posterior state and the full-load temperature

General form. The three estimation questions, the resistance of the air path and the hall temperature at the test and at full load, share one form: the posterior of a quantity $f(\mathbf{z})$, equation (2.9), summarized by its moments (2.10) and its central interval (2.11).

Specialized form. Equations (2.10) and (2.11) with $f(\mathbf{z}) = z_i$, the entry asked about, and $p(\mathbf{z} \mid \mathbf{y})$ the Gaussian posterior of step 3: the mean is $\hat{z}_i$, the variance Σ_{ii}, by the marginals (4.31), and the interval is equation (4.56). The rows are linear and Gaussian, so these are the exact moments, with no approximation to set beside them.

Building blocks. The MAP state and the limit covariance (4.35), with its limit values of Section 4.2.10; the prior marginal $\mathcal{N}(27.6,\ 1.5^2)$ of Section 4.3.5, for the prior interval.

The posterior state. Table 4.6 collects it: each row is one entry of $\hat{\mathbf{z}}$, with its posterior mean and its standard deviation $\sqrt{\Sigma_{ii}}$, in the limit $\sigma_p \to 0$; the full-load spread is twice the test spread, the ratio 60/30 of the entries of $\mathbf{v}$.

Table 4.6: The posterior state of the equipment hall after the test reading, in the limit of exact physical rows.

State entry	Posterior mean	Standard deviation
ρ [K/kW]	0.16923	0.01387
T_1 [°C; sd in K]	23.077	0.416
T_2 [°C; sd in K]	28.154	0.832

The diagonal of the limit covariance. The limit covariance (4.35) gives $\Sigma_{11} = 1/5200 = 1.9231 \times 10^{-4}\,(\mathrm{K/kW})^2$, $\Sigma_{22} = 900/5200 = 0.17308\,\mathrm{K}^2$ ($\sqrt{\Sigma_{22}} = 0.4160\,\mathrm{K}$) and

$$\Sigma_{33} = \mathrm{Var}(T_2 \mid \mathbf{y}) = \frac{3600}{5200} = 0.69231\,\mathrm{K}^2, \qquad \sqrt{\Sigma_{33}} = 0.8321\,\mathrm{K}. \tag{4.60}$$

The credible interval. By equation (4.56), $28.1538 \pm 1.960 \times 0.8321 = 28.1538 \pm 1.6308$, so that

$$T_2 \in [26.52,\ 29.78]^\circ\mathrm{C}$$

with 95 % posterior probability under the stated model; the prior interval was $[24.66,\ 30.54]^\circ$C, 1.8 times as wide.

The conductance beside the resistance. The conductance is a derived quantity, $G(\mathbf{z}) = 1/z_1$, so the delta method (4.53) reads it from the same posterior, with

$$\mathbf{g} = \frac{\partial G}{\partial \mathbf{z}} = \begin{bmatrix} -1/\hat{z}_1^2 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -34.92 \\ 0 \\ 0 \end{bmatrix} (\mathrm{kW/K})^2, \qquad \mathrm{Var}(G \mid \mathbf{y}) \approx \mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g} = \frac{\Sigma_{11}}{\hat{z}_1^4},$$

which gives $\hat{G} = 1/\hat{\rho} = 5.909\,\mathrm{kW/K}$ with $\mathrm{sd}(G \mid \mathbf{y}) \approx 0.484\,\mathrm{kW/K}$ and the 5–95 % band $[5.113,\ 6.706]\,\mathrm{kW/K}$. The exact band, by the change of variables (4.55), is $[5.207,\ 6.830]\,\mathrm{kW/K}$, higher by 0.095 and 0.124 kW/K; the median is 5.909 kW/K on both routes.

4.3.7 The threshold probability at full load

General form. A requirement that a quantity stay below a limit is the posterior probability of the event, equation (2.12), compared with an assurance level κ.

Specialized form. Equation (2.12) with $f(\mathbf{z}) = -T_2 = -z_3$, $f^* = -T_{\max}$ and $p(\mathbf{z} \mid \mathbf{y})$ the Gaussian posterior of step 3 becomes the normal tail (4.57) of the third entry. The commissioning test states no assurance level κ; the probability is reported without an accept verdict. The posterior is exactly Gaussian, so the tail is the exact probability under the model.

Building blocks. The third entry $\hat{z}_3$ and the variance Σ_{33} of equation (4.60), and the standard normal distribution function Φ.

The probability. With $\hat{z}_3 = 28.1538\,^\circ\mathrm{C}$ and $\sqrt{\Sigma_{33}} = 0.8321\,\mathrm{K}$, the standardized distance in equation (4.57) is $(28 - 28.1538)/0.8321 = -0.1849$, and

$$\Pr(T_2 < 28\,^\circ\mathrm{C} \mid \mathbf{y}) = \Phi(-0.1849) = 0.4267. \tag{4.61}$$

The probability of exceeding the limit is $1 - 0.4267 = 0.5733$. The fit at $\sigma_p = 0.01\,\mathrm{K}$ gives 0.4267 as well, with $\sqrt{\Sigma_{33}} = 0.83223\,\mathrm{K}$ against the limit's $0.83205\,\mathrm{K}$.

Against the prior. Under the prior marginal $\mathcal{N}(27.6,\ 1.5^2)$ of Section 4.3.5, before the test reading, the probability that the full-load temperature stays below the limit was

$$\Pr(T_2 < 28) = \Phi\left(\frac{28 - 27.6}{1.5}\right) = \Phi(0.2667) = 0.6051.$$

Figure 4.5 draws the answer. The horizontal axis is the full-load temperature and the vertical axis its density in K^{-1}. The solid curve is the posterior marginal, the dashed curve the prior marginal, and the vertical line the limit $T_{\max}$. The shaded area under the posterior to the left of the line is the probability (4.61); the unshaded area under it to the right is the probability of exceeding the limit. The posterior is narrower than the prior and centered to the right of it, across the limit.

4.4 The checks (step 5)

Obtaining a posterior is not the same as establishing that it is credible.

4.4.1 Algebraic and physical verification

Algebraic verification. At the fitted state of equation (4.49), with $\sigma_p = 0.01\,\mathrm{K}$, the residual vector (4.23) is $\mathbf{r}(\hat{\mathbf{z}}) = (0.0049225,\ 0,\ -0.24612,\ 0.36919)^\mathsf{T}$, and the gradient (4.28), column by column of $\mathbf{J}$,

$$\mathbf{J}^\mathsf{T}\mathbf{r}(\hat{\mathbf{z}}) = \begin{bmatrix} -3000 \times 0.0049225 + 40 \times 0.36919 \\ 100 \times 0.0049225 + 2 \times (-0.24612) \\ 100 \times 0 \end{bmatrix} = \begin{bmatrix} -14.767 + 14.767 \\ 0.4922 - 0.4922 \\ 0 \end{bmatrix} = \mathbf{0},$$

to numerical precision: the state satisfies the first-order optimality condition. Each line reads as a balance of rows on one entry: the prior's pull on ρ is held by the first physical row, which in turn is held at T_1 by the reading row; nothing pulls on T_2 but its own physical row. The information matrix (4.27) has positive eigenvalues, so the objective is strictly convex and the stationary point is its unique minimum.

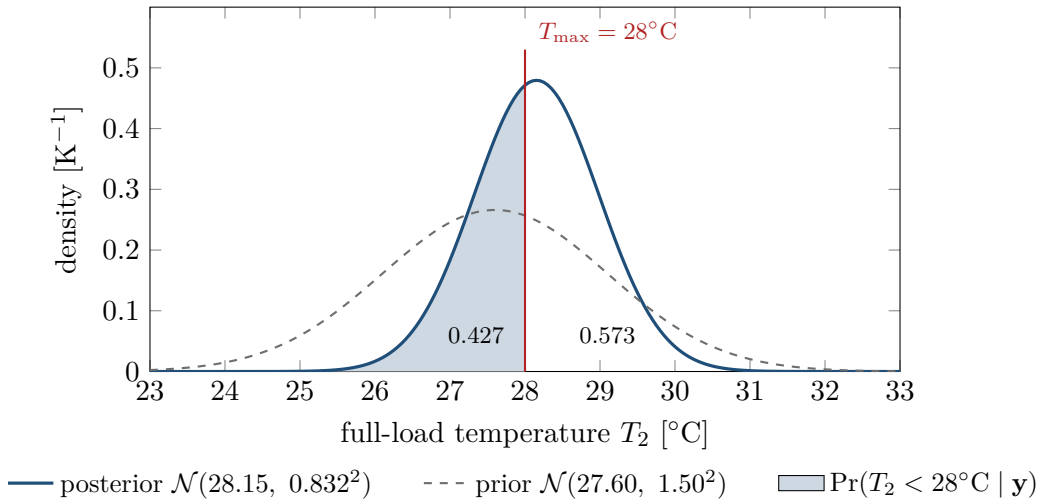


Figure 4.5: The threshold probability of equation (4.57). Posterior marginal of the full-load temperature, $\mathcal{N}(28.1538, 0.8321^2)$, with the area below $T_{\max} = 28°\mathrm{C}$ shaded ($\Phi(-0.1849) = 0.4267$), and the prior marginal $\mathcal{N}(27.6, 1.5^2)$, under which the same probability is $\Phi(0.2667) = 0.6051$.

Physical consistency. The physical rows at the fitted state, in kelvin, are $\sigma_p r_1 = 4.9 \times 10^{-5}\,\mathrm{K}$ and $\sigma_p r_2 = 0$: both halls balance to well within the tolerance, and the first imbalance shrinks with σ_p. The difference between the observed and the fitted test temperature is $23.2 - 23.0769 = 0.1231\,\mathrm{K}$, a quarter of the assumed sensor standard deviation of $0.5\,\mathrm{K}$. The model is consistent with this reading. Agreement with one reading does not validate the physical model: with one reading and one free direction, the fit has nearly enough freedom to reproduce any value.

4.4.2 Prior predictive check

Before the reading was used, the model and the prior predicted a distribution for it. By equation (4.11), $Y = \mathbf{h}^{\mathsf{T}}\mathbf{z} + \epsilon$, with the state under its prior $\mathcal{N}(\mathbf{z}^{(0)}, \boldsymbol{\Sigma}_{\mathrm{p}})$ of Section 4.3.5 and ϵ independent of it, so Y is Gaussian with

$$\mathbb{E}[Y] = \mathbf{h}^{\mathsf{T}}\mathbf{z}^{(0)} = 22.8°\mathrm{C}, \quad \mathrm{Var}(Y) = \mathbf{h}^{\mathsf{T}}\boldsymbol{\Sigma}_{\mathrm{p}}\mathbf{h} + \sigma_T^2 = (\boldsymbol{\Sigma}_{\mathrm{p}})_{22} + \sigma_T^2 = 0.5625 + 0.25 = 0.8125\,\mathrm{K}^2,$$

the variances adding because the two terms are independent, and $(\boldsymbol{\Sigma}_{\mathrm{p}})_{22} = 30^2 \times 0.025^2$ in the limit of exact rows. Thus

$$Y \sim \mathcal{N}(22.8,\ 0.9014^2). \tag{4.62}$$

The observed reading lies $(23.2 - 22.8)/0.9014 = 0.444$ standard deviations from the predicted mean, an unremarkable distance: the reading is compatible with the prior information and the model. This check is distinct from computing the posterior. It asks whether the evidence was plausible under the assumptions before those assumptions were used to interpret it; a reading several predictive standard deviations away would have signaled a wrong prior, a wrong model or a faulty sensor before any posterior was formed.

4.4.3 The evidence by two routes

The evidence (4.18) of the model is the prior predictive density (4.62) evaluated at the reading,

$$p(y_T \mid M) = \mathcal{N}(y_T;\ \mathbf{h}^{\mathsf{T}}\mathbf{z}^{(0)},\ \mathbf{h}^{\mathsf{T}}\boldsymbol{\Sigma}_{\mathrm{p}}\mathbf{h} + \sigma_T^2),$$

with $\mathcal{N}(y;\mu,\sigma^2)$ the Gaussian density of Section 4.2.1. With $\mu = 22.8°\mathrm{C}$ and $\sigma^2 = 0.8125\,\mathrm{K}^2$,

$$p(y_T = 23.2 \mid M) = \frac{1}{\sqrt{2\pi \times 0.8125}} \exp\left[-\frac{(23.2-22.8)^2}{2\times 0.8125}\right] = 0.44259 \times 0.90623 = 0.4011\,\mathrm{K}^{-1}. \tag{4.63}$$

The Laplace form (4.19) gives the same number from the fit. At $\sigma_p = 0.01\,\mathrm{K}$, $\det(\boldsymbol{\Sigma})^{1/2} = 1.3867 \times 10^{-6}$ and $\mathcal{L}(\hat{\mathbf{z}}) = 0.09845$ (Table 4.4), so

$$p(y_T \mid M) = 0.39894 \times \frac{\det(\boldsymbol{\Sigma})^{1/2}}{\sigma_p^2} \times \frac{\mathrm{e}^{-\mathcal{L}(\hat{\mathbf{z}})}}{\sigma_T\,\sigma_0} = 0.39894 \times 0.013867 \times 0.90624 \times 80 = 0.4011\,\mathrm{K}^{-1},$$

with $\det(\boldsymbol{\Sigma})^{1/2}/\sigma_p^2 = 0.013867$ against its limit $1/\sqrt{5200} = 0.013868$, the width of the posterior along the free direction.

Common mistake

Reading an evidence value as a probability that the model is correct. The density $0.4011\,\mathrm{K}^{-1}$ is not a 40 % confidence in the thermal model: it is a density, in inverse units of the reading, whose value changes by a factor of 1000 if the reading is expressed in millikelvin. A posterior probability of a model requires the evidence of the competing models and their prior probabilities. Likewise, the posterior probability 0.4267 is a confidence in the claim about T_2 under the model, while the evidence ratio is a confidence in the model the claim was computed under; the two answer different questions.

4.4.4 Posterior predictive validation with a new observation

Suppose a further, independent reading is to be taken at the same 30 kW condition. Before its value is seen, its distribution given the evidence already in hand is the posterior predictive. The new reading is $Y_{\mathrm{new}} = \mathbf{h}^{\mathsf{T}}\mathbf{z} + \epsilon_{\mathrm{new}}$, with $\mathbf{z} \mid \mathbf{y} \sim \mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ and $\epsilon_{\mathrm{new}} \sim \mathcal{N}(0, \sigma_T^2)$ independent of it, so

$$Y_{\mathrm{new}} \mid \mathbf{y} \sim \mathcal{N}\big(\mathbf{h}^{\mathsf{T}}\hat{\mathbf{z}},\ \mathbf{h}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{h} + \sigma_T^2\big) = \mathcal{N}\big(23.0769,\ 0.17308 + 0.25\big) = \mathcal{N}\big(23.0769,\ 0.6504^2\big), \tag{4.64}$$

with $\mathbf{h}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{h} = \Sigma_{22}$ from equation (4.35). The predictive variance has two parts: the remaining uncertainty about the test temperature, $0.173\,\mathrm{K}^2$, and the error of the new reading itself, $0.25\,\mathrm{K}^2$. A further reading of the same sensor can therefore never be predicted more precisely than σ_T, however well the resistance becomes known.

An independent new reading compared with this distribution assesses the predictive performance of the model: a reading a fraction of a predictive standard deviation from 23.0769°C passes, and one several standard deviations away does not. Repeated large discrepancies, or a systematic sign in them, would indicate that the model, the prior, the measurement-error assumptions or the representation of the operating condition need revision.

4.4.5 The limit of exact rows

Table 4.7 repeats the fit at three tolerances and in the limit $\sigma_p \to 0$. Each column is one tolerance; the rows give the MAP resistance and full-load temperature, their standard deviations, the correlation of the two hall temperatures, the threshold probability of Section 4.3.5, and the condition number of the information matrix. By $\sigma_p = 0.01\,\mathrm{K}$ the full-load temperature and the threshold probability have stopped moving at the printed precision; the resistance and its spread still move in their sixth decimal, the full-load spread in its fourth; at $\sigma_p = 0.001\,\mathrm{K}$ all have stopped. The fit at $\sigma_p = 0.001\,\mathrm{K}$ reproduces the limit values of Section 4.2.10 to every printed digit.

Table 4.7: The fit of the hall state as the model tolerance of the physical rows shrinks. The last column is the limit of exact rows.

Quantity	$\sigma_p = 0.1\,\mathrm{K}$	$0.01\,\mathrm{K}$	$0.001\,\mathrm{K}$	$\sigma_p \to 0$
$\hat{\rho}$ [K/kW]	0.169119	0.169230	0.169231	0.169231
$\sqrt{\Sigma_{11}}$ [K/kW]	0.014056	0.013869	0.013868	0.013868
$\hat{T}_2$ [°C]	28.1471	28.1538	28.1538	28.1538
$\sqrt{\Sigma_{33}}$ [K]	0.84926	0.83223	0.83205	0.83205
correlation of T_1, T_2 [–]	0.96522	0.99964	0.999996	1
$\Pr(T_2 < 28°\mathrm{C})$ [–]	0.4312	0.4267	0.4267	0.4267
condition number of $\mathbf{J}^\mathsf{T}\mathbf{J}$ [–]	4.0×10^5	3.9×10^7	3.9×10^9	–

The $\sqrt{\Sigma_{11}}$ row widens with the tolerance as equation (4.33) requires, the tolerance acting on the reading as extra sensor variance, and the limit column is the matrix of equation (4.35); its entry for $\sqrt{\Sigma_{11}}$ is the square root of $\boldsymbol{\Sigma}_{\mathrm{red}}$, the covariance of the reduced state (Section 4.2.11). The fitted covariances satisfy equation (4.34) at every tolerance of the table ($0.69260 = 0.69250 + 0.0001$ at $\sigma_p = 0.01\,\mathrm{K}$, $0.72125 = 0.71125 + 0.01$ at $0.1\,\mathrm{K}$). A tighter tolerance than $0.001\,\mathrm{K}$ buys nothing and costs conditioning: the condition number grows a hundredfold per decade, and at $\sigma_p = 10^{-6}\,\mathrm{K}$ it reaches 4×10^{15} and round-off already moves $\sqrt{\Sigma_{33}}$ from its limit by more than $10^{-6}\,\mathrm{K}$, though by less than $10^{-3}\,\mathrm{K}$.

4.4.6 Second routes to the full-load answers

The propagated full-load temperature. As the third entry, with $\mathbf{g} = (0,\ 0,\ 1)^\mathsf{T}$, the full-load temperature has $\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g} = \Sigma_{33} = 0.69231\,\mathrm{K}^2$; as the prediction $18{+}60\rho$, with $\mathbf{g} = (60,\ 0,\ 0)^\mathsf{T}$, it has $\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g} = 60^2\,\Sigma_{11} = 3600/5200 = 0.69231\,\mathrm{K}^2$, the same number, as Section 4.3.1 requires in the limit.

The threshold through the resistance. By equation (4.58), $t^* = (28 - 28.1538)/60 = -0.0025641\,\mathrm{K/kW}$ and

$$\rho^* = \hat{z}_1 + t^* = 0.169231 - 0.002564 = 0.166667\,\mathrm{K/kW},$$
$$\Pr(\rho < \rho^* \mid \mathbf{y}) = \Phi\!\left(\frac{-0.002564}{0.013868}\right) = \Phi(-0.1849) = 0.4267,$$

the number of equation (4.61).

4.4.7 The propagation check against draws

For a linear output the delta method is exact, and the check is the algebraic agreement of the entry and the gradient already shown in Section 4.4.6. For a nonlinear output the linearized result must be compared with the full nonlinear evaluation at draws from the posterior. The procedure is the following.

1. Factor the covariance as $\boldsymbol{\Sigma} = \mathbf{L}\mathbf{L}^\mathsf{T}$ with $\mathbf{L}$ lower triangular (the Cholesky factor, which exists at any finite tolerance), and draw N_d independent states $\mathbf{z}^{(j)} = \hat{\mathbf{z}} + \mathbf{L}\boldsymbol{\varepsilon}^{(j)}$, $j = 1, \ldots, N_\mathrm{d}$, each $\boldsymbol{\varepsilon}^{(j)}$ a vector of three independent standard normal numbers, so that $\mathbf{z}^{(j)} \sim \mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$. At $\sigma_p = 0.01\,\mathrm{K}$,

$$\mathbf{L} = \begin{bmatrix} 0.013869 & 0 & 0 \\ 0.41592 & 0.0099980 & 0 \\ 0.83217 & 0 & 0.010000 \end{bmatrix},$$

whose first column is $\sqrt{\Sigma_{11}}$ times (nearly) $\mathbf{v}$ and whose other two columns have the width σ_p of the physical rows. (In the limit of exact rows the factor has the single column $\mathbf{v}/\sqrt{5200}$, and a draw is $\hat{\mathbf{z}} + \mathbf{v}\,\varepsilon^{(j)}/\sqrt{5200}$, along the free direction.)

2. Evaluate the output at each drawn state through the full model, for example $T_2^{(j)} = z_3^{(j)}$ or $G^{(j)} = 1/z_1^{(j)}$.
3. Estimate a threshold probability as the fraction of draws satisfying the event, $\hat{P} = N^{-1}\sum_j \mathbf{1}[f(\mathbf{z}^{(j)}) < f^*]$, whose standard error is $\sqrt{\hat{P}(1-\hat{P})/N}$; estimate quantiles as the empirical quantiles of the $f(\mathbf{z}^{(j)})$.
4. Compare with the closed forms of the delta method. Agreement within the standard error means the closed form can be trusted.

A few hundred draws resolve a probability near one half to about ± 0.03; $N_\mathrm{d} = 10^5$ resolves it to about ± 0.002.

Table 4.8 applies the procedure to the full-load temperature, which is linear in ρ, and to the conductance $G = 1/\rho$, which is not. The first column names the quantity; the second gives the delta-method value, from a Gaussian with mean $f(\hat{\mathbf{z}})$ and standard deviation $\sqrt{\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}}$; the third gives the exact value of the push-forward of the posterior, the limit to which the draw estimates converge, computed here in closed form because both outputs are monotone in ρ (the q-quantile of G is $1/(\hat{\rho} - z_q\sqrt{\Sigma_{11}})$ with z_q the standard normal quantile); the fourth gives one run of $N_\mathrm{d} = 10^5$ drawn states at $\sigma_p = 0.01\,\mathrm{K}$, generated with a seeded pseudo-random generator so that the run can be repeated exactly. The conductance carries the gradient of Section 4.3.6.

Table 4.8: The propagation check: delta method against the exact push-forward of the posterior, whose first entry is $\mathcal{N}(0.169231, 0.013868^2)$, and one run of 10^5 states drawn as $\hat{\mathbf{z}} + \mathbf{L}\boldsymbol{\varepsilon}$ (standard error of a probability about 0.0015). The threshold $G = 6.25\,\mathrm{kW/K}$ is the pre-test conductance $1/\mu_0$.

Quantity	Delta method	Exact	10^5 draws
T_2, median [°C]	28.154	28.154	28.152
T_2, 5 % and 95 % [°C]	26.785, 29.522	26.785, 29.522	26.786, 29.516
$\Pr(T_2 < 28°\mathrm{C})$ [–]	0.4267	0.4267	0.4269
G, median [kW/K]	5.909	5.909	5.910
G, 5 % and 95 % [kW/K]	5.113, 6.706	5.207, 6.830	5.210, 6.828
$\Pr(G > 6.25\,\mathrm{kW/K})$ [–]	0.2407	0.2528	0.2527

For the temperature the three columns agree to within the sampling error, and the closed form needs no correction. For the conductance the median is right in all three columns, since $1/\rho$ is monotone and maps the median of ρ to the median of G, but the delta-method band is shifted: the exact band $[5.207, 6.830]$ is skewed to the right of $\hat{G}$, while the linearized band $[5.113, 6.706]$ is symmetric about it, and the delta method understates the probability that the conductance exceeds its pre-test value by about 0.012. The draws reproduce the exact column to within their standard error. Figure 4.6 shows the cause. The horizontal axis is the conductance and the vertical axis its density; the solid curve is the exact density of $G = 1/\rho$, obtained by the change of variables $p_G(G) = p(\rho = 1/G \mid \mathbf{y})/G^2$ of Section 4.3.3, and the dashed curve the delta-method Gaussian. The exact density has its mode to the left of 5.909 and a longer right tail, and the Gaussian cannot represent the asymmetry.

The conductance also shows why "the posterior mean of G" must be handled with care. Under a Gaussian posterior for ρ, the mean of $1/\rho$ does not exist in the strict sense, because the Gaussian has positive density at $\rho = 0$, where $1/\rho$ is not integrable; the mass involved here is about 10^{-34} and has no practical effect, but it is a reminder that moments of a transformed quantity are properties of the parameterization. Quantiles and threshold probabilities, which

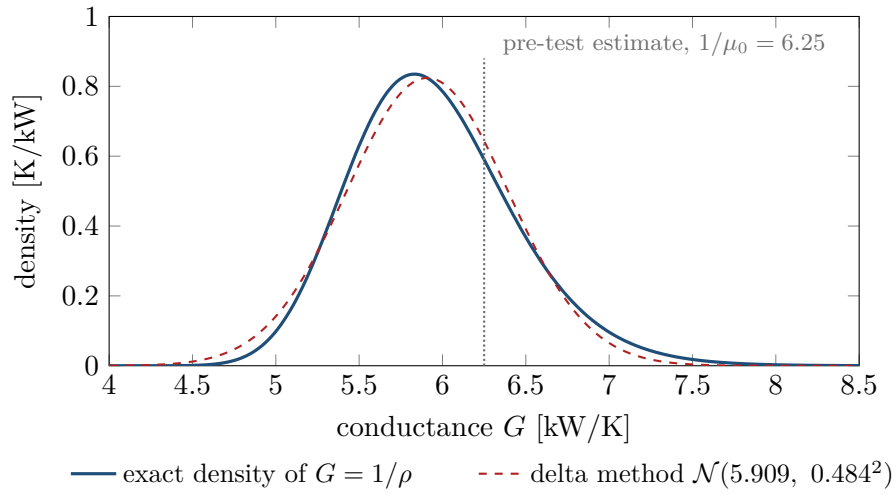


Figure 4.6: Posterior density of the conductance $G = 1/\rho$ under $\rho \mid \mathbf{y} \sim \mathcal{N}(0.169231, 0.013868^2)$: exact, by change of variables, $p_G(G) = \varphi((1/G - 0.169231)/0.013868)/(0.013868\, G^2)$ with φ the standard normal density; and the delta-method Gaussian $\mathcal{N}(1/\hat{\rho}, \Sigma_{11}/\hat{\rho}^4)$. Dotted line: the pre-test conductance.

the push-forward preserves exactly for a monotone output, are the robust summaries, and they are what the check compares.

4.4.8 Sensitivity to the value read

Table 4.9 repeats the assessment for other values of the single test reading. The first numeric column is the prior; each further column is the posterior after a different value of the reading, the actual one, 23.2°C, among them. The rows give the MAP resistance from equation (4.29), the mean and standard deviation of the full-load temperature, and the probability (4.57). The standard deviation is the same for every reading, and the mean moves along the gain vector (4.59), which with the limit covariance (4.35) is

$$\frac{\partial \hat{\mathbf{z}}}{\partial y_T} = \frac{\boldsymbol{\Sigma}\mathbf{h}}{\sigma_T^2} = 4 \begin{bmatrix} 0.0057692 \\ 0.17308 \\ 0.34615 \end{bmatrix} = \begin{bmatrix} 0.023077 \\ 0.69231 \\ 1.38462 \end{bmatrix},$$

four times the second column of the limit covariance, in K/kW per K and K per K. One column of the table, for $y_T = 22.4$°C, is the actual state moved by -0.8 K along it: $\hat{\rho} = 0.169231 - 0.8 \times 0.023077 = 0.150769$, $\mathbb{E}[T_2 \mid \mathbf{y}] = 28.1538 - 0.8 \times 1.38462 = 27.046$°C, and $\Phi((28 - 27.046)/0.8321) = \Phi(1.1464) = 0.8742$.

Table 4.9: The full-load assessment before the test and after a single test reading of various values ($\sigma_T = 0.5$ K, prior $\mathcal{N}(0.160, 0.025^2)$). The column 23.2°C is the observed reading.

Quantity	Prior	$y_T = 22.4$	22.8	23.2	23.6	24.0
$\hat{\rho}$ [K/kW]	0.16000	0.15077	0.16000	0.16923	0.17846	0.18769
$\mathbb{E}[T_2]$ [°C]	27.60	27.05	27.60	28.15	28.71	29.26
sd(T_2) [K]	1.500	0.832	0.832	0.832	0.832	0.832
$\Pr(T_2 < 28$°C) [–]	0.6051	0.8742	0.6846	0.4267	0.1975	0.0647

A reading of 22.8°C, equal to what the prior predicted, leaves the mean where it was and only narrows the distribution; because the mean then lies below the limit, narrowing alone

raises the probability from 0.605 to 0.685. The observed 23.2°C moves the mean across the limit and lowers the probability to 0.427. The posterior probability equals the prior one for a reading of 22.93°C and falls below one half for any reading above 23.09°C.

4.5 The report (step 6)

One reading at the commissioning test overturns the record's pre-test estimate. The pre-test resistance put the hall at 27.6°C at full load, under the limit, with probability 0.6051 of staying below it; the reading of 23.2°C moves the resistance up, from 0.160 to 0.169 K/kW, and leaves the probability at 0.4267, near even odds, although it narrows the full-load spread from 1.5 to 0.832 K. The answers to the questions of Table 4.2 are given in the units in which the questions were asked, conditional on the steady-state model, the test reading and the pre-test prior; the release decision and the entry in the maintenance log follow from them.

4.5.1 The answers and their bands

Table 4.10 gives the answers. Each row is a question; the second column gives the answer in the limit of exact physical rows, and the third its band: the posterior standard deviation of an estimate, with the central 95 % credible interval of the full-load temperature, and for the probability the range over the model tolerances of Table 4.7.

Table 4.10: The report for the equipment hall at 60 kW, after the commissioning test reading of 23.2°C. Answers in the limit of exact physical rows; bands are posterior standard deviations, a credible interval, or the range over model tolerances σ_p from 0 to 0.1 K.

Asset manager's question	Answer	Band
What is the thermal resistance of the hall's air path at commissioning?	0.16923 [K/kW]; conductance 5.909 [kW/K]	standard deviation 0.01387 [K/kW]; conductance 5–95 % band 5.113 to 6.706 [kW/K], exact 5.207 to 6.830 [kW/K]
What was the hall temperature during the commissioning test?	23.077 [°C]	standard deviation 0.416 [K]
What will the hall temperature be at full load?	28.15 [°C]	standard deviation 0.83 [K]; 26.52 to 29.78 [°C], central 95 % interval
Can the hall be released for 60 kW within the 28°C acceptance limit?	probability 0.427 [–]; of exceeding, 0.573 [–]	0.4267 to 0.4312 [–] over the model tolerances

The observed temperature was slightly warmer than the prior model predicted (22.8°C, the second entry of $\mathbf{z}^{(0)}$), so the inferred resistance rose above its pre-test value and the implied conductance fell below 6.25 kW/K. The inferred test temperature is not the observed 23.2°C: the reading is noisy, and the prior, carried to T_1 by the physical row, also contributes information. The conductance in the first row is $1/\hat{\rho}$, the conductance evaluated at the estimated resistance; it is not the posterior mean of the conductance (Section 4.4.7). With correlations above 0.9996 the three entries are known almost only as one combination, the free direction $\mathbf{v}$, and the full-load interval contains temperatures both below and above the limit.

4.5.2 The release decision and the entry in the maintenance log

Conditional on the steady-state model, the reading and the prior, the estimated full-load temperature lies slightly above the limit, with substantial uncertainty on both sides of it. The reading reduced the uncertainty and, at the same time, lowered the probability of staying

within the acceptance limit, from 0.605 before it to 0.427. The evidence does not support a high-confidence claim that the release keeps the hall within the limit: the posterior probability of doing so is about 0.43. A release for full load would require an assurance level κ, stated explicitly, and a consideration of the consequences of exceeding the limit. Further measurements or more cooling capacity may be appropriate, but their value should be computed, not assumed. That is what an inference report gives the asset manager: the physical interpretation of a mathematical result, its uncertainty, and its implications for a decision.

The acceptance is the first entry of the hall's maintenance log. The resistance it reads on the pre-test prior answers the release question only; the commissioning baseline stays Chapter 1's $\hat{g}_0 = -0.0139$, the fit of the same test on the handbook prior. In the values of Table 4.10, the entry reads:

> *Window 0, acceptance against the record's* 28°C *criterion.*
> Release for 60 kW not supported, $P = 0.427$ (thermal resistance of the air path 0.16923 ± 0.01387 K/kW).
> Next measurement: one supervised reading at full load.

Table 4.9 shows that the direction of the change was set by the value read, not by the act of reading. New evidence does not necessarily make the plant look safer: it can make it look less safe, or move the probability in either direction, depending on the value observed. The purpose of inference is to improve the knowledge of the plant, not to produce reassuring probabilities.

Principle 4.2. *A measurement narrows the posterior by an amount fixed before it is taken, but the direction in which it moves a condition or acceptance probability is set by the value it returns. Evidence improves knowledge; it does not promise reassurance.*

4.6 The next measurement (step 7)

The report leaves the acceptance question near even odds, with a band that a single reading at the test condition has set. A further measurement is one more row of the residual vector, and the information it adds is known before its value is.

4.6.1 The preposterior covariance

A further reading at the test condition is the row $(T_1 - y)/\sigma_T$, whose line of the Jacobian is $\mathbf{j} = (0,\ 2,\ 0)$, the line of the first reading; a reading at full load, by a sensor of the same spread, is the row $(T_2 - y)/\sigma_T$, with line $\mathbf{j}_\mathrm{f} = (0,\ 0,\ 2)$. The information that a reading adds, $\mathbf{j}^\mathsf{T}\mathbf{j}$, is known before the reading is taken, because its Jacobian depends on the sensor and on where it is placed, not on the value. For a candidate measurement a with information $\mathbf{\Lambda}_a$ the preposterior covariance is

$$\mathbf{\Sigma}_a = \left(\mathbf{\Sigma}^{-1} + \mathbf{\Lambda}_a\right)^{-1}, \tag{4.65}$$

formed at a finite tolerance, where $\mathbf{\Sigma}^{-1} = \mathbf{J}^\mathsf{T}\mathbf{J}$ exists, the object on which candidate measurements are ranked. In the limit of exact rows the posterior stays on the free direction, and a candidate adds to the information $\|\mathbf{J}\mathbf{v}\|^2 = 5200$ along it the square of its own line times $\mathbf{v}$: $(\mathbf{j}\mathbf{v})^2 = 60^2 = 3600$ for a reading at 30 kW and $(\mathbf{j}_\mathrm{f}\mathbf{v})^2 = 120^2 = 14\,400$ for a reading at 60 kW. The full-load standard deviation after the candidate is 60 divided by the square root of the total. For n further readings at 30 kW, $\mathbf{\Lambda}_a = n\,\mathbf{j}^\mathsf{T}\mathbf{j}$, which adds $4n$ to the (T_1, T_1) entry, and the third diagonal entry of $\mathbf{\Sigma}_a$ gives the standard deviation of the full-load temperature: 0.640 K for $n = 1$, 0.539 K for $n = 2$, 0.429 K for $n = 4$ and 0.325 K for $n = 8$. More readings resolve the question; they do not decide its answer. If the full-load temperature truly lies above 28°C, the probability tends to zero as the readings accumulate, and the decision then rests on the air path, not on the measurement.

4.6.2 The candidate measurements

Table 4.11 ranks three candidates. Each row is a measurement the asset manager can order; the columns give the information along $\mathbf{v}$ after it, the standard deviations of the resistance and of the full-load temperature that the preposterior covariance (4.65) predicts, and the width it removes from the current full-load standard deviation of 0.8321 K. One reading at full load carries four times the information of one at the test condition, because the full-load temperature moves twice as far as the test temperature along $\mathbf{v}$, and it removes exactly the width of four test readings.

Table 4.11: Candidate measurements for the hall, in the limit of exact physical rows, from the posterior after the test reading. Each reading has $\sigma_T = 0.5\,\mathrm{K}$; the width removed is taken from the current $\sqrt{\Sigma_{33}} = 0.8321\,\mathrm{K}$.

Candidate	Information along $\mathbf{v}$ [(kW/K)^2]	sd(ρ) after [K/kW]	sd(T_2) after [K]	Width removed [K]
One further reading at 30 kW	8800	0.01066	0.640	0.192
Four further readings at 30 kW	19 600	0.00714	0.429	0.403
One reading at 60 kW	19 600	0.00714	0.429	0.403

4.6.3 The recommendation

Four further readings at the test condition and one reading at full load narrow the answer equally, and the choice between them rests on what else each tests. Readings at 30 kW need no operation at the load of the release, but repeated readings of one sensor share its bias, and no number of them can show whether the resistance at 60 kW is the resistance at 30 kW. A reading at full load tests exactly that assumption, on which the prediction rests (Section 4.6.7), at the cost of running the hall at the load in question. The next measurement is therefore one reading at full load, taken in a supervised run, the load raised in steps and held at 60 kW until the hall settles, which can be stopped if the hall approaches its limit; where such a run is not acceptable, it is four further readings at the test condition, by more than one sensor where possible.

4.6.4 A second observation as an update

Before any run at full load, the first candidate of Table 4.11 arrives: a second, independent reading at the test condition, $y_{T,2} = 23.0^\circ\mathrm{C}$, with the same standard deviation $\sigma_T = 0.5\,\mathrm{K}$; the first reading is now written $y_{T,1} = 23.2^\circ\mathrm{C}$. Against the posterior predictive (4.64) it lies $(23.0 - 23.0769)/0.6504 = -0.12$ standard deviations from the predicted mean and passes that check. The posterior after the first reading serves as the prior for the second.

The precision form. Information from independent rows adds. The new reading is one more row of the residual vector, $(T_1 - y_{T,2})/\sigma_T$, whose line of the Jacobian is $\mathbf{j} = (0,\ 2,\ 0)$, the same as the first reading's. Its contribution to the information matrix is the outer product $\mathbf{j}^\mathsf{T}\mathbf{j}$, which changes the (T_1, T_1) entry alone,

$$\boldsymbol{\Lambda}_2 = \mathbf{J}^\mathsf{T}\mathbf{J} + \mathbf{j}^\mathsf{T}\mathbf{j} = \mathbf{J}^\mathsf{T}\mathbf{J} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad \boldsymbol{\Sigma}_2 = \boldsymbol{\Lambda}_2^{-1}.$$

The old rows are balanced at $\hat{\mathbf{z}}$, $\mathbf{J}^\mathsf{T}\mathbf{r}(\hat{\mathbf{z}}) = \mathbf{0}$, so the gradient of the enlarged objective at $\hat{\mathbf{z}}$ comes from the new row alone, $\mathbf{j}^\mathsf{T}(23.0769 - 23.0)/0.5 = (0,\ 0.30769,\ 0)^\mathsf{T}$, and one Gauss–Newton

step (4.29) from $\hat{\mathbf{z}}$ gives $\hat{\mathbf{z}}_2 = \hat{\mathbf{z}} - 0.30769\,(\boldsymbol{\Sigma}_2)_{\cdot 2}$, with $(\boldsymbol{\Sigma}_2)_{\cdot 2}$ the second column of $\boldsymbol{\Sigma}_2$. In the limit of exact rows this column is $(3.4091 \times 10^{-3},\ 0.10227,\ 0.20455)^{\mathsf{T}}$, and

$$\hat{\mathbf{z}}_2 = \begin{bmatrix} 0.169231 \\ 23.0769 \\ 28.1538 \end{bmatrix} - 0.30769 \begin{bmatrix} 0.0034091 \\ 0.10227 \\ 0.20455 \end{bmatrix} = \begin{bmatrix} 0.168182 \\ 23.0455 \\ 28.0909 \end{bmatrix}.$$

The first entry of $\boldsymbol{\Sigma}_2$ is $1.1364 \times 10^{-4}\,(\mathrm{K/kW})^2$, so that $\sqrt{(\boldsymbol{\Sigma}_2)_{11}} = 0.010660\,\mathrm{K/kW}$.

The rank-one form. The same result follows from the first posterior alone, without forming and inverting $\boldsymbol{\Lambda}_2$. The second reading measures $\mathbf{h}^{\mathsf{T}}\mathbf{z} = T_1$ with $\mathbf{h} = (0,\ 1,\ 0)^{\mathsf{T}}$; its predicted value under the first posterior is $\mathbf{h}^{\mathsf{T}}\hat{\mathbf{z}} = 23.0769\,^{\circ}\mathrm{C}$ and its predictive variance, from equation (4.64), is $S_{\mathrm{in}} = \mathbf{h}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{h} + \sigma_T^2 = \Sigma_{22} + \sigma_T^2 = 0.42308\,\mathrm{K}^2$. Conditioning the Gaussian on the new row gives the update

$$\hat{\mathbf{z}}_2 = \hat{\mathbf{z}} + \boldsymbol{\Upsilon}\,(y_{T,2} - \mathbf{h}^{\mathsf{T}}\hat{\mathbf{z}}), \qquad \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma} - S_{\mathrm{in}}\,\boldsymbol{\Upsilon}\boldsymbol{\Upsilon}^{\mathsf{T}}, \qquad \boldsymbol{\Upsilon} = \frac{\boldsymbol{\Sigma}\mathbf{h}}{S}, \tag{4.66}$$

where $\boldsymbol{\Upsilon}$ is the gain vector, one entry per state entry, and $y_{T,2} - \mathbf{h}^{\mathsf{T}}\hat{\mathbf{z}}$ the innovation, the part of the reading the first posterior did not predict. Because $\boldsymbol{\Sigma}_2\mathbf{h} = \boldsymbol{\Sigma}\mathbf{h} - \boldsymbol{\Upsilon}\,\mathbf{h}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{h} = (\sigma_T^2/S_{\mathrm{in}})\,\boldsymbol{\Sigma}\mathbf{h}$, the gain is also $\boldsymbol{\Upsilon} = \boldsymbol{\Sigma}_2\mathbf{h}/\sigma_T^2$, the same vector as the gain (4.59) of the posterior that includes the new reading. The gain follows from writing the posterior mean as $\hat{\mathbf{z}} + \mathrm{Cov}(\mathbf{z}, Y_{\mathrm{new}})\,(y_{T,2} - \mathbb{E}[Y_{\mathrm{new}}])/S_{\mathrm{in}}$ for the jointly Gaussian state and new reading, with $\mathrm{Cov}(\mathbf{z}, Y_{\mathrm{new}}) = \boldsymbol{\Sigma}\mathbf{h}$, the second column of $\boldsymbol{\Sigma}$. With the limit covariance (4.35),

$$\boldsymbol{\Upsilon} = \frac{1}{0.42308} \begin{bmatrix} 0.0057692 \\ 0.17308 \\ 0.34615 \end{bmatrix} = \begin{bmatrix} 0.013636 \\ 0.40909 \\ 0.81818 \end{bmatrix},$$

in kW^{-1} for the first entry and dimensionless for the other two; the innovation is $23.0 - 23.0769 = -0.0769\,\mathrm{K}$, and $\hat{\mathbf{z}}_2 = \hat{\mathbf{z}} - 0.0769\,\boldsymbol{\Upsilon} = (0.168182,\ 23.0455,\ 28.0909)^{\mathsf{T}}$, the state of the precision form. The covariance update subtracts the outer product $S_{\mathrm{in}}\boldsymbol{\Upsilon}\boldsymbol{\Upsilon}^{\mathsf{T}}$; on its first and last diagonal entries,

$$(\boldsymbol{\Sigma}_2)_{11} = 1.9231 \times 10^{-4} - 0.42308 \times 0.013636^2 = 1.1364 \times 10^{-4}\,(\mathrm{K/kW})^2,$$
$$(\boldsymbol{\Sigma}_2)_{33} = 0.69231 - 0.42308 \times 0.81818^2 = 0.40909\,\mathrm{K}^2,$$

the values of the precision form. Because $S_{\mathrm{in}}\,\boldsymbol{\Upsilon}^{\mathsf{T}} = \mathbf{h}^{\mathsf{T}}\boldsymbol{\Sigma}$, the covariance update is also

$$\boldsymbol{\Sigma}_2 = (\mathbf{I} - \boldsymbol{\Upsilon}\mathbf{h}^{\mathsf{T}})\,\boldsymbol{\Sigma},$$

with $\mathbf{I}$ the identity matrix: each column of $\boldsymbol{\Sigma}$ loses the gain vector times that column's second entry. The correction is the outer product of one vector, which is the reason for the name rank-one update. For a linear-Gaussian model it is exact, and one more independent reading costs one product of $\boldsymbol{\Sigma}$ with a vector rather than a new inference. For a strongly nonlinear model, or after a shift of the estimate large enough to make the Jacobian at the old estimate stale, the model is linearized again or the full fit is repeated.

The updated assessment. Table 4.12 follows the assessment through the two readings. Each column is a state of knowledge, before any reading, after the first, after both; the rows are the information about the resistance, the reciprocal of the first diagonal entry of the covariance, the MAP resistance, the mean and standard deviation of the full-load temperature, the threshold probability, and the 95 % credible interval.

Table 4.12: The full-load assessment before the test, after the first reading (23.2°C) and after the second (23.0°C).

Quantity	Prior	One reading	Two readings
$1/\Sigma_{11}$ [(kW/K)^2]	1600	5200	8800
$\hat{\rho}$ [K/kW]	0.160000	0.169231	0.168182
$\mathbb{E}[T_2]$ [°C]	27.600	28.154	28.091
sd(T_2) [K]	1.500	0.832	0.640
Pr($T_2 < 28$°C) [–]	0.6051	0.4267	0.4435
95 % interval [°C]	24.66, 30.54	26.52, 29.78	26.84, 29.34

After the second reading the full-load temperature is the third entry of $\hat{\mathbf{z}}_2$, 28.091°C, with standard deviation $\sqrt{(\boldsymbol{\Sigma}_2)_{33}} = \sqrt{0.40909} = 0.640\,\text{K}$, and the probability of staying below the limit becomes $\Phi(-0.1421) = 0.4435$. The interval narrowed by a quarter, yet the probability moved only from 0.427 to 0.443, because two effects acted in opposite directions. The narrowing alone, with the mean held at 28.154°C, would have lowered the probability to 0.405, since the mean lies above the limit; the shift of the mean alone, with the spread held at 0.832 K, would have raised it to 0.456. The effect of actual new evidence cannot be described by its reduction of the covariance alone: the value observed matters.

4.6.5 A later window and the trend rows

A second reading in the same window is one more row on the same entry, because the air path it reads is the path the first reading read. A later condition test is the next measurement in time, and it cannot be folded into the same entry: by then the path may have moved.

The later condition test. Two years after commissioning, in window 24 (Dec Y2), the asset manager repeats the commissioning test as a condition test: the same load of 30 kW, the supply air at 18°C, the same hall sensor. It reads 23.5°C. Subscripted now by window, the two readings are $y_{T,0} = 23.2$°C and $y_{T,24} = 23.5$°C. The trend chart takes each test temperature as exact and plots the conductance $Q_1/(y_T - T_c)$: from $30/5.2 = 5.7692$ kW/K at commissioning to $30/5.5 = 5.4545$ kW/K two years later, a ratio of 0.9455 and a loss of 2.77 %/yr, with no band. What the chart leaves open is the asset manager's question: *how fast is the air path losing conductance, and is it losing any at all?* A loss is a negative trend, so its general form is the first line of equation (2.21) with the sign of the trend entry reversed,

$$\Pr(\gamma < 0 \mid \mathbf{y}) = \int \mathbf{1}[\gamma < 0]\; p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}, \tag{4.67}$$

with γ the trend entry of a state that carries the path's condition in every window.

The choices of the window state. Five choices enter here. The conductance is carried as Chapter 1's log-conductance $g = \ln(G/G_{\text{ref}})$, not as the resistance ρ: this keeps the conductance positive, makes the conductance at a later window against the commissioning baseline a difference of entries, $G_{24}/G_0 = \mathrm{e}^{g_{24}-g_0}$, and makes the trend entry a fractional loss per year; it costs the exactly Gaussian posterior of the resistance, because the rows become nonlinear in the state, which brings in the Laplace approximation (4.38) and its curvature term. Each window's hall temperature is the function of its g_t that the exact physical row gives, on the reduced state of Section 4.2.8: $T_t = T_c + (Q_1/G_{\text{ref}})\,\mathrm{e}^{-g_t} = (18 + 5\,\mathrm{e}^{-g_t})$°C, a nonlinear physical row evaluated point by point as that section's paragraph on nonlinear physical rows describes; this keeps one entry per window, and it breaks if the physical row holds only up to a finite

discrepancy σ_p, which the checks below set against the full state. The handbook prior of Chapter 1, $g_0 \sim \mathcal{N}(\ln(7.5/6) = 0.2231,\ 0.3^2)$, is a row on g_0 alone: it describes the path as built, and it would mislead as a statement about a later window, which the path may have left; on g this prior takes the place of the pre-test resistance of the main example. The trend entry has no prior row: the two tests alone say what it is, and with nothing bounding g_{24} from above the posterior is improper in a far tail, whose mass the checks below measure; a prior on it from the coils of this design is the partial pooling of Chapter 9. Both tests run at the same load and supply air: the raw test temperature and g then carry the same information, and they part when the inputs move with the season, which Chapter 13 takes up.

The trend rows. The windows are months, indexed t, with window 0 the commissioning test and $\Delta t = 1/12\,\mathrm{yr}$ their spacing. The state carries one log-conductance per window and the trend entry γ, in yr^{-1}, and a trend row ties each window's entry to the next:

$$\mathbf{z} = \begin{bmatrix} g_0 \\ g_1 \\ \vdots \\ g_{24} \\ \gamma \end{bmatrix}, \qquad r_{q,t}(\mathbf{z}) = \frac{g_{t+1} - g_t - \gamma\,\Delta t}{\sigma_q}, \quad t = 0, \ldots, 23. \tag{4.68}$$

The spread σ_q is that of one window's departure from the line, the wander of the path from one month to the next; it is dimensionless, per monthly window, and set here at 0.005. Each trend row has the Jacobian line $(\ldots,\ -1/\sigma_q,\ +1/\sigma_q,\ \ldots,\ -\Delta t/\sigma_q)$, its two window entries in the columns of g_t and g_{t+1} and the last in the column of γ. The rows are linear, so they add nothing to the curvature term $\mathbf{C}$ of equation (4.44). The record has readings at windows 0 and 24 only, so the state has 26 entries and its Jacobian is 27×26: two reading rows, one prior row and 24 trend rows. Figure 4.7 draws the graph across windows. The top row is the chain of window entries, with the 23 windows that have no reading drawn as one dashed box and the trend rows as black squares between neighbors; the node above is the trend entry, linked to every trend row; the lower row holds the two tests, each a hall node tied to its window's entry by the physical row, with the load and the supply air pinned and the reading attached.

The unread windows integrated out. Windows 1 to 23 carry no reading, and their entries appear in trend rows only. Write $e_t = \sigma_q\, r_{q,t}$ for the departure of window t; under the rows the 24 departures are independent with variance σ_q^2 each, and their sum telescopes, $\sum_{t=0}^{23} e_t = g_{24} - g_0 - 24\,\Delta t\,\gamma$. Integrating the 23 unread entries out of the Gaussian product of the trend rows is the Schur complement of their block (Section 3.5.3); because the rows are linear and those entries appear in no other row, what remains is the Gaussian of the sum, a single row of variance $24\,\sigma_q^2$,

$$r_q(\mathbf{z}) = \frac{g_{24} - g_0 - 24\,\Delta t\,\gamma}{\sigma_q\sqrt{24}}, \qquad \sigma_q\sqrt{24} = 0.024495, \quad 24\,\Delta t = 2\,\mathrm{yr}, \tag{4.69}$$

on the equivalent state $\mathbf{z} = (g_0,\ g_{24},\ \gamma)^{\mathsf{T}}$. The 26-entry fit and the three-entry fit agree in $\hat{\gamma}$, in its standard deviation and in $\hat{g}_{24}$ to less than 10^{-8}. This is the form a record with gaps uses: a run of n_{w} windows without a reading is one row of spread $\sigma_q\sqrt{n_{\mathrm{w}}}$.

The rows, the Jacobian and its pattern. On the equivalent state the rows are the two readings, the prior and the trend row,

$$\mathbf{r}(\mathbf{z}) = \begin{bmatrix} (T_c + (Q_1/G_{\mathrm{ref}})\,\mathrm{e}^{-g_0} - y_{T,0})/\sigma_T \\ (T_c + (Q_1/G_{\mathrm{ref}})\,\mathrm{e}^{-g_{24}} - y_{T,24})/\sigma_T \\ (g_0 - \mu_g)/\sigma_g \\ (g_{24} - g_0 - 24\,\Delta t\,\gamma)/(\sigma_q\sqrt{24}) \end{bmatrix} \begin{matrix} \text{reading, window 0} \\ \text{reading, window 24} \\ \text{prior} \\ \text{trend} \end{matrix}, \tag{4.70}$$

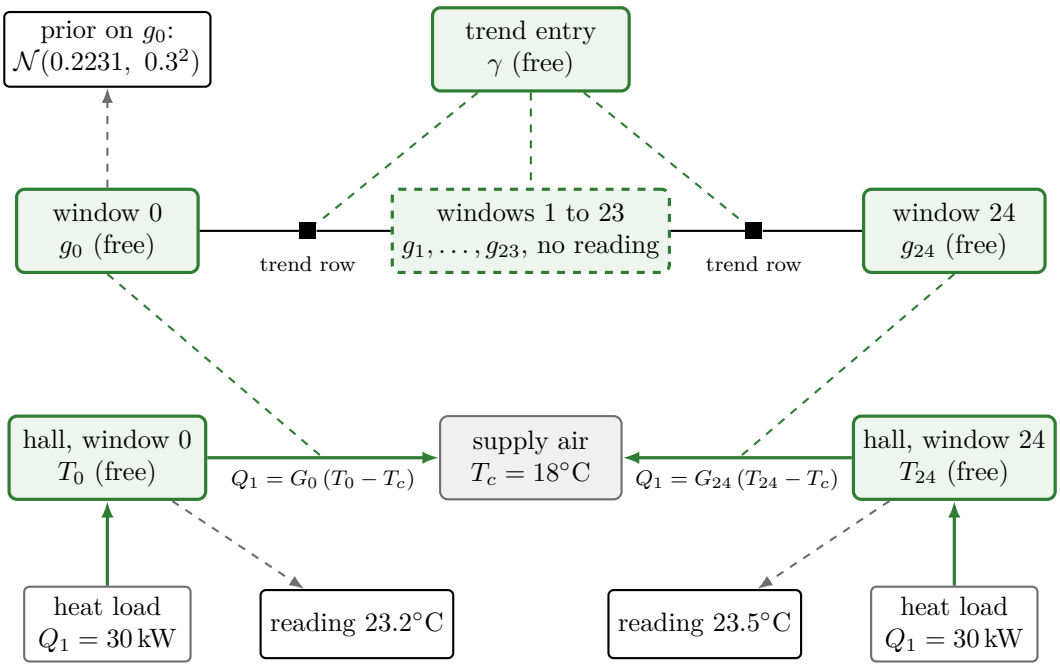


Figure 4.7: The plant graph of the hall across windows. Free (green): the log-conductances g_0 and g_{24} of the two tested windows, the entries g_1 to g_{23} of the windows without a reading (one dashed box), the trend entry γ, and the hall temperatures T_0 and T_{24}. Black squares: trend rows, each tying a window's entry to the next; dashed green links tie γ to every trend row and each g_t to its window's physical row, with $G_t = G_{\mathrm{ref}}\,\mathrm{e}^{g_t}$. Pinned: the supply air T_c and the test load Q_1 (gray and white). Observed: T_0 and T_{24}, by one reading each (dashed arrows); g_0 carries the handbook prior, and no other entry has a prior.

and they act on the entries in the pattern

$$\mathbf{J} = \begin{bmatrix} * & 0 & 0 \\ 0 & * & 0 \\ * & 0 & 0 \\ * & * & * \end{bmatrix}, \qquad \text{columns } g_0, g_{24}, \gamma.$$

The third column has one mark: the trend entry is seen by the trend row and by nothing else. The first two columns share only that row, so the trend row is the one path from one window to the other. At the MAP state, with $\ell_t = (Q_1/G_{\text{ref}})\,\mathrm{e}^{-\hat{g}_t}/\sigma_T$ the magnitude of the slope of window t's reading row, the 4×3 Jacobian, its first two columns dimensionless and the third in yr, is

$$\mathbf{J} = \begin{bmatrix} -\ell_0 & 0 & 0 \\ 0 & -\ell_{24} & 0 \\ 1/\sigma_g & 0 & 0 \\ -1/(\sigma_q\sqrt{24}) & 1/(\sigma_q\sqrt{24}) & -(2\,\text{yr})/(\sigma_q\sqrt{24}) \end{bmatrix} = \begin{bmatrix} -10.1402 & 0 & 0 \\ 0 & -11.0000 & 0 \\ 3.3333 & 0 & 0 \\ -40.8248 & 40.8248 & -81.6497 \end{bmatrix}, \tag{4.71}$$

and the information is

$$\mathbf{J}^{\mathsf{T}}\mathbf{J} = \begin{bmatrix} 1780.6021 & -1666.6667 & 3333.3333 \\ -1666.6667 & 1787.6667 & -3333.3333 \\ 3333.3333 & -3333.3333 & 6666.6667 \end{bmatrix}.$$

The fit. The Gauss–Newton iteration (4.29) starts at the as-built state, both windows at the handbook value and no aging: $\mathbf{z}^{(0)} = (0.2231,\ 0.2231,\ 0)^{\mathsf{T}}$. Table 4.13 lists the iterates: each line gives the three entries, the objective and the largest entry of the gradient (4.28). The objective falls from 7.380000 to 0.570211, 0.346140 and 0.345972, where it stays. The gradient falls from 3.07×10^{1} to below 10^{-8} at iterate 7, and from iterate 3 on by a constant factor of about 0.023 per iterate: Gauss–Newton converges linearly here, not in one step, because the window-0 row keeps a residual at the minimum and the curvature term the step leaves out is not zero. The factor is close to that term's share of the curvature of window 0's own fit, $2.6341/111.3013 = 0.0237$.

Table 4.13: Gauss–Newton iterations for the window state (g_0, g_{24}, γ) from the as-built state, with $\sigma_q = 0.005$ per window.

Iterate l	g_0 [–]	g_{24} [–]	γ [yr^{-1}]	$\mathcal{L}$ [–]	$\max_i \lvert(\mathbf{J}^{\mathsf{T}}\mathbf{r})_i\rvert$ [– or yr]
0	0.223144	0.223144	0.000000	7.380000	3.07×10^{1}
1	−0.032478	−0.151856	−0.059689	0.570211	7.74
2	−0.014499	−0.096879	−0.041190	0.346140	2.01×10^{-1}
3	−0.013939	−0.095311	−0.040686	0.345972	1.50×10^{-3}
4	−0.013926	−0.095310	−0.040692	0.345972	3.45×10^{-5}
5	−0.013926	−0.095310	−0.040692	0.345972	7.97×10^{-7}
6	−0.013926	−0.095310	−0.040692	0.345972	1.85×10^{-8}
7	−0.013926	−0.095310	−0.040692	0.345972	$< 10^{-8}$

The fit returns $\hat{g}_0 = -0.013926$ ($G = 5.9170\,\text{kW/K}$), $\hat{g}_{24} = -0.095310$ ($G = 5.4545\,\text{kW/K}$) and $\hat{\gamma} = -0.040692\,\text{yr}^{-1}$. The window-0 entry is Chapter 1's commissioned value on its handbook prior, close to the 5.909 kW/K that the pre-test resistance gave the main example, the two priors being different. At the MAP state the window-0 reading's residual is −0.259768 and the

prior's -0.790231; the window-24 reading and the trend row are below 10^{-8}. With no prior on the trend entry nothing resists those two rows, and the later reading is fitted to round-off: $\hat{g}_{24} = \ln(5/5.5)$, its reading-only value. The objective at the minimum is 0.345972.

The curvature term. Only the reading rows are nonlinear, each in its own window's entry, with second derivative $(Q_1/G_{\mathrm{ref}})\,\mathrm{e}^{-g_t}/\sigma_T$, the first Jacobian entry of its line with the sign changed. The curvature term of equation (4.44) is therefore

$$\mathbf{C} = \sum_k r_k \frac{\partial^2 r_k}{\partial \mathbf{z}\,\partial \mathbf{z}^{\mathsf{T}}} = \mathrm{diag}(-0.259768 \times 10.1402,\ 0,\ 0) = \mathrm{diag}(-2.6341,\ 0,\ 0),$$

the window-24 term vanishing with its residual. It is the curvature term of Chapter 1's commissioning fit, and it is kept: the covariance is $\mathbf{\Sigma} = \mathbf{H}^{-1} = (\mathbf{J}^{\mathsf{T}}\mathbf{J} + \mathbf{C})^{-1}$, the Laplace covariance (4.38), with the Gauss–Newton form $\mathbf{\Sigma}_{\mathrm{GN}} = (\mathbf{J}^{\mathsf{T}}\mathbf{J})^{-1}$ beside it.

The covariance read. With rows and columns ordered (g_0, g_{24}, γ),

$$\mathbf{\Sigma} = \begin{bmatrix} 8.984619 \times 10^{-3} & 0 & -4.492310 \times 10^{-3} \\ 0 & 8.264463 \times 10^{-3} & 4.132231 \times 10^{-3} \\ -4.492310 \times 10^{-3} & 4.132231 \times 10^{-3} & 4.462271 \times 10^{-3} \end{bmatrix}, \tag{4.72}$$

with the third row and column in yr^{-1} and the corner in yr^{-2}. The standard deviations are 0.094787, 0.090909 and $0.066800\,\mathrm{yr}^{-1}$; $\mathbf{\Sigma}_{\mathrm{GN}}$ gives $0.066410\,\mathrm{yr}^{-1}$ for the trend entry. The zero in the corner Σ_{12} says that the two windows are independent a posteriori, and the pattern says why: the trend entry's column has the trend row alone, so whatever the two windows' entries, a value of γ zeros that row, and with no prior on γ the row puts no constraint on $g_{24} - g_0$. The window-0 entry and its spread are then those of window 0 alone, Chapter 1's posterior of g: -0.013926 with standard deviation 0.094787 on the reduced state (Chapter 1's full state gives -0.013916 at $\sigma_p = 0.01\,\mathrm{K}$). The later test does not revise the commissioning baseline. The correlations read by sign: -0.7095 between γ and g_0, since a higher baseline, with the later window held, means a faster loss; $+0.6805$ between γ and g_{24}, since a higher later value means a slower one.

The variance of the trend entry. Because the trend row is the only row on γ, the trend entry is that row solved for it, $\gamma = (g_{24} - g_0 - \varepsilon_q)/(24\,\Delta t)$, with $\varepsilon_q = \sum_t e_t$ the summed departure, independent of both windows. Its gradient with respect to $(g_0, g_{24}, \varepsilon_q)$ is $\mathbf{g} = (-1,\ 1,\ -1)^{\mathsf{T}}/(2\,\mathrm{yr})$, and the three have the diagonal covariance $\mathbf{D} = \mathrm{diag}(\Sigma_{11}, \Sigma_{22}, 24\,\sigma_q^2)$, so

$$\begin{aligned} \mathrm{Var}(\gamma \mid \mathbf{y}) = \mathbf{g}^{\mathsf{T}}\mathbf{D}\,\mathbf{g} &= \frac{\mathrm{Var}(g_0 \mid \mathbf{y}) + \mathrm{Var}(g_{24} \mid \mathbf{y}) + 24\,\sigma_q^2}{(2\,\mathrm{yr})^2} \\ &= \frac{8.984619 + 8.264463 + 0.600000}{4} \times 10^{-3}\,\mathrm{yr}^{-2} = 4.462271 \times 10^{-3}\,\mathrm{yr}^{-2}, \end{aligned} \tag{4.73}$$

the corner of $\mathbf{\Sigma}$. The wander contributes 3.36 % of it: the readings' errors, not the wander, set the band.

The straight line as a limit. The exact limit $\sigma_q \to 0$ of the trend row is a constraint, $g_{24} - g_0 - 24\,\Delta t\,\gamma = 0$, and its reduced state follows from Section 4.2.8. The row's line over (g_0, g_{24}, γ) is $\mathbf{G} = (-1,\ 1,\ -2)/(\sigma_q\sqrt{24})$, with the 2 in yr; with (g_0, γ) free and g_{24} dependent, $\mathbf{G}_{\mathrm{dep}} = 1/(\sigma_q\sqrt{24})$ is invertible and equation (4.45) gives $\mathbf{N}_{\mathrm{b}} = (1,\ 2)$, so

$$g_{24} = g_0 + (2\,\mathrm{yr})\,\gamma, \tag{4.74}$$

Chapter 2's straight line (2.19) through the two windows. The line state (g_0, γ) gives $\hat{\gamma} = -0.040692\,\mathrm{yr}^{-1}$ with standard deviation 0.065668 ($\boldsymbol{\Sigma}_{\mathrm{GN}}$: 0.065271), and a correlation of -0.7217 between g_0 and γ. Centered at window 12, the level of equation (2.19) is $\alpha = g_0 + (1\,\mathrm{yr})\,\gamma = -0.054618$, with standard deviation 0.065668 and a correlation of -0.0418 with γ. The two standard deviations are equal because with two windows α and $(1\,\mathrm{yr})\,\gamma$ are half the sum and half the difference of two independent entries.

Table 4.14 repeats the fit as σ_q shrinks. Each column is one spread, the last the straight line; the rows give the trend entry and its standard deviation, the window-24 entry and its standard deviation, and the probability of a loss. The trend entry and the window-24 entry do not move: with two windows, σ_q moves the band and not the mean. Once three windows share the trend entry it moves the mean too (Exercise 4.7), which is the case of Chapter 10.

Table 4.14: The window state as the spread σ_q of the trend row shrinks. The last column is the straight line, the limit $\sigma_q \to 0$.

Quantity	$\sigma_q = 0.02$	0.01	0.005	0.001	$\sigma_q \to 0$
$\hat{\gamma}$ [yr^{-1}]	-0.040692	-0.040692	-0.040692	-0.040692	-0.040692
$\mathrm{sd}(\gamma \mid \mathbf{y})$ [yr^{-1}]	0.081928	0.070088	0.066800	0.065714	0.065668
$\hat{g}_{24}$ [–]	-0.095310	-0.095310	-0.095310	-0.095310	-0.095310
$\mathrm{sd}(g_{24} \mid \mathbf{y})$ [–]	0.090909	0.090909	0.090909	0.090909	0.090909
$\Pr(\gamma < 0 \mid \mathbf{y})$ [–]	0.6903	0.7192	0.7288	0.7321	0.7323

Figure 4.8 draws the path of the condition. The horizontal axis is the window and the vertical axis the log-conductance. The shaded band is the line's mean with one standard deviation either side, in the limit $\sigma_q \to 0$, and the dashed curves are the band of the monthly chain at $\sigma_q = 0.02$, the largest spread of Table 4.14, drawn there because at the reported 0.005 it would lie on the shaded edge; the dots with bars are the two window posteriors, the open circles the reading-only values $\ln(5/5.2) = -0.039221$ and $\ln(5/5.5) = -0.095310$, and the dotted line the stated truth of the constructed test. The dashed band is wider than the line's by the bridge term $\sigma_q^2\, t(24-t)/24$ in variance, the variance of the departures summed up to window t given their total. It vanishes at the two tested windows and is largest at window 12: 2.400×10^{-3} at $\sigma_q = 0.02$, which widens the standard deviation there from 0.065668 to 0.081928, and 1.500×10^{-4} at the reported 0.005. It collapses onto the line's band as $\sigma_q \to 0$. The second open circle lies on the window-24 posterior, which the fit placed on the reading-only value; the first lies below the window-0 posterior, because the handbook prior pulls that entry up.

The trend entry's posterior. Equation (4.67) with $p(\mathbf{z} \mid \mathbf{y})$ the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ of the window state at $\sigma_q = 0.005$ and $\gamma = z_3$ is the normal tail of the third entry, with mean $\hat{z}_3 = \hat{\gamma}$ and variance Σ_{33} of equation (4.72):

$$\Pr(\gamma < 0 \mid \mathbf{y}) \approx \Phi\!\left(\frac{-\hat{\gamma}}{\sqrt{\Sigma_{33}}}\right) = \Phi\!\left(\frac{0.040692}{0.066800}\right) = \Phi(0.6092) = 0.7288, \tag{4.75}$$

with 0.7258 beside it from the exact posterior by quadrature (Table 4.15). The trend entry is $\hat{\gamma} = -0.040692\,\mathrm{yr}^{-1}$, with standard deviation 0.066800 and 5–95 % band $[-0.150569,\ 0.069184]\,\mathrm{yr}^{-1}$. In conductance terms, through the monotone map $100(\mathrm{e}^{\gamma} - 1)$, which carries quantiles over exactly, the median loss is 3.99 %/yr and the band runs from a loss of 13.98 %/yr to a gain of 7.16 %/yr, 21.14 points wide. The conductance at window 24 against the commissioning baseline, $G_{24}/G_0 = \mathrm{e}^{g_{24} - g_0}$, is read through the gradient $\mathbf{g} = (-1,\ 1,\ 0)^{\mathsf{T}}$ of its logarithm: $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g} = 1.724908 \times 10^{-2}$, a standard deviation of 0.131336, and through the exponential a median of 0.9218 with 5–95 % band [0.7427, 1.1441].

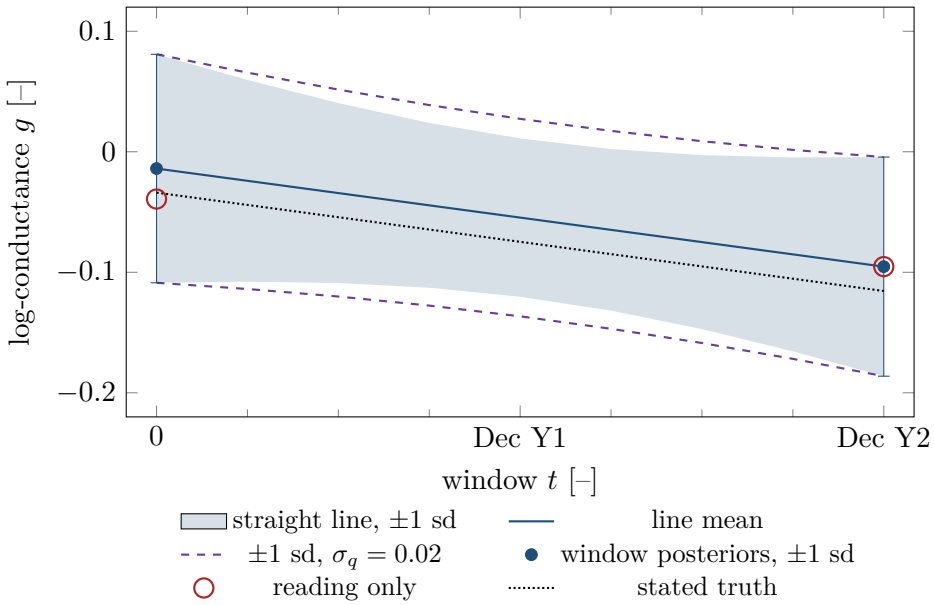


Figure 4.8: The path of the log-conductance across windows. Shaded: the straight line $g_t = \hat{g}_0 + \hat{\gamma}\, t\, \Delta t$ with one standard deviation either side, in the limit $\sigma_q \to 0$. Dashed: the band of the monthly chain at $\sigma_q = 0.02$, chosen so that the bridge term $\sigma_q^2\, t(24 - t)/24$ is visible, at most 2.400×10^{-3} in variance at window 12; at the reported $\sigma_q = 0.005$ it is 1.500×10^{-4} and the band lies on the shaded edge. Dots with bars: the window posteriors, -0.013926 ± 0.094787 and -0.095310 ± 0.090909. Open circles: the reading-only values $\ln(5/5.2)$ and $\ln(5/5.5)$. Dotted: the stated truth, $G_0 = 5.80\,\text{kW/K}$ and a loss of 4 %/yr.

Set against the trend chart's 2.77 %/yr, the posterior median is steeper, 3.99 %/yr. The handbook prior raises the commissioning entry, from its reading-only value $\ln(5/5.2) = -0.0392$ (the fit without the prior row of Section 1.4.4) to -0.0139, and nothing raises the window-24 entry. The band is the larger finding: two tests on a 0.5 K sensor cannot resolve a loss of a few percent a year.

The condition test is the second entry of the maintenance log, read against the commissioning baseline:

> *Window 24 (Dec Y2), condition test.*
> Conductance 5.4545 kW/K against the baseline's 5.9170 kW/K, a ratio of 0.9218 (5–95 %: 0.7427 to 1.1441).
> Degradation rate −3.99 %/yr (5–95 %: −13.98 to +7.16 %/yr); probability of a loss 0.729.
> Next measurement: a further condition test before the rate is acted on.

The stated truth. The later test is constructed. Its stated truth is a commissioned conductance $G_0 = 5.80\,\text{kW/K}$ ($g_0 = -0.033902$) and a trend $\gamma = \ln 0.96 = -0.040822\,\text{yr}^{-1}$, a loss of 4 %/yr on a straight line with no wander. The commissioning reading lies +0.055 standard deviations of the sensor from the truth's test temperature, and the window-24 reading is drawn from the truth with seed 4: its error is −0.225 standard deviations as reported, −0.175 before rounding to the sensor's tenth of a degree. The true trend lies 0.002 posterior standard deviations from the median: that closeness is the luck of two draws, not the precision of the fit, whose band spans $[-0.150569,\ 0.069184]\,\text{yr}^{-1}$. The same truth checks the main example: its resistance $1/5.80 = 0.172414\,\text{K/kW}$ lies +0.23 posterior standard deviations from $\hat{\rho} = 0.169231$, and its full-load temperature, 28.34°C, is above the limit, so the release truly fails, an outcome to which the posterior gave the probability $1 - 0.4267 = 0.5733$.

The checks. Table 4.15 sets the reported route against the exact posterior of the trend entry, at $\sigma_q = 0.005$ and on the straight line, with the stated truth in the last row. Each row is a route; the columns give the median of γ, its 5 % and 95 % quantiles and the probability of a loss. The exact posterior is computed by quadrature. Since the two windows are independent a posteriori, the exact posteriors of g_0 and g_{24} are each computed on a grid from their own rows; the density of $g_{24} - g_0$ is their convolution, and convolved again with $\mathcal{N}(0, 24\,\sigma_q^2)$ and divided by 2 yr it is the density of γ; its quantiles are interpolated in the cumulative distribution. The far tail named among the choices of the window state is measured here: as g_{24} grows the window-24 likelihood tends to $\exp[-(y_{T,24} - T_c)^2/(2\sigma_T^2)] = 5.3 \times 10^{-27}$ of its peak rather than to zero, and at that height the mass beyond the grid is negligible.

Table 4.15: The trend entry by route, with the stated truth of the constructed later test.

Route	Median γ [yr^{-1}]	5 % [yr^{-1}]	95 % [yr^{-1}]	$\Pr(\gamma < 0 \mid \mathbf{y})$ [–]
Laplace, $\sigma_q = 0.005$ (reported)	−0.0407	−0.1506	0.0692	0.7288
exact posterior by quadrature, $\sigma_q = 0.005$	−0.0406	−0.1528	0.0713	0.7258
Laplace, straight line	−0.0407	−0.1487	0.0673	0.7323
exact posterior by quadrature, straight line	−0.0406	−0.1510	0.0695	0.7292
stated truth	−0.0408	–	–	–

The Laplace route exceeds the exact one in the probability of a loss by 0.0030 at both spreads, and its band is the narrower at both ends. The gap comes from the spreads: the exact standard deviations of the two windows, 0.096695 and 0.093093, are 2.0 and 2.4 % wider than the Laplace 0.094787 and 0.090909, so the exact posterior of γ is wider and puts a little more mass above zero. The medians agree because the skews of the two windows largely cancel in their difference: the exact posterior means of g_0 and g_{24} are −0.001342 and −0.082572, above their MAP values by 0.0126 and 0.0127. The monthly chain of 26 entries and the three-entry state agree to less than 10^{-8}. The reduced state of the readings is checked against the full state $(T_0, g_0, T_{24}, g_{24}, \gamma)$ with physical rows of tolerance σ_p: at $\sigma_p = 0.1$, 0.01 and 0.001 K the trend entry is −0.041164, −0.040697 and $-0.040692\,\mathrm{yr}^{-1}$ with standard deviations 0.068055, 0.066813 and 0.066800, against the reduced state's −0.040692 and 0.066800. The physical row's curvature term on g_0 tends to −2.6341 and does not vanish in the limit: it is the curvature of the constraint curve, the term the reduced state carries on its reading row.

The rows later chapters cite. The monthly trend rows (4.68), the marginal row (4.69) across windows without a reading, the straight line (4.74) as their limit $\sigma_q \to 0$ through the reduced state, the variance of the trend entry (4.73), and the trend entry read through equation (2.21) are what the later chapters cite. Chapter 10 puts a reading in each of 36 windows; the rows are unchanged, and only the record grows.

4.6.6 What the analysis showed

The pre-test estimate said the hall passes at 27.6°C, and the reading scaled to full load said it fails at 28.4°C; the analysis says neither. With the reading, the pre-test estimate and the physics combined, the probability that the hall stays below 28°C at 60 kW is 0.4267, less than even, and the full-load temperature is not known well enough to call it either way. One reading at the test condition fixes the thermal resistance of the air path to about ±8 % and the full-load temperature to about ±0.8 K, one standard deviation each, although the hall has never run at 60 kW: the physics carries the reading along the free direction $\mathbf{v}$, from T_1 to ρ and from ρ to

T_2. The posterior is a ridge along that direction, the three entries known almost only as one combination, and the reduced state is that ridge named by its one free entry: every question asked of the hall is a question about where it sits on the ridge. The 60 kW question in particular is a resistance question: the hall stays below 28°C exactly when the resistance stays below ρ^*, and the reading at 30 kW places the resistance close to ρ^*, which leaves the question near even odds. Further readings at the test condition narrow the ridge and resolve the question without deciding its answer. A reading at full load would act on the third column of $\mathbf{J}$ directly, and its worth is what its preposterior covariance $\mathbf{\Sigma}_a$ says before it is taken: as much as four readings at the test condition, with the shared resistance itself put to the test. A second reading in the same window narrows the same entry; a later window adds a new entry and a trend row. Two tests two years apart make a loss more likely than not (0.73), but they cannot say how fast: the band runs from a loss of 14 %/yr to a gain of 7 %/yr.

4.6.7 Limits of the model

The steady-state balance (4.2) neglects the heat stored in the hall and its equipment, and it describes the temperature long after a load change, not minutes after it; Exercise 4.6 formulates the transient model that a question about the first minutes of full load would need. The single effective path assumes that the resistance is the same at 60 kW as at 30 kW; a path whose resistance changes with load, for example because the unit's fans or its coil change state, would break the link drawn in Figure 4.2, and one test reading cannot detect it. The supply air is pinned at 18°C by the discharge controller, which holds it only while the unit has capacity to spare; a unit at its limit at full load lets the supply air rise, which shifts the full-load temperature up, not only widens it; an uncertain supply air adds its variance through the budget (4.54).

The probabilistic assumptions have limits of their own. The Gaussian prior on a positive resistance is harmless here, with a negative mass of 8×10^{-11}, but a wide prior on a small quantity would place real probability below zero; a logarithmic parameterization avoids this at the price of a nonlinear model and an approximate posterior. The sensor error is taken as Gaussian, unbiased and of known spread; a bias would shift every inferred quantity, and Section 4.2.2 names the comparison of models that would detect it. The physical rows are taken in the limit of zero tolerance; a real model discrepancy is a finite σ_p, which widens Σ_{11} and adds σ_p^2 to the full-load variance by equation (4.34), and the reduced state of Section 4.2.8 is exact only in that limit. The record is constructed, not field data: the stated truth of Section 4.6.5 covers both windows, the commissioning reading of 23.2°C sits near it and the later reading is drawn from it, and the checks compare the answers with it. The trend rows rest on the choices named with them there; of these, the straight line is the one two windows cannot test, since a path that bends or steps between the tests is drawn as the line through them. Each of these is a statement about the model, and the probability 0.4267 is conditional on all of them.

Chapter 5 grows the hall: two coils fed with chilled water produce the supply air, now the cold aisle, and three racks draw from it. Each coil's health index is read against its own commissioning fit; the method is unchanged, and it runs on the reduced state of this chapter.

4.7 What to retain

Definition 4.1 (The anatomy of an inference)**.** A condition or acceptance probability over a plant graph is obtained in two operations. Inference sums one term per row into the objective $\mathcal{L} = -\ln p(\mathbf{z} \mid \mathbf{y})$ up to a constant, and uses only its gradient and Hessian. A Gaussian row's term is half its squared residual; a count row's is its negative log-probability, which joins the Jacobian through its working residual and line, reproducing the row's gradient and information but not its value. The physical rows, the readings and the priors, each divided by its standard deviation, stack into one residual vector $\mathbf{r}(\mathbf{z})$ with Jacobian $\mathbf{J}$; the normal

equations $(\mathbf{J}^\mathsf{T}\mathbf{J})\,\Delta\mathbf{z} = -\mathbf{J}^\mathsf{T}\mathbf{r}$ lead to the MAP state $\hat{\mathbf{z}}$, and the inverse of the information matrix, $\boldsymbol{\Sigma} = (\mathbf{J}^\mathsf{T}\mathbf{J})^{-1}$, is the posterior covariance, exactly for a linear-Gaussian model and under the Laplace approximation, in its Gauss–Newton form, otherwise. An exact physical relation is the limit of a row's tolerance, never a substitution; in that limit the state reduces to its free entries, with Jacobian $\mathbf{J}_{\text{rows}}\mathbf{N}$ on the null space $\mathbf{N}$ of the physical rows. A later window adds a new entry of the parameter and a trend row $(g_{t+1} - g_t - \gamma\,\Delta t)/\sigma_q$ tying it to the last, and the straight line is that row's exact limit $\sigma_q \to 0$. Query evaluation then reads the answer off the posterior: an entry of the state with its variance Σ_{ii}, or a derived quantity $f(\mathbf{z})$ with variance $\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}$, and a threshold probability as the normal tail $\Phi\big((f^* - f(\hat{\mathbf{z}}))/\sqrt{\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}}\big)$, which for an entry is $\Phi\big((f^* - \hat{z}_i)/\sqrt{\Sigma_{ii}}\big)$.

Three conclusions carry beyond the example. Physical relations connect the quantities that are measured to those that are not, which is what allows an unobserved quantity, here a full-load temperature never reached, to be inferred from indirect evidence. Uncertainty is not an error bar attached to a point estimate but a probability distribution, and it must be propagated through the physical model to answer an asset manager's question. And the posterior is reusable: once the joint posterior is in hand, many questions are answered without repeating the inference, and new evidence updates it rather than replacing it. A derived quantity is propagated in one of two ways: its spread by the gradient, $\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}$, and its density by the change of variables, which multiplies the density by the change-of-variable factor of the map and so keeps the skew and the bounds that a linearization loses.

The inference of the hall returns the posterior state $\hat{\mathbf{z}}$ of equation (4.49), its covariance $\boldsymbol{\Sigma}$ of equations (4.50) and (4.35), and the Jacobian $\mathbf{J}$ of equation (4.25), from which the gradients of derived quantities are formed. Table 4.16 lists what the chapter has computed from them without repeating the inference: the four answers of Table 4.2, three further quantities that later chapters reuse, and the trend entry, read from the posterior of the window state. Each row is a quantity, the calculation on the posterior that gives it, and the place where it is carried out.

Table 4.16: The four answers of Table 4.2, three further quantities reused later, all computed from one posterior of the equipment hall, and the trend entry, from the posterior of the window state.

Quantity	Calculation on the posterior	Section
The answers		
Thermal resistance of the air path	the entry $\hat{z}_1$ and $\sqrt{\Sigma_{11}}$	4.3.6
Hall temperature at the test condition	the entry $\hat{z}_2$ and $\sqrt{\Sigma_{22}}$	4.3.6
Hall temperature at full load	the entry $\hat{z}_3$, $\sqrt{\Sigma_{33}}$ and the credible interval	4.3.6
Probability of staying below 28°C at 60 kW	the Gaussian tail $\Phi((T_{\max} - \hat{z}_3)/\sqrt{\Sigma_{33}})$	4.3.7
Further quantities reused later		
Spread of the conductance	$\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}$ with $\mathbf{g} = (-1/\hat{z}_1^2, 0, 0)^\mathsf{T}$, and the exact push-forward	4.3.6
The entry that drives the temperature	the gradient and the variance budget	4.3.2
Effect of another reading	the information matrix plus the reading's $\mathbf{j}^\mathsf{T}\mathbf{j}$	4.6.1
The window state		
Trend entry (degradation rate)	$\hat{\gamma}$ and $\Sigma_{\gamma\gamma}$ of the window state, read as $\Pr(\gamma < 0 \mid \mathbf{y})$	4.6.5

The table is one inference followed by a column of cheap queries on its result, the division of Section 4.1.2. Two kinds of question fall outside it: a question that compares two models, which needs the evidence of each (Section 4.2.2), and the arrival of new evidence, which changes the posterior itself (Section 4.6.4).

Table 4.17 lists the five mathematical objects of the chapter that recur throughout the book, with the form of each and its purpose. The first three are the working objects when the posterior is well represented by a local Gaussian. The posterior distribution itself is the more fundamental object: it may be Gaussian, multimodal, discrete, a mixture or otherwise non-Gaussian, and its structure decides which computational method is appropriate.

Table 4.17: The five mathematical objects of an inference.

Object	Form	Purpose
State estimate	$\hat{\mathbf{z}}$	the most probable (or otherwise selected) central state
Posterior covariance	$\boldsymbol{\Sigma}$	local uncertainty of the state and the dependence among its entries
Output gradient	$\mathbf{g} = \nabla f(\hat{\mathbf{z}})$	connects state uncertainty to a derived quantity
Posterior distribution	$p(\mathbf{z} \mid \mathbf{y}, M)$	the uncertainty about the state given the evidence
Model evidence	$p(\mathbf{y} \mid M)$	normalizes the posterior and compares models

Table 4.18 lists the terms of the chapter with their expressions, units and the subsection that defines each.

Table 4.18: The terms of the anatomy of an inference and where each is defined. Units for the equipment hall.

Term	Expression	Units	Meaning	Defined in
Thermal resistance	$\rho = 1/G$, $T = T_c + Q\rho$	K/kW	makes every row linear in the state	4.1.4
Physical rows	$r_1 = T_1 - 18 - 30\rho$, r_2	K	conservation at each hall node	4.1.5
The question as one marginal	$\Pr(T_2 < T_{\max})$ from $\hat{z}_3$, Σ_{33}	–	one entry's mean and variance	4.1.6
Model tolerance	σ_p; exact rows as $\sigma_p \to 0$	K	model discrepancy of a physical row	4.1.5
Likelihood	$p(y_T \mid \mathbf{z})$, eq. (4.14)	K^{-1}	the reading as a function of the state	4.2.1
Residual vector	$\mathbf{r}(\mathbf{z})$, eq. (4.23)	–	four rows, each over its standard deviation	4.2.3
Objective	$\mathcal{L} = \frac{1}{2}\mathbf{r}^{\mathsf{T}}\mathbf{r}$	–	minus log posterior	4.2.3
Jacobian	$\mathbf{J} = \partial\mathbf{r}/\partial\mathbf{z}$, 4×3	mixed	one line per row, one column per entry	4.2.3
Information matrix	$\mathbf{J}^{\mathsf{T}}\mathbf{J} = \sum_k \mathbf{j}_k^{\mathsf{T}}\mathbf{j}_k = \tilde{\mathbf{J}}^{\mathsf{T}}\mathbf{W}\tilde{\mathbf{J}}$	mixed	the Hessian when $\mathbf{C} = \mathbf{0}$	4.2.3
General objective	$\mathcal{L} = \sum_k \mathcal{L}_k$, $\mathcal{L}_k = -\ln p_k$	–	one term per row, any distribution	4.2.7
Count row, working form	$r_k = (\mu_k - n_k)/\sqrt{\mu_k}$, $\mathbf{j}_k = \sqrt{\mu_k}\,\nabla \ln \mu_k^{\mathsf{T}}$	–	gradient and information, not the term	4.2.7
Hessian	$\mathbf{H} = \mathbf{J}^{\mathsf{T}}\mathbf{J} + \mathbf{C}$	mixed	the information plus a named curvature term	4.2.7
Normal equations	$(\mathbf{J}^{\mathsf{T}}\mathbf{J})\,\Delta\mathbf{z} = -\mathbf{J}^{\mathsf{T}}\mathbf{r}$	mixed	Gauss–Newton step to $\hat{\mathbf{z}}$	4.2.4
Posterior covariance	$\mathbf{\Sigma} = (\mathbf{J}^{\mathsf{T}}\mathbf{J})^{-1}$	mixed	spread is inverse information	4.2.5
Laplace approximation	$\mathcal{N}(\hat{\mathbf{z}}, \mathbf{H}^{-1})$	–	local Gaussian posterior	4.2.6
Free direction	$\mathbf{v}$, $\mathbf{J}_{\mathrm{phys}}\mathbf{v} = \mathbf{0}$	mixed	$\mathbf{\Sigma} \to \mathbf{v}\mathbf{v}^{\mathsf{T}}/5200$, rank one	4.2.4
Delta method	$\mathrm{Var}(f) \approx \mathbf{g}^{\mathsf{T}}\mathbf{\Sigma}\mathbf{g}$	varies	linearized propagation	4.3.1
Change of variables	$p_w(w) = p_x(h^{-1}(w))\,\lvert \mathrm{d}x/\mathrm{d}w \rvert$, eq. (4.55)	varies	density of a transformed quantity	4.3.3
Variance budget	eq. (4.54)	K^2	one term per entry plus cross terms	4.3.2
Threshold probability	$\Phi\big((T_{\max} - \hat{z}_3)/\sqrt{\Sigma_{33}}\big)$	–	normal tail at the limit	4.3.5
Credible interval	$\hat{z}_3 \pm 1.960\sqrt{\Sigma_{33}}$	°C	central 95 % posterior range	4.3.4
Reduced state	$\mathbf{N}$, $\mathbf{J}_{\mathrm{red}} = \mathbf{J}_{\mathrm{rows}}\mathbf{N}$, $\mathbf{\Sigma}_{\mathrm{red}}$	mixed	the free entries with exact physical rows	4.2.8
Prior predictive	$\mathcal{N}(22.8, 0.9014^2)$	°C	reading predicted before use	4.4.2
Posterior predictive	$\mathcal{N}(23.077, 0.6504^2)$	°C	next reading predicted	4.4.4
Evidence	$p(y_T \mid M)$, eq. (4.18)	K^{-1}	normalizer; compares models	4.2.2
Bayes factor	$B_{\mathrm{two,one}} = p(y_T \mid M_{\mathrm{two}})/p(y_T \mid M_{\mathrm{one}})$	–	evidence ratio of two models	4.2.2
Rank-one update	eq. (4.66)	–	adds one reading to a posterior	4.6.4
Preposterior covariance	$\mathbf{\Sigma}_a = (\mathbf{\Sigma}^{-1} + \mathbf{\Lambda}_a)^{-1}$, eq. (4.65)	mixed	ranks candidate measurements	4.6.1
Trend row	$(g_{t+1} - g_t - \gamma\,\Delta t)/\sigma_q$, eq. (4.68)	–	ties one window's entry to the next	4.6.5
Marginal row across unread windows	$(g_{t+n_{\mathrm{w}}} - g_t - n_{\mathrm{w}}\,\Delta t\,\gamma)/(\sigma_q\sqrt{n_{\mathrm{w}}})$, eq. (4.69)	–	a gap of n_{w} windows as one row	4.6.5
Trend entry	γ, $\mathrm{Var}(\gamma)$ of eq. (4.73)	yr^{-1}	degradation rate; $\Pr(\gamma < 0 \mid \mathbf{y})$	4.6.5
Straight line	$\sigma_q \to 0$: $g_{24} = g_0 + (2\,\mathrm{yr})\,\gamma$	–	the trend row as a constraint, on its reduced state	4.6.5

4.8 Exercises

Exercise 4.1 (Application: a complete inference)**.** A different equipment hall obeys $T = T_c + Q\rho$ with its supply air held at $T_c = 20°\mathrm{C}$ and a test load $Q = 20\,\mathrm{kW}$. The prior on the resistance is

$\rho \sim \mathcal{N}(0.20, 0.04^2)$ in K/kW, and a sensor with $\sigma_T = 0.4\,\mathrm{K}$ reads $y_T = 24.5°\mathrm{C}$. The question is the temperature at a requested load of 40 kW. (a) Write the state $\mathbf{z} = (\rho, T_1, T_2)^\mathsf{T}$, with T_1 the test temperature and T_2 the temperature at 40 kW, the four normalized rows, and the 4×3 Jacobian at $\sigma_p = 0.01\,\mathrm{K}$. ($\mathbf{r} = ((T_1 - 20 - 20\rho)/\sigma_p,\ (T_2 - 20 - 40\rho)/\sigma_p,\ (T_1 - 24.5)/0.4,\ (\rho - 0.20)/0.04)^\mathsf{T}$; $\mathbf{J}$ has the lines $(-2000, 100, 0)$, $(-4000, 0, 100)$, $(0, 2.5, 0)$ and $(25, 0, 0)$.) (b) Take one Gauss–Newton step from the handbook state and compute $\boldsymbol{\Sigma}$ in the limit of exact rows. (At $\mathbf{z}^{(0)} = (0.20,\ 24.0,\ 28.0)^\mathsf{T}$ only the reading row is nonzero, -1.25, and $\mathbf{J}^\mathsf{T}\mathbf{r} = (0,\ -3.125,\ 0)^\mathsf{T}$; the normal equations give $\hat{\mathbf{z}} = (0.2200,\ 24.400,\ 28.800)^\mathsf{T}$, and a second step does not move it. The free direction is $\mathbf{v} = (1,\ 20,\ 40)^\mathsf{T}$, with $\mathbf{J}\mathbf{v} = (0,\ 0,\ 50,\ 25)^\mathsf{T}$ and information $50^2 + 25^2 = 3125$, so $\boldsymbol{\Sigma} \to \mathbf{v}\mathbf{v}^\mathsf{T}/3125$: $\Sigma_{11} = 3.2 \times 10^{-4}$, $\sqrt{\Sigma_{11}} = 0.01789\,\mathrm{K/kW}$; at $\sigma_p = 0.01\,\mathrm{K}$ the fit gives $\Sigma_{11} = 3.2016 \times 10^{-4}$ and a correlation of 0.9996 between ρ and T_1.) (c) Read off the posterior mean and standard deviation of the temperature at 40 kW, and the probability that it stays below 30°C. ($\hat{z}_3 = 28.8°\mathrm{C}$ and $\sqrt{\Sigma_{33}} = \sqrt{1600/3125} = 0.7155\,\mathrm{K}$; $\Pr(T_2 < 30 \mid \mathbf{y}) = \Phi(1.2/0.7155) = \Phi(1.677) = 0.9532$, against $\Phi(2/1.6) = 0.8944$ under the prior alone.)

Exercise 4.2 (Derivation: a variance budget with correlated entries)**.** For $T = T_c + Q\rho$ at the known load $Q = 50\,\mathrm{kW}$, the posterior covariance of (T_c, ρ) is $\boldsymbol{\Sigma} = \begin{bmatrix} 0.25 & -0.003 \\ -0.003 & 0.0001 \end{bmatrix}$, in K^2, K·K/kW and $(\mathrm{K/kW})^2$. (a) Derive the temperature variance from $\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}$ and compute the standard deviation. ($\mathbf{g} = (1, 50)^\mathsf{T}$; $\mathrm{Var}(T) = 0.25 + 50^2 \times 0.0001 + 2 \times 50 \times (-0.003) = 0.25 + 0.25 - 0.30 = 0.20\,\mathrm{K}^2$, $\mathrm{sd}(T) = 0.4472\,\mathrm{K}$.) (b) State the result if the covariance term were ignored, and explain the difference. ($\sqrt{0.50} = 0.7071\,\mathrm{K}$, 58 % too large; the posterior correlation is $-0.003/\sqrt{0.25 \times 0.0001} = -0.6$, and with sensitivities of the same sign the negative cross term cancels part of the two diagonal terms: the evidence pins the combination $T_c + 50\rho$ better than either entry.)

Exercise 4.3 (Derivation: a nonlinear derived quantity)**.** With $\rho \mid \mathbf{y} \sim \mathcal{N}(0.169231, 0.013868^2)$ from the chapter's example, derive the first-order approximation to the variance of the conductance $G = 1/\rho$, compute its standard deviation, and explain why the result is only an approximation. (With $G(\mathbf{z}) = 1/z_1$, $\mathbf{g} = (-1/\hat{z}_1^2,\ 0,\ 0)^\mathsf{T}$, so $\mathrm{Var}(G \mid \mathbf{y}) \approx \mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g} = \Sigma_{11}/\hat{z}_1^4$; $\hat{G} = 1/0.169231 = 5.909\,\mathrm{kW/K}$ and $\mathrm{sd}(G \mid \mathbf{y}) \approx 0.013868/0.169231^2 = 0.484\,\mathrm{kW/K}$. The map $1/\rho$ is nonlinear, so G is not Gaussian: its exact 5–95 % band is $[5.207, 6.830]$, skewed, against the linearized $[5.113, 6.706]$; strictly, the moments of $1/\rho$ do not exist under a Gaussian ρ, which has positive density at zero, so positivity and the transformed distribution must be handled explicitly when exact conductance statements are needed.)

Exercise 4.4 (Formulation: observability and the plant graph)**.** In $T = T_c + Q\rho$ suppose that both T_c and ρ are unknown, the discharge controller's set point not being trusted, and that there is a single temperature reading at one known load. (a) Write the observation Jacobian with respect to (T_c, ρ) and state whether the reading can determine both unknowns without further information. ($y = T_c + Q\rho + \epsilon$, so $\mathbf{J} = [1 \ \ Q]$, of rank one for two unknowns: the reading determines only the combination $T_c + Q\rho$, and every pair with the same combination explains it equally well.) (b) Name a reading or operating condition that resolves the ambiguity. (A second reading at a different known load $Q_2 \neq Q_1$ gives $\mathbf{J} = \begin{bmatrix} 1 & Q_1 \\ 1 & Q_2 \end{bmatrix}$, with determinant $Q_2 - Q_1 \neq 0$ and full rank; alternatively, a direct reading of the supply air, by a discharge sensor, a new node observed in the graph, removes the ambiguity.)

Exercise 4.5 (Application: Bayesian evidence)**.** A model predicts $Y \sim \mathcal{N}(24, 1^2)$ in °C, and the reading $y = 25°\mathrm{C}$ arrives. (a) Compute the marginal likelihood $p(y \mid M)$. ($\frac{1}{\sqrt{2\pi}} \exp[-(25 - 24)^2/2] = 0.39894 \times 0.60653 = 0.2420\,\mathrm{K}^{-1}$.) (b) Explain why this value is not the probability

that M is correct. (It is a probability density in inverse temperature units, 2.420×10^{-4} per millikelvin for the same reading, not $\Pr(M \mid y)$; a posterior model probability requires prior probabilities and the evidence of the alternative models under comparison.)

Exercise 4.6 (Formulation: model validity and thermal storage)**.** The chapter assumed the steady-state balance $Q = G(T - T_c)$. The load now steps from 30 kW to 60 kW in the release test, and the asset manager asks for the temperature five minutes later. (a) Explain why the steady-state equation may be insufficient. (It neglects the energy stored in the air, structure and equipment of the hall, so it gives only the temperature reached long after the change; five minutes may be short against the thermal time constant.) (b) Starting from an energy balance, formulate a model with thermal storage and name the new parameter and the condition it needs. (With an effective thermal capacitance C_{th} in kJ/K, $C_{\mathrm{th}}\,\mathrm{d}T/\mathrm{d}t = Q - G(T - T_c)$; the new parameter is C_{th} and an initial temperature $T(0)$ is required. For constant Q, G and T_c after the step, $T(t) = T_c + Q/G + [T(0) - T_c - Q/G]\,\mathrm{e}^{-Gt/C_{\mathrm{th}}}$, with time constant C_{th}/G; the probability for the temperature at five minutes propagates the posterior of G, C_{th} and $T(0)$ through this time-dependent model rather than through the steady-state equation.)

Exercise 4.7 (Application: a third condition test)**.** Between the two tests of Section 4.6.5, a condition test in window 12 (Dec Y1), at the same load and supply air, reads 23.1°C. (a) State what the trend chart concludes from it. ($G = 30/5.1 = 5.8824\,\mathrm{kW/K}$, above the commissioning value of 5.7692 kW/K: an apparent gain.) (b) Write the state and the trend rows. ($\mathbf{z} = (g_0,\ g_{12},\ g_{24},\ \gamma)^{\mathsf{T}}$, with the reading rows of equation (4.70) at the three windows, the prior on g_0, and two marginal trend rows of equation (4.69), $(g_{12} - g_0 - 12\,\Delta t\,\gamma)/(\sigma_q\sqrt{12})$ and $(g_{24} - g_{12} - 12\,\Delta t\,\gamma)/(\sigma_q\sqrt{12})$.) (c) Fit the straight line and say why the commissioning entry now moves. (On the line state (g_0, γ), $\hat{g}_0 = -0.002291$ and $\hat{\gamma} = -0.041158\,\mathrm{yr}^{-1}$, standard deviation 0.066712, $\Pr(\gamma < 0 \mid \mathbf{y}) = 0.7314$. The commissioning entry moves from -0.0139 to -0.0023 because the line now ties three windows, and the cool reading at window 12 pulls it up; with two windows the trend entry absorbed the trend row and window 0 stood alone.) (d) Repeat at $\sigma_q = 0.005$ and 0.02 and say why the spread now moves the mean. (At 0.005, $\hat{\gamma} = -0.041153\,\mathrm{yr}^{-1}$ with standard deviation 0.067815 and $\Pr(\gamma < 0 \mid \mathbf{y}) = 0.7280$; at 0.02, $\hat{\gamma} = -0.041087\,\mathrm{yr}^{-1}$ with standard deviation 0.082638. With three windows the wander lets window 12 leave the line, so the spread sets how hard the middle reading pulls on the trend, and the mean moves by $7.1 \times 10^{-5}\,\mathrm{yr}^{-1}$ from the line to $\sigma_q = 0.02$; the probability of a loss at $\sigma_q = 0.02$ is 0.6905.)

Part II

Assessing Present Condition

Chapter 5

Health Indices from Physical Parameters

5.1 The dashboard's health index and the coil's

The hall was handed over eighteen months ago. Its commissioning test, window 0, ran at about 110 kW on the hall's meter (110.92 kW) with the chilled water at 14°C, and it is the baseline every later condition reading is measured against. Today is June of year 2, window 18. The condition-monitoring dashboard reads each coil's approach, the discharge air less the chilled-water supply, against its commissioning value. Coil 1's approach has risen from 2.86 to 3.60 K, by 0.74 K or 25.9 %; coil 2's from 3.08 to 4.36 K, by 1.28 K or 41.6 %. The health-index tile reports the ratio, the commissioning approach over today's: coil 1 at 0.79, "lost 20.6 %", and coil 2 at 0.71, "lost 29.4 %". Both tiles are red. The conclusion the dashboard invites is that both coils have fouled badly, coil 2 worse, and that both are due for cleaning.

The tile cannot say which part of that fall is the coils. The load has grown since handover, from 110.92 to 130.99 kW on the hall's meter; the season has moved the other heat and the chilled-water supply; and a coil's approach depends on the other coil as well, through the return air the two share. Nor can the tile say how well the commissioning test and window 18, eighteen months apart, know a change of a few percent. The question of this chapter is

how a coil's health index is read as a physical parameter of the plant, against the commissioning fit, with its band, and how much of what the dashboard calls aging is the coil.

It is answered with the method of Chapter 4, on the reduced state of the two-unit hall, carried now in two windows at once.

Jargon: health index against commissioning, derate factor, condition assessment, approach temperature, raw-performance index, joint state of two windows, commissioning floor, water-side heat meter

The *health index against commissioning* sets a parameter against its value at the commissioning test; the *derate factor* of a coil sets its conductance against the maker's rated value. A *condition assessment* reads every health index of an asset in one window against its commissioning baseline. The *approach temperature* of a cooling unit is its discharge air less the chilled-water supply, $T_{d,u} - T_s$: a performance figure of the unit, not a parameter of it. A *raw-performance index* is a performance figure at a later window set against its value at commissioning, here the ratio of two approaches; condition-monitoring dashboards display it as a health index. The *joint state of two windows* carries the entries of both windows in one state vector, so that one posterior covers the commissioning test and today. The *commissioning floor* is the part of a later health index's variance that is the commissioning test's own, which no measurement after handover removes. A *water-side heat meter* reads the heat a coil passes to the chilled water, from the water's flow and its temperature rise.

Key idea

A health index is the ratio of one physical parameter in two windows, $\mathrm{e}^{g_{u,18}-g_{u,0}}$ for coil u. It is read from one posterior over both windows, so its band carries the commissioning test's error as well as today's. The load enters the physics as an entry of each window and the season as known inputs, and both leave the parameter; a raw-performance index, a ratio of performance figures, cannot remove them.

5.1.1 The health index as a parameter

Chapter 1 read the hall's index against design, $h = G/G_{\mathrm{ref}}$. Chapter 4 read it against the commissioning fit, over two condition tests run at the same load and the same supply air, and there the raw test temperature and the log-conductance carried the same information. Here the inputs move between the two windows: the IT load has grown, and in June, winter in the southern hemisphere where the hall stands, the chilled water is colder and the other heat lower. The approach of a coil depends on all of them, and on the other coil; the coil's conductance does not, because it is a property of its tubes and fins, the same quantity at any load. The parameter, not the performance figure, is what stays comparable from window to window.

The index of coil u in window t against commissioning is

$$h_{u,t} = \frac{G_{u,t}}{G_{u,0}} = \frac{\bar{G}_u\,\mathrm{e}^{g_{u,t}}}{\bar{G}_u\,\mathrm{e}^{g_{u,0}}} = \mathrm{e}^{g_{u,t}-g_{u,0}}, \qquad \ln h_{u,t} = g_{u,t} - g_{u,0}, \tag{5.1}$$

with $\bar{G}_u$ the rated conductance and $g_{u,t} = \ln(G_{u,t}/\bar{G}_u)$ the log-derate. The rating cancels: the index against commissioning does not depend on the rated value, while the derate factor $d_{u,t} = \mathrm{e}^{g_{u,t}}$, the index against rating, does. The logarithm of the index is the change since commissioning of Chapter 3, a difference of two entries of the state, and both entries are estimates: the commissioning test is a fit like any other, with an error of its own.

Principle 5.1. *A health index is the posterior of a parameter's change since commissioning, with the variance of a difference, equation* (3.38)*. A performance figure is a derived quantity of the whole state, and its change mixes every entry it depends on.*

5.1.2 The two-unit hall in two windows (step 0)

The plant is the two-unit hall of *Operations and Control by Derivation*, Chapter 5, read as the reduced state of Chapter 4 reads a plant with nonlinear physical rows (Section 4.2.8). Part I's single air path is split in two. Two cooling coils, each with a rated conductance $\bar{G}_u = 7.0\,\mathrm{kW/K}$, take heat from the return air, the warm air the racks give off, at T_r, into chilled water supplied at $T_s(t)$. The air crosses each coil at its unit's air capacity rate $C_u = 10.0\,\mathrm{kW/K}$, the air mass flow times the specific heat of air, pinned at the unit's fixed fan speed, and leaves it as the discharge air, at $T_{d,u}$. The two discharge streams mix in the cold aisle in front of the racks, at T_c, and the racks draw their inlet air from it, each with a recirculation fraction β_j, the share of return air at its inlet, which an airflow survey estimated at handover; rack 1 has an inlet sensor. The heat to remove in window t is the IT load Q_t, an entry of the state read by the hall's load meter, and the other heat $Q_o(t)$ of the lighting, the fans and the envelope.

The physical rows. In each window five relations hold: the hall balance, by which the heat put in leaves through the two coils; the air side of each unit, by which the air crossing coil u gives up what the coil passes to the water; the mixing of the two discharge streams in the cold aisle; and the inlet of rack 1, cold-aisle air with a share β_1 of return air. With $G_{u,t} = \bar{G}_u\,\mathrm{e}^{g_{u,t}}$ and the temperatures those of window t, written as rows that vanish when they hold,

$$\begin{aligned}
r_{\mathrm{h},t} &= \big[(G_{1,t}+G_{2,t})\,(T_r - T_s(t)) - Q_t - Q_o(t)\big]/\sigma_p && && \text{hall balance,}\\
r_{\mathrm{a},u,t} &= \big[C_u\,(T_r - T_{d,u}) - G_{u,t}\,(T_r - T_s(t))\big]/\sigma_p, && u = 1,2 && \text{air side of unit } u,\\
r_{\mathrm{m},t} &= \big[(C_1+C_2)\,T_c - C_1 T_{d,1} - C_2 T_{d,2}\big]/\sigma_p && && \text{cold-aisle mixing,}\\
r_{\mathrm{i},t} &= \big[T_1 - T_c - \beta_1\,(T_r - T_c)\big]/\sigma_p && && \text{rack-1 inlet,}
\end{aligned} \tag{5.2}$$

the first four in kW and the last in K, each over its tolerance σ_p; the limit of exact rows takes the tolerances to zero together. The five rows fix the five temperatures $(T_r, T_{d,1}, T_{d,2}, T_c, T_1)$ of the window. Solved in that order, the hall balance first, they give the forward equations of the operations book, written here with the window's known inputs,

$$\begin{aligned}
T_r &= T_s(t) + \frac{Q_t + Q_o(t)}{G_{1,t}+G_{2,t}}, \qquad T_{d,u} = T_r - \frac{G_{u,t}\,(T_r - T_s(t))}{C_u},\\
T_c &= \frac{C_1 T_{d,1} + C_2 T_{d,2}}{C_1 + C_2}, \qquad T_1 = T_c + \beta_1\,(T_r - T_c),
\end{aligned} \tag{5.3}$$

and the two air sides summed with the mixing row and the hall balance give the air balance of the two units together,

$$(C_1 + C_2)\,(T_r - T_c) = Q_t + Q_o(t), \tag{5.4}$$

by which the rise from the cold aisle to the return is set by the heat and the airflow alone, not by the coils.

Modeling choices at the graph. Four choices that no question owns make this graph. The hall is taken in steady state within each window: this makes every row of equation (5.2) algebraic, so that one window's readings fix one state, and it fails in the minutes after a load step, while the hall's air and equipment still store or release heat. The airflow is pinned at the units' fixed fan speed and nothing controls the discharge air within the window: this is what lets the discharge readings inform the coils, since a coil that passes less heat leaves its discharge warmer, and it fails under an active discharge controller, which would make the discharge report its set point and move a valve instead (Chapter 6 takes this up). The chilled-water supply $T_s(t)$ and the other heat $Q_o(t)$ are known inputs: the supply is read by a precise supply sensor, and the other heat is the hall's lighting and fan schedule with the envelope gain. This is what takes

the season out of the parameter, and it fails if either is wrong, because the error then goes into g. Each coil carries one effective conductance between the return air and the chilled water, Part I's single path split in two: this keeps one entry per coil, and it fails if a coil's heat transfer depends on its discharge temperature in a way one conductance cannot carry.

The record. Table 5.1 gives the two windows the chapter reads. Each row is a quantity; the first two are the known inputs of the window, and the other five the readings of the hall's sensors, with their standard deviations in the last column. Between the two windows the supply is 1.06 K colder and the other heat 2.12 kW lower, because June is winter in the southern hemisphere where the hall stands, while the meter reads a load 20.07 kW higher.

Table 5.1: The two windows of the condition record read in this chapter: the commissioning test and today.

Quantity [unit]	Window 0 (commissioning)	Window 18 (Jun Y2)	sd [K or kW]
Chilled-water supply $T_s(t)$, known input [°C]	14.00	12.94	–
Other heat $Q_o(t)$, known input [kW]	8.00	5.88	–
Return air y_{T_r} [°C]	22.71	23.48	0.3
Discharge air of unit 1, $y_{T_{d,1}}$ [°C]	16.86	16.54	0.3
Discharge air of unit 2, $y_{T_{d,2}}$ [°C]	17.08	17.30	0.3
Hall load meter y_Q [kW]	110.92	130.99	3.0
Rack-1 inlet y_{T_1} [°C]	18.08	18.26	0.3

The graph. The rows of the two windows are set side by side in Figure 5.1, window 0 on the left and window 18 on the right. In each copy the IT load Q_t (green) and the other heat $Q_o(t)$ feed the return air; the return air feeds the two coils (green, each carrying its free log-derate $g_{u,t}$), which pass heat to the pinned chilled water (gray) and send their discharge air to the cold aisle; the cold aisle feeds the inlet of rack 1. Blue nodes are temperatures fixed by the rows (5.2). The dashed arrows leave each copy for its five readings, with their values. One green node below the two copies, the recirculation fraction β_1 of rack 1, is linked to the rack-1 inlet of both windows. The rating prior is attached to the coils of window 0 only, and the survey prior to β_1.

The joint state. In the limit of exact rows the five rows of each window make its five temperatures functions of the free entries, and the fit runs on the reduced state of Section 4.2.8, here over both windows,

$$\mathbf{z} = (Q_0,\ g_{1,0},\ g_{2,0},\ Q_{18},\ g_{1,18},\ g_{2,18},\ \beta_1)^{\mathsf{T}}, \tag{5.5}$$

the loads in kW and the other five entries dimensionless; z_1 to z_7 in that order. The full state of the two windows adds each window's five temperatures, seventeen entries in all, under ten physical rows; its point on the set where the rows hold is equation (5.3) applied in each window at that window's inputs. Racks 2 and 3 are not drawn and their fractions are not entries: they enter no reading row of the record and no question of this chapter, and their block of the posterior is their survey prior alone, which separates from the rest (Section 5.4.4 shows the check). Leaving them out keeps every entry of the state in use, and it fails for a question about rack 2 or rack 3, which *Operations and Control by Derivation*, Chapter 5, answers on the same hall.

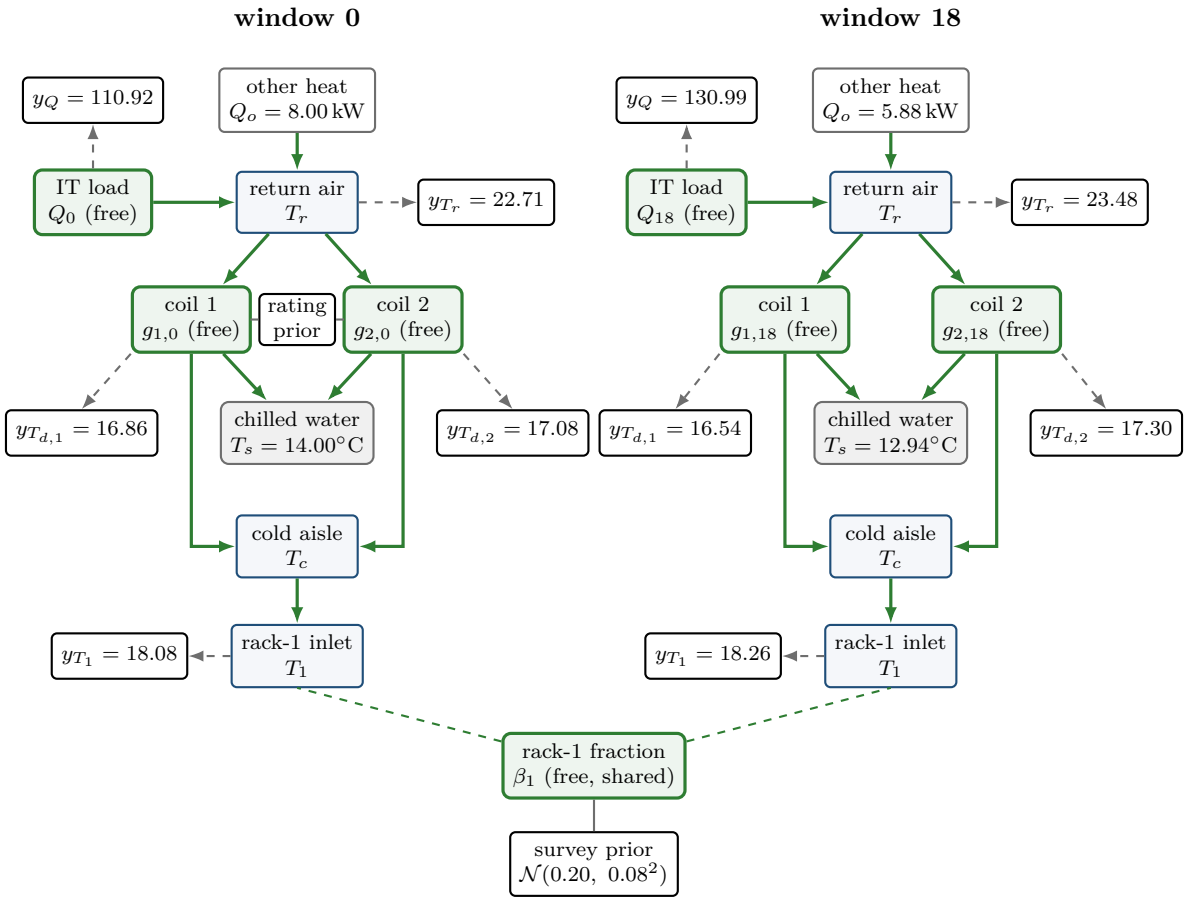


Figure 5.1: The rows of the two-unit hall in window 0 (left) and window 18 (right). Free (green): each window's IT load Q_t and the log-derates $g_{1,t}$, $g_{2,t}$ of its coils, and the recirculation fraction β_1 of rack 1, one entry shared by both windows (dashed green links). Blue: temperatures fixed by the physical rows (5.2). Known inputs of each window, pinned in its fit: the chilled-water supply $T_s(t)$ (gray) and the other heat $Q_o(t)$ (white); also pinned, not drawn, the ratings $\bar{G}_u = 7.0\,\mathrm{kW/K}$ and the capacity rates $C_u = 10.0\,\mathrm{kW/K}$. Observed (dashed): the return air, the two discharge temperatures and rack 1's inlet in °C, and the load meter in kW. Priors (solid gray): the rating prior on the two coils of window 0 only, and the survey prior on β_1.

5.1.3 The asset manager's questions (step 1)

The asset manager asks five questions of the two windows, and step 1 writes each as a probability over the joint state (5.5), with the unknown on the left of the bar and the ten readings and the priors on the right. Table 5.2 lists them. The first column is the question in the asset manager's words, the first of them the drill question; the second the general form of its category from Chapter 2, for a quantity $f(\mathbf{z})$ derived from the state and a level f^* of it; the third its mathematical form for the hall; the fourth its category; the fifth the computation of Section 2.3.6 that answers it; and the last the subsection where it is answered. "This coil" in the drill question is each coil in turn, because the state carries both. The second and third questions share a general form and one answer. The capacity question is a preview, today's margin only: when the margin is used up is Chapter 14's question. Which measurement would narrow the index asks nothing new of today's evidence; it is the next measurement for these five questions, Section 5.6.

Table 5.2: The questions asked of the two-unit hall at window 18 against its commissioning test. $\mathbf{y}$ is the ten readings and the three prior rows; $h_{u,18}$ the health index (5.1); $A_{u,t}$ the approach of unit u in window t; $G^* = 10.761\,\mathrm{kW/K}$ the capacity requirement.

Asset manager's question	General form	Mathematical form	Category	Computation	Answered in
How far has this coil's conductance fallen from its commissioned value?	$\Pr(f(\mathbf{z}) \le f^* \mid \mathbf{y})$; moments; credible interval, (2.9)–(2.11)	$p(h_{u,18} \mid \mathbf{y})$, $h_{u,18} = \mathrm{e}^{g_{u,18}-g_{u,0}}$, $u = 1, 2$: median and 5–95 % band	estimation	posterior	5.3.1
Has it lost any conductance at all, beyond what the readings' scatter allows?	$\Pr(f(\mathbf{z}) < f^* \mid \mathbf{y})$, (2.12)	$\Pr(g_{u,18} - g_{u,0} < 0 \mid \mathbf{y})$	prediction under a threshold	posterior	5.3.2
Which coil has fallen more, and how sure is that?	$\Pr(f(\mathbf{z}) < f^* \mid \mathbf{y})$, (2.12)	$\Pr(\ln h_{2,18} < \ln h_{1,18} \mid \mathbf{y})$	prediction under a threshold	posterior	5.3.2
Is the dashboard's 21 % and 29 % the coils, or the load and the season?	moments and credible interval of a derived quantity, (2.10), (2.11)	$p(h_u^{\mathrm{raw}} \mid \mathbf{y})$, $h_u^{\mathrm{raw}} = A_{u,0}(\mathbf{z})/A_{u,18}(\mathbf{z})$, and its value with $g_{\cdot,18} = g_{\cdot,0}$	estimation	posterior	5.3.3
Do the coils still carry the hall's capacity scenario today?	$\Pr(f(\mathbf{z}) \ge f^* \mid \mathbf{y})$, (2.12)	$\Pr(G_{1,18} + G_{2,18} \ge G^* \mid \mathbf{y})$	prediction under a threshold	posterior	5.3.4

5.1.4 The evidence as rows (step 2)

Step 2 attaches the evidence of Table 5.1 to the graph as rows on the joint state. Each window contributes five readings: the return air, the two discharge temperatures and rack 1's inlet, by sensors of standard deviation 0.3 K, and the hall load meter, of standard deviation 3.0 kW, which reads that window's load directly. Three priors carry what was known before the readings. The maker's rating puts each coil as built at its rating to within about 15 %, $g_{u,0} \sim \mathcal{N}(0,\ 0.15^2)$, a row on each commissioning entry; the airflow survey gives $\beta_1 \sim \mathcal{N}(0.20,\ 0.08^2)$. The loads have no prior row, because the meter reads them, and today's log-derates have none.

Which row acts on which entry. On the reduced state a reading row sees the free entries through the row of $\mathbf{N}_\mathrm{b}$ of the temperature it reads, equation (4.47). By equation (5.3) the return air and the two discharge temperatures of window t depend on $(Q_t, g_{1,t}, g_{2,t})$, rack 1's

inlet on those three and β_1, and the meter on Q_t alone. With $*$ for an entry that can be nonzero and the columns in the order of equation (5.5), the 13×7 Jacobian has the pattern

$$\mathbf{J} = \begin{bmatrix} * & * & * & 0 & 0 & 0 & 0 \\ * & * & * & 0 & 0 & 0 & 0 \\ * & * & * & 0 & 0 & 0 & 0 \\ * & 0 & 0 & 0 & 0 & 0 & 0 \\ * & * & * & 0 & 0 & 0 & * \\ 0 & 0 & 0 & * & * & * & 0 \\ 0 & 0 & 0 & * & * & * & 0 \\ 0 & 0 & 0 & * & * & * & 0 \\ 0 & 0 & 0 & * & 0 & 0 & 0 \\ 0 & 0 & 0 & * & * & * & * \\ 0 & * & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & * & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & * \end{bmatrix} \begin{array}{c} \text{window 0: return air } y_{T_r} \\ \text{discharge } y_{T_{d,1}} \\ \text{discharge } y_{T_{d,2}} \\ \text{load meter } y_Q \\ \text{rack-1 inlet } y_{T_1} \\ \text{window 18: return air } y_{T_r} \\ \text{discharge } y_{T_{d,1}} \\ \text{discharge } y_{T_{d,2}} \\ \text{load meter } y_Q \\ \text{rack-1 inlet } y_{T_1} \\ \text{rating prior on } g_{1,0} \\ \text{rating prior on } g_{2,0} \\ \text{survey prior on } \beta_1 \end{array}, \tag{5.6}$$

$$\mathbf{g}_{\ln h_1}^{\mathsf{T}} = (0,\ -1,\ 0,\ 0,\ 1,\ 0,\ 0), \qquad \mathbf{g}_{\ln h_2}^{\mathsf{T}} = (0,\ 0,\ -1,\ 0,\ 0,\ 1,\ 0).$$

The second line is not a row: it gives the gradients of the two questions' quantities $\ln h_{u,18}$, from equation (5.1). Two facts of the graph are read here, before any number. The window-0 rows touch only the first three columns and the window-18 rows only the next three, apart from the rack-1 row of each window, which also touches β_1: the two windows meet in the column of β_1 and nowhere else. And each index's gradient reaches across that block structure, one entry in each window's block, so its variance takes one variance from each window.

Modeling choices at the rows. Four choices make these rows. Each reading is taken as Gaussian, unbiased and independent of the others given the state, the reading rows of Chapter 4: this makes each reading one normalized row and the objective a sum of squares, and it fails if a sensor drifts, which is the question of Chapter 7. The two windows share one recirculation fraction for rack 1, because the racks were not rearranged between handover and today: this lets the rack-1 reading of both windows inform one fraction, and it fails if blanking panels or rack positions changed, when each window would carry its own β_1. The rating prior goes on the commissioning entries only: the maker's rating, to about 15 %, describes the coil as built, and put on today's entries it would pull each of them toward the rating: the index of the coil farther from the rating would move toward no change, and through the anticorrelation of the two coils within a window the other coil's could move the opposite way. Today's entries are fixed by today's readings alone, and Section 5.4.3 sets the other choice beside this one. The rows carry no trend rows: Chapter 4's trend rows (4.68) tie windows together, but with a free rate per coil each rate absorbs its row, Chapter 4's finding that the two windows stay independent ($\Sigma_{12} = 0$ in equation (4.72)), so they would add two entries and two rows and leave the seven entries' posterior where it is (Section 5.4.4 shows it). The rows that tie windows exist, and this chapter needs no tie beyond the parameter's own two entries; the rate is Chapter 10's.

5.2 The posterior of the state (step 3)

All five questions of Table 5.2 are read from one posterior, that of the joint state (5.5) under the thirteen rows of step 2, with the known inputs of Table 5.1 and the pinned quantities of Section 5.1.2.

5.2.1 The rows on the joint state

The rows of the reduced state are the evidence rows at the point of the set where the physical rows hold, equation (4.46), with each window's temperatures from equation (5.3) at that window's inputs. Write $\mathbf{z}_t = (Q_t, g_{1,t}, g_{2,t}, \beta_1)$ for the entries window t sees. Each window gives five rows, each divided by its standard deviation,

$$\mathbf{r}_t(\mathbf{z}_t) = \begin{bmatrix} (T_r(\mathbf{z}_t) - y_{T_r,t})/0.3 \\ (T_{d,1}(\mathbf{z}_t) - y_{T_{d,1},t})/0.3 \\ (T_{d,2}(\mathbf{z}_t) - y_{T_{d,2},t})/0.3 \\ (Q_t - y_{Q,t})/3.0 \\ (T_1(\mathbf{z}_t) - y_{T_1,t})/0.3 \end{bmatrix}, \qquad t = 0, 18, \tag{5.7}$$

and the residual vector stacks the two windows and the three priors in the order of the pattern (5.6),

$$\mathbf{r}(\mathbf{z}) = \begin{bmatrix} \mathbf{r}_0(\mathbf{z}_0) \\ \mathbf{r}_{18}(\mathbf{z}_{18}) \\ g_{1,0}/0.15 \\ g_{2,0}/0.15 \\ (\beta_1 - 0.20)/0.08 \end{bmatrix}, \qquad \mathcal{L}(\mathbf{z}) = \tfrac{1}{2}\,\mathbf{r}(\mathbf{z})^{\mathsf{T}}\mathbf{r}(\mathbf{z}), \tag{5.8}$$

thirteen dimensionless rows on seven entries. The reduced Jacobian is $\mathbf{J} = \mathbf{J}_{\text{rows}}\mathbf{N}$ of equation (4.47), formed window by window: each reading row of window t is its temperature's row of that window's $\mathbf{N}_{\text{b}}$ over the sensor's 0.3 K, the meter row is $1/3.0$ in the column of Q_t, and each prior row is the reciprocal of its standard deviation in its own column. The rows are nonlinear in the state through $\mathrm{e}^{g_{u,t}}$ and the product $\beta_1(T_r - T_c)$, so $\mathbf{J}$ is evaluated at each iterate, as the paragraph on nonlinear physical rows of Section 4.2.8 describes.

5.2.2 The normal equations and the covariance

The MAP state $\hat{\mathbf{z}}$ minimizes the objective (5.8). Each Gauss–Newton iteration solves the normal equations of equation (4.29) on the reduced rows,

$$(\mathbf{J}^{\mathsf{T}}\mathbf{J})\,\Delta\mathbf{z} = -\mathbf{J}^{\mathsf{T}}\mathbf{r}(\mathbf{z}), \qquad \mathbf{z}^{(l+1)} = \mathbf{z}^{(l)} + \Delta\mathbf{z}, \tag{5.9}$$

with $\mathbf{J}$ and $\mathbf{r}$ at $\mathbf{z}^{(l)}$ and each window's temperatures solved again by equation (5.3). The iteration starts from the state the record suggests, the meter's loads, the coils at their ratings and the survey's fraction, $\mathbf{z}^{(0)} = (110.92,\ 0,\ 0,\ 130.99,\ 0,\ 0,\ 0.20)^{\mathsf{T}}$. With every row normalized the information matrix is $\mathbf{\Lambda} = \mathbf{J}^{\mathsf{T}}\mathbf{J}$; with unnormalized rows and their weights it is $\mathbf{J}^{\mathsf{T}}\mathbf{W}\mathbf{J}$. At the minimum the Laplace covariance is

$$\mathbf{\Sigma} = (\mathbf{J}^{\mathsf{T}}\mathbf{J})^{-1}\big|_{\hat{\mathbf{z}}}, \tag{5.10}$$

the Gauss–Newton form of equation (4.38), which drops the curvature term $\mathbf{C}$ of equation (4.44); Section 5.4.2 sets the full Hessian beside it. By equation (4.48) this $\mathbf{\Sigma}$ is the free-entry block of the seventeen-entry full state's covariance in the limit of exact rows.

The block structure. The pattern (5.6) fixes the shape of $\mathbf{\Lambda}$ before any number. A product of two columns of $\mathbf{J}$ is zero unless some row touches both, and only the rack-1 row of each window and the survey prior touch β_1, so

$$\mathbf{\Lambda} = \begin{bmatrix} \mathbf{\Lambda}_{0,0} & \mathbf{0} & \mathbf{\Lambda}_{0,\beta} \\ \mathbf{0} & \mathbf{\Lambda}_{18,18} & \mathbf{\Lambda}_{18,\beta} \\ \mathbf{\Lambda}_{0,\beta}^{\mathsf{T}} & \mathbf{\Lambda}_{18,\beta}^{\mathsf{T}} & \Lambda_{\beta\beta} \end{bmatrix}, \tag{5.11}$$

with $\mathbf{\Lambda}_{t,t}$ the 3×3 block of window t's entries $(Q_t, g_{1,t}, g_{2,t})$, $\mathbf{\Lambda}_{t,\beta}$ their coupling to β_1 through the rack-1 row, and $\Lambda_{\beta\beta}$ the information on β_1 from both rack-1 rows and the survey. The zero blocks say that no row reads both windows, so that given β_1 the two windows are independent. With β_1 integrated out they are not: each window's entries move with β_1, and the uncertainty of the β_1 both windows share leaves a covariance between a window-0 entry and a window-18 entry, the smaller the better β_1 is known.

One state over two windows. Chapter 2's commissioning row, $(g - \hat{g}_0)/\sigma_{\hat{g}_0}$, carries an earlier fit's marginal into a later one, and it is the right tool when the earlier record is no longer at hand. Here it is at hand, and the commissioning test is a posterior like today's. The joint state makes the commissioning entries entries of the one posterior that every answer is read from, so their covariance with today's entries, and with each other, is computed rather than assumed. A row carrying the commissioning fit forward reproduces that posterior only if it carries the full covariance of every entry the windows share: here the two commissioning log-derates, which window 0's return-air reading leaves negatively correlated, and β_1, which both windows' rack-1 rows inform. Marginal rows on $g_{1,0}$ and $g_{2,0}$ alone drop that correlation and that information. Section 5.4.4 sets both forms against the joint state.

5.2.3 The fit of the two windows

The iterations. Table 5.3 runs the normal equations (5.9) from $\mathbf{z}^{(0)}$. Each row is an iteration, with the objective and the norm of its gradient $\mathbf{J}^{\mathsf{T}}\mathbf{r}$. The objective falls from 25.838001 to 0.511964 in the first step and to 0.492426 in the second, and the gradient is below 10^{-4} from the third.

Table 5.3: Gauss–Newton iterations of the joint state of windows 0 and 18, equation (5.9), from the meter's loads, the rated coils and the survey's fraction.

Iteration l [–]	Objective $\mathcal{L}(\mathbf{z}^{(l)})$ [–]	$\lVert\mathbf{J}^{\mathsf{T}}\mathbf{r}\rVert$ [–]
0	25.838001	3.306×10^{2}
1	0.511964	1.030×10^{1}
2	0.492426	1.353×10^{-2}
3	0.492426	$< 10^{-4}$
4	0.492426	$< 10^{-4}$

The Jacobian at the MAP state. With the columns of equation (5.5) and the rows in the order of the pattern (5.6),

$$\mathbf{J}(\hat{\mathbf{z}}) = \begin{bmatrix} 0.2480 & -14.9336 & -14.4459 & 0 & 0 & 0 & 0 \\ 0.0785 & -24.8037 & -4.5758 & 0 & 0 & 0 & 0 \\ 0.0841 & -5.0635 & -24.3160 & 0 & 0 & 0 & 0 \\ 0.3333 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.1153 & -14.9336 & -14.4459 & 0 & 0 & 0 & 19.7457 \\ 0 & 0 & 0 & 0.2624 & -18.8479 & -16.8534 & 0 \\ 0 & 0 & 0 & 0.0864 & -30.1496 & -5.5517 & 0 \\ 0 & 0 & 0 & 0.1051 & -7.5461 & -28.1552 & 0 \\ 0 & 0 & 0 & 0.3333 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.1298 & -18.8479 & -16.8534 & 22.6742 \\ 0 & 6.6667 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 6.6667 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 12.5 \end{bmatrix}, \tag{5.12}$$

the load columns in kW^{-1} and the others dimensionless. The zeros sit where the pattern places them. Each discharge row has its largest entry in its own coil's column, and the return-air row sees both coils nearly alike, because the return air depends on their sum. The rack-1 row repeats the return-air row in the log-derate columns: by the air balance (5.4) a derate leaves the rise unchanged and moves the inlet as it moves the return air. Window 18's lines are those of window 0 in its own columns, with larger slopes because the heat to remove is larger. The ten reading rows have rank 7 of 7: each window is observable from its own readings, and today's entries need no prior.

The information. The information matrix has the block form (5.11),

$$\mathbf{J}^{\mathsf{T}}\mathbf{J} = \begin{bmatrix} 0.199 & -7.800 & -7.652 & 0 & 0 & 0 & 2.277 \\ -7.800 & 1131.332 & 668.080 & 0 & 0 & 0 & -294.874 \\ -7.652 & 668.080 & 1074.021 & 0 & 0 & 0 & -285.244 \\ 0 & 0 & 0 & 0.215 & -10.791 & -10.048 & 2.942 \\ 0 & 0 & 0 & -10.791 & 1676.424 & 1015.148 & -427.360 \\ 0 & 0 & 0 & -10.048 & 1015.148 & 1391.612 & -382.138 \\ 2.277 & -294.874 & -285.244 & 2.942 & -427.360 & -382.138 & 1060.260 \end{bmatrix},$$

with the two window blocks on the diagonal, zeros between them, and the coupling to β_1 in the last row and column.

The covariance. Its inverse, the Laplace covariance (5.10), is given here in the blocks of (5.11), the rows and columns of each window in the order $(Q_t, g_{1,t}, g_{2,t})$, the load entries in kW and kW^2 and the others dimensionless:

$$\boldsymbol{\Sigma}_{0,0} = \begin{bmatrix} 7.5965 & 3.2638 \times 10^{-2} & 3.4435 \times 10^{-2} \\ 3.2638 \times 10^{-2} & 1.5696 \times 10^{-3} & -6.8900 \times 10^{-4} \\ 3.4435 \times 10^{-2} & -6.8900 \times 10^{-4} & 1.6598 \times 10^{-3} \end{bmatrix},$$

$$\boldsymbol{\Sigma}_{18,18} = \begin{bmatrix} 7.7031 & 2.9026 \times 10^{-2} & 3.5372 \times 10^{-2} \\ 2.9026 \times 10^{-2} & 1.2080 \times 10^{-3} & -6.1601 \times 10^{-4} \\ 3.5372 \times 10^{-2} & -6.1601 \times 10^{-4} & 1.4797 \times 10^{-3} \end{bmatrix},$$

$$\boldsymbol{\Sigma}_{0,18} = \begin{bmatrix} 6.5092 \times 10^{-3} & 3.9062 \times 10^{-4} & 3.9598 \times 10^{-4} \\ 5.8158 \times 10^{-4} & 3.4901 \times 10^{-5} & 3.5380 \times 10^{-5} \\ 5.8124 \times 10^{-4} & 3.4880 \times 10^{-5} & 3.5359 \times 10^{-5} \end{bmatrix},$$

$$\boldsymbol{\Sigma}_{0,\beta} = \begin{bmatrix} 2.3086 \times 10^{-3} \\ 2.0626 \times 10^{-4} \\ 2.0614 \times 10^{-4} \end{bmatrix}, \qquad \boldsymbol{\Sigma}_{18,\beta} = \begin{bmatrix} 3.3760 \times 10^{-3} \\ 2.0260 \times 10^{-4} \\ 2.0538 \times 10^{-4} \end{bmatrix}, \qquad \mathrm{Var}(\beta_1 \mid \mathbf{y}) = 1.1973 \times 10^{-3},$$

with $\boldsymbol{\Sigma}_{0,18}$ the covariance of window 0's entries (rows) with window 18's (columns), and $\boldsymbol{\Sigma}_{t,\beta}$ the covariance of window t's entries with β_1. The correlations read by sign, and the rows fix each sign. The two log-derates of one window are correlated negatively, -0.4269 at window 0 and -0.4607 at window 18: the return air sees the sum $G_1 + G_2$ and each discharge sees its own coil, all at 0.3 K, so a higher g_1 is paid for by a lower g_2 at the same return air, and the readings know the hall's total better than its split. Each load is correlated positively with its window's log-derates, about $+0.30$: a larger load needs a larger conductance to give the same rise of the return air over the supply. The log-derates are correlated positively with β_1, about $+0.15$: a larger fraction warms the rack-1 inlet, which a cooler cold aisle, that is, larger conductances, offsets. And a coil's two log-derates are correlated by only $+0.0253$ (coil 1) and $+0.0226$ (coil 2), through β_1 alone, the one shared column: the two windows are nearly independent.

The MAP state. Table 5.4 gives the MAP state with the posterior spreads $\sqrt{\Sigma_{ii}}$ and the prior spread of each entry; the last column names the rows that inform it. Today's log-derates, with no prior, are known better than the commissioning entries, because today's heat is larger and the slopes of equation (5.12) with it. The fraction β_1 is known to 0.03460, against 0.04711 from window 0's rows and the survey alone: the second window's rack-1 reading is the gain from sharing it. The conductances are $G_{1,0} = 6.8325$, $G_{2,0} = 6.6093$, $G_{1,18} = 6.7059$ and $G_{2,18} = 5.9963\,\mathrm{kW/K}$. The fitted discharge of unit 2 at window 18, which Exercise 5.5 reads against a reference sensor, is 17.2275°C.

Table 5.4: The MAP state of the joint state of windows 0 and 18, with the posterior and prior spreads of each entry.

Entry [unit]	$\hat{z}_i$ [entry's]	$\sqrt{\Sigma_{ii}}$ [entry's]	Prior sd [entry's]	Informed by
Q_0 [kW]	110.47391	2.75617	none	window 0's meter and temperature readings
$g_{1,0}$ [–]	−0.02423	0.03962	0.15	window 0's readings; the rating prior
$g_{2,0}$ [–]	−0.05743	0.04074	0.15	window 0's readings; the rating prior
Q_{18} [kW]	130.16652	2.77545	none	window 18's meter and temperature readings
$g_{1,18}$ [–]	−0.04293	0.03476	none	window 18's readings
$g_{2,18}$ [–]	−0.15477	0.03847	none	window 18's readings
β_1 [–]	0.20403	0.03460	0.08	both rack-1 inlet readings; the survey

The residuals. At the MAP state twice the objective is $2\mathcal{L}(\hat{\mathbf{z}}) = 0.9849$ on $13 - 7 = 6$ degrees of freedom, and $\Pr(\chi^2_6 > 0.98) = 0.986$: the readings sit well inside their errors, the luck of this record's draws rather than a property of the sensors. The largest normalized residual is window 18's return air, 0.5658.

5.2.4 The full posterior by reweighted draws

The Laplace normal $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ is the reported posterior. Beside it stands the full posterior of the same rows, without the Gaussian approximation, computed by reweighted draws (importance sampling) with the Laplace normal as the proposal, the route that Table 4.1 lists. The procedure is the following.

1. Draw N_d states $\mathbf{z}^{(j)} \sim \mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$, $j = 1, \dots, N_\mathrm{d}$, each through the Cholesky factor of $\boldsymbol{\Sigma}$ as in Section 4.4.7.
2. Weight each draw by the ratio of the posterior to the proposal, both known up to a constant,

$$\varpi_j \propto \frac{\exp\!\left[-\mathcal{L}(\mathbf{z}^{(j)})\right]}{p_\mathrm{L}(\mathbf{z}^{(j)})}, \qquad p_\mathrm{L}(\mathbf{z}) \propto \exp\!\left[-\tfrac{1}{2}(\mathbf{z}-\hat{\mathbf{z}})^\mathsf{T}\boldsymbol{\Sigma}^{-1}(\mathbf{z}-\hat{\mathbf{z}})\right], \qquad \sum_j \varpi_j = 1, \quad (5.13)$$

   with $\mathcal{L}$ the objective (5.8) evaluated through the exact forward equations (5.3).
3. Read a probability as $\sum_j \varpi_j \, \mathbf{1}[\text{event at } \mathbf{z}^{(j)}]$, a mean as $\sum_j \varpi_j f(\mathbf{z}^{(j)})$, and a quantile from the weighted cumulative distribution of the $f(\mathbf{z}^{(j)})$.
4. State the effective sample size $N_\mathrm{eff} = 1/\sum_j \varpi_j^2$, which equals N_d when the proposal is the posterior, and the Monte Carlo standard error of each result by batch means: the draws split into 20 batches, the result computed in each with its own normalized weights, and the spread of the 20 results divided by $\sqrt{20}$.

The result is the full posterior of the stated model up to its Monte Carlo error, and it differs from the Laplace answer by whatever the Gaussian approximation and the dropped curvature term leave out. For this record, $N_{\mathrm{d}} = 200\,000$ draws from a seeded generator give $N_{\mathrm{eff}} = 197\,321$, 98.7 % of N_{d}, and a largest weight of 3.29×10^{-5}: the Laplace normal is close to the posterior, and no draw dominates.

5.3 The answers (step 4)

Every answer is read from the pair $(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ of Section 5.2.3, and that value is the one reported. A derived quantity f is carried by its gradient $\mathbf{g} = \partial f / \partial \mathbf{z}$ at $\hat{\mathbf{z}}$, with variance $\mathbf{g}^{\mathsf{T}} \boldsymbol{\Sigma} \mathbf{g}$. The full posterior of Section 5.2.4 stands beside each value, a median beside a median.

5.3.1 The health index of each coil

Asset manager's question. How far has this coil's conductance fallen from its commissioned value? It is asked of each coil in turn.

General form. An estimation question about a derived quantity asks for its posterior distribution, equation (2.9),

$$\Pr(f(\mathbf{z}) \le f^* \mid \mathbf{y}) = \int \mathbf{1}\big[f(\mathbf{z}) \le f^*\big]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z},$$

summarized by its moments, equation (2.10), and by a credible interval, equation (2.11), here with the levels 0.05 and 0.95 in place of 0.025 and 0.975, the 5–95 % band of Chapter 4.

Specialized form. The question is equations (2.9), (2.10) and (2.11) with $\mathbf{z}$ the joint state (5.5), $f(\mathbf{z}) = \ln h_{u,18}$ of equation (5.1), that is $z_5 - z_2$ for coil 1 and $z_6 - z_3$ for coil 2, and $p(\mathbf{z} \mid \mathbf{y})$ the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ of Section 5.2.2. One choice enters here: the index is carried as its logarithm, which the Laplace route makes normal. This is what makes the answer a closed form in $\hat{\mathbf{z}}$ and $\boldsymbol{\Sigma}$, and it is accurate only if the posterior of $\ln h$ is close to normal; the full posterior beside it tests that. The quantity is linear in the state, so its gradient is the second line of the pattern (5.6) and its variance, by the delta method (4.53), is exact under the normal. For coil 1, with the entries numbered as in equation (5.5),

$$\widehat{\ln h}_{1,18} = \hat{z}_5 - \hat{z}_2, \qquad \mathrm{Var}(\ln h_{1,18} \mid \mathbf{y}) = \mathbf{g}_{\ln h_1}^{\mathsf{T}} \boldsymbol{\Sigma}\, \mathbf{g}_{\ln h_1} = \Sigma_{22} + \Sigma_{55} - 2\Sigma_{25}, \tag{5.14}$$

the change-since-commissioning variance (3.38) of the coil's two entries, and for coil 2 the same with the entries z_3 and z_6. The three terms are the budget of the index: the commissioning entry's variance, today's, and the cross term of the two, whose sign is that of $-\Sigma_{25}$. Quantiles pass unchanged through the increasing map $h = \mathrm{e}^{\ln h}$ (Section 4.3.3), so the index has the median and the band

$$\mathrm{med}(h_{u,18} \mid \mathbf{y}) = \mathrm{e}^{\widehat{\ln h}_{u,18}}, \qquad \Big[\mathrm{e}^{\widehat{\ln h}_{u,18} - 1.6449\,\mathrm{sd}(\ln h_{u,18} \mid \mathbf{y})},\ \mathrm{e}^{\widehat{\ln h}_{u,18} + 1.6449\,\mathrm{sd}(\ln h_{u,18} \mid \mathbf{y})}\Big], \tag{5.15}$$

with $\mathrm{sd}(\ln h_{u,18} \mid \mathbf{y})$ the square root of the variance (5.14) and 1.6449 the 95 % quantile of the standard normal. The band is asymmetric in h, wider above the median than below it. Its density is a lognormal by the change of variables (4.55),

$$p(h_{u,18} \mid \mathbf{y}) = \frac{1}{h_{u,18}\,\mathrm{sd}(\ln h_{u,18} \mid \mathbf{y})}\, \varphi\!\left(\frac{\ln h_{u,18} - \widehat{\ln h}_{u,18}}{\mathrm{sd}(\ln h_{u,18} \mid \mathbf{y})}\right).$$

Building blocks. The MAP state $\hat{\mathbf{z}}$ and the covariance $\boldsymbol{\Sigma}$ of Section 5.2.3; the gradients of the pattern (5.6); the variance (5.14); the quantiles of equation (5.15).

The index of each coil. For coil 1, equation (5.14) gives $\widehat{\ln h}_{1,18} = -0.04293 - (-0.02423) = -0.01870$ and

$$\mathrm{Var}(\ln h_{1,18} \mid \mathbf{y}) = 1.5696 \times 10^{-3} + 1.2080 \times 10^{-3} - 2\,(3.4901 \times 10^{-5}) = 2.7078 \times 10^{-3} = 0.05204^2.$$

By equation (5.15) the index has median 0.9815 and 5–95 % band [0.9010, 1.0692]: a median loss of 1.85 %, with the band running from a loss of 9.90 % to a gain of 6.92 %. The full posterior gives a median of 0.9815 and a band of [0.9008, 1.0699]. For coil 2, $\widehat{\ln h}_{2,18} = -0.15477 - (-0.05743) = -0.09734$ and

$$\mathrm{Var}(\ln h_{2,18} \mid \mathbf{y}) = 1.6598 \times 10^{-3} + 1.4797 \times 10^{-3} - 2\,(3.5359 \times 10^{-5}) = 3.0688 \times 10^{-3} = 0.05540^2,$$

so the index has median 0.9072 and band [0.8282, 0.9938]: a median loss of 9.28 %, with the band running from a loss of 0.62 % to one of 17.18 %. The full posterior gives 0.9070 and [0.8279, 0.9939].

The two lognormal densities are plotted in Figure 5.2: the horizontal axis is the health index against commissioning, the vertical axis its posterior density; the shaded areas are the 5–95 % bands, the solid gray line marks $h = 1$, no change, the dashed lines the dashboard's two tiles, and the dotted lines the stated truth of the record, check markers for Section 5.4.5. Coil 1's density straddles $h = 1$; coil 2's lies almost wholly below it. The dashboard's two values lie far below both densities.

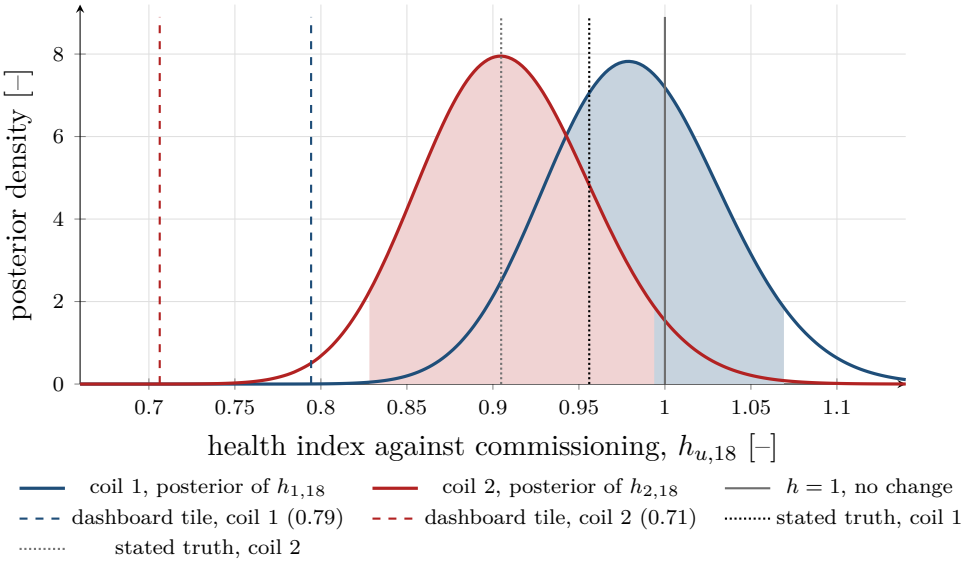


Figure 5.2: The posterior densities of the health index of each coil at window 18 against commissioning, the lognormals of equation (5.15): coil 1 with $\ln h \sim \mathcal{N}(-0.01870,\ 0.05204^2)$, coil 2 with $\ln h \sim \mathcal{N}(-0.09734,\ 0.05540^2)$. Shaded: the 5–95 % bands. Solid gray: no change. Dashed: the dashboard's raw-performance tiles. Dotted: the stated truth of the record, a check marker (Section 5.4.5).

The budget of the variance. Table 5.5 splits each variance into the three terms of equation (5.14); each row is a term, with its value and its share of the total for each coil. The shares appear as bars in Figure 5.3, the horizontal axis in percent of the variance.

More than half of each index's variance is the commissioning test's own: 58.0 % for coil 1 and 54.1 % for coil 2, against 44.6 and 48.2 % for today's entries. The cross term is small and negative, −2.6 and −2.3 %: the two entries of a coil are correlated positively through β_1, and the difference cancels the part they share. Read against rating, the derate factors $d_{u,t} = \mathrm{e}^{\hat{g}_{u,t}}$

Table 5.5: The variance budget of the logarithm of each health index, equation (5.14), at the MAP state.

Term	Coil 1 [–]	Share [%]	Coil 2 [–]	Share [%]
Commissioning entry, $\mathrm{Var}(g_{u,0} \mid \mathbf{y})$	1.5696×10^{-3}	58.0	1.6598×10^{-3}	54.1
Today's entry, $\mathrm{Var}(g_{u,18} \mid \mathbf{y})$	1.2080×10^{-3}	44.6	1.4797×10^{-3}	48.2
Cross term, $-2\,\mathrm{Cov}(g_{u,0}, g_{u,18} \mid \mathbf{y})$	-6.9801×10^{-5}	-2.6	-7.0718×10^{-5}	-2.3
Total, $\mathrm{Var}(\ln h_{u,18} \mid \mathbf{y})$	2.7078×10^{-3}	100	3.0688×10^{-3}	100

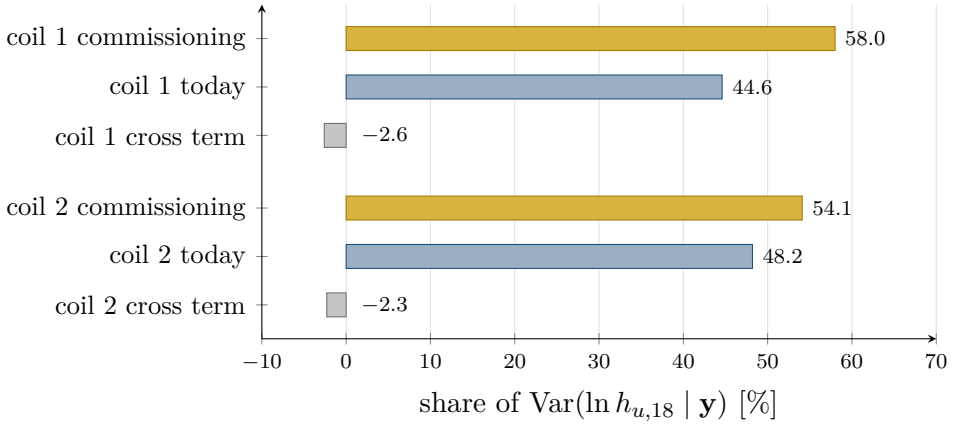


Figure 5.3: The variance budget of each coil's health index (Table 5.5): the commissioning entry's term (gold), today's entry's term (blue) and the cross term (gray), in percent of $\mathrm{Var}(\ln h_{u,18} \mid \mathbf{y})$. The gold bars are the part of the index that only the commissioning test can set.

are 0.9761 at commissioning and 0.9580 today for coil 1, and 0.9442 and 0.8566 for coil 2; the capacity question of Section 5.3.4 uses today's.

5.3.2 Loss beyond the scatter, and the coil that fell more

Asset manager's questions. Has it lost any conductance at all, beyond what the readings' scatter allows? Which coil has fallen more, and how sure is that? The first asks whether each index is below one with a probability the record supports; the second whether the dashboard's ranking of the two coils is right.

General form. A prediction under a threshold is a tail of a derived quantity under the posterior, equation (2.12),

$$\Pr(f(\mathbf{z}) \geq f^* \mid \mathbf{y}) = \int \mathbf{1}[f(\mathbf{z}) \geq f^*]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}.$$

Specialized form. Both questions are equation (2.12) with $p(\mathbf{z} \mid \mathbf{y})$ the Laplace normal of Section 5.2.2. For the second question, $f(\mathbf{z}) = -\ln h_{u,18}$ and $f^* = 0$: the event is a loss, $\ln h_{u,18} < 0$, on the gradient of the pattern (5.6), and with the variance (5.14),

$$\Pr(\ln h_{u,18} < 0 \mid \mathbf{y}) = \Phi\left(\frac{-\widehat{\ln h}_{u,18}}{\mathrm{sd}(\ln h_{u,18} \mid \mathbf{y})}\right). \tag{5.16}$$

One choice enters here: "beyond the scatter" is read as the posterior probability of any loss, the level $f^* = 0$. This is what makes the question one the record can answer without a decision rule,

and it is not a decision: a loss large enough to act on needs a decision level, Chapter 19's question, and the assurance level κ of equation (2.12) is not set here. For the third question, $f(\mathbf{z}) = \ln h_{1,18} - \ln h_{2,18}$ and $f^* = 0$: the event is that coil 2 has fallen more, $D_h = \ln h_{2,18} - \ln h_{1,18} < 0$, a linear quantity with the gradient $\mathbf{g}_{D_h} = \mathbf{g}_{\ln h_2} - \mathbf{g}_{\ln h_1} = (0,\ 1,\ -1,\ 0,\ -1,\ 1,\ 0)^\mathsf{T}$, and

$$\begin{aligned} \mathrm{Var}(D_h \mid \mathbf{y}) &= \mathrm{Var}(\ln h_{2,18} \mid \mathbf{y}) + \mathrm{Var}(\ln h_{1,18} \mid \mathbf{y}) - 2\,\mathrm{Cov}(\ln h_1, \ln h_2), \\ \mathrm{Cov}(\ln h_1, \ln h_2) &= \mathbf{g}_{\ln h_1}^\mathsf{T} \boldsymbol{\Sigma}\, \mathbf{g}_{\ln h_2} = \Sigma_{56} + \Sigma_{23} - \Sigma_{53} - \Sigma_{26}, \end{aligned} \tag{5.17}$$

with $\Pr(D_h < 0 \mid \mathbf{y}) = \Phi(-\hat{D}_h / \sqrt{\mathrm{Var}(D_h \mid \mathbf{y})})$. The covariance of the two indices collects the within-window covariances of the two coils, Σ_{56} today and Σ_{23} at commissioning, both negative by Section 5.2.3, less the two small cross-window terms.

Building blocks. The pair $(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ of Section 5.2.3; the indices and their variances of Section 5.3.1; the gradient $\mathbf{g}_{D_h}$ and the covariance (5.17); the normal distribution function Φ.

Loss beyond the scatter. By equation (5.16),

$$\begin{aligned} \Pr(\ln h_{1,18} < 0 \mid \mathbf{y}) &= \Phi(0.0187/0.0520) = \Phi(0.3593) = 0.640, \\ \Pr(\ln h_{2,18} < 0 \mid \mathbf{y}) &= \Phi(0.0973/0.0554) = \Phi(1.7572) = 0.961, \end{aligned}$$

with 0.63976 and 0.96007 from the full posterior (Monte Carlo standard errors 0.00121 and 0.00063). The record cannot say that coil 1 has lost anything; for coil 2 a loss is likely.

The coil that fell more. By equation (5.17), with the entries of $\boldsymbol{\Sigma}$ of Section 5.2.3,

$$\begin{aligned} \mathrm{Cov}(\ln h_1, \ln h_2) &= -6.1601 \times 10^{-4} - 6.8900 \times 10^{-4} - 3.4880 \times 10^{-5} - 3.5380 \times 10^{-5} \\ &= -1.3753 \times 10^{-3}, \end{aligned}$$

a correlation of -0.4771, and

$$\mathrm{Var}(D_h \mid \mathbf{y}) = 3.0688 \times 10^{-3} + 2.7078 \times 10^{-3} + 2.7505 \times 10^{-3} = 8.5271 \times 10^{-3} = 0.09234^2.$$

With $\hat{D}_h = -0.07864$,

$$\Pr(\ln h_{2,18} < \ln h_{1,18} \mid \mathbf{y}) = \Phi(0.8516) = 0.803,$$

and the full posterior gives 0.802758 (standard error 0.000847). The ratio of the two indices, $h_{2,18}/h_{1,18} = \mathrm{e}^D$, has median 0.9244 and 5–95 % band [0.7941, 1.0760]. The cross term is what sets this answer. It comes from the anticorrelation of the two coils within each window, and because the gradient $\mathbf{g}_{D_h}$ takes the two indices with opposite signs, it adds to the variance of the difference, which is wider than either index's: 0.0923 against 0.0520 and 0.0554. The record knows each coil's change better than it knows which coil changed more.

The joint posterior of the two logarithms is shown in Figure 5.4. The horizontal axis is $\ln h_{1,18}$ and the vertical axis $\ln h_{2,18}$; the two ellipses are the contours at one and two standard deviations of the Laplace normal, tilted by the correlation -0.4771; the diagonal is $\ln h_2 = \ln h_1$, and the shaded region below it is the event of the third question; the thin gray lines are the zeros of the two axes, the edges of the second question's events; the dot is the MAP pair and the cross the stated truth. Most of both ellipses lies below the diagonal; the inner one lies wholly below the horizontal zero line, coil 2's loss, and straddles the vertical one, coil 1's.

5.3.3 The dashboard's index as a quantity of the state

Asset manager's question. Is the dashboard's 21 % and 29 % the coils, or the load and the season? The question asks what the dashboard's raw-performance index measures, and how much of the fall it shows would be there if the coils had not aged at all.

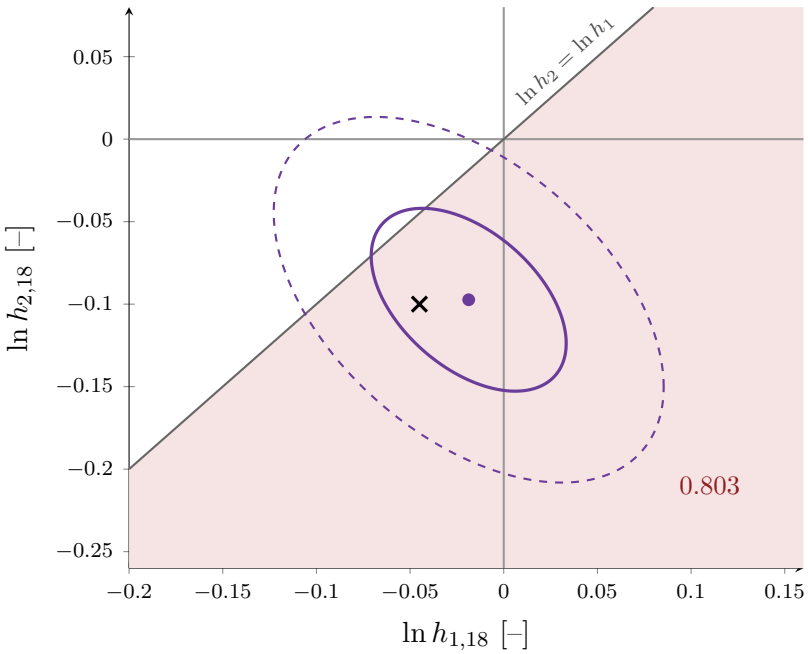


Figure 5.4: The joint posterior of the logarithms of the two health indices under the Laplace normal: means $(-0.01870,\ -0.09734)$, standard deviations 0.05204 and 0.05540, correlation -0.4771. Solid and dashed ellipses: the contours at one and two standard deviations, drawn as $\hat{\mathbf{m}} + n_\sigma\,\mathbf{L}(\cos\phi, \sin\phi)^\mathsf{T}$ with $\mathbf{L}$ the Cholesky factor of the 2×2 covariance and $n_\sigma = 1, 2$. Shaded: the event $\ln h_2 < \ln h_1$, probability 0.803. Gray lines: no change of either coil. Dot: the MAP pair. Cross: the stated truth, a check marker (Section 5.4.5).

General form. An estimation question about a derived quantity is summarized by its moments, equation (2.10),

$$\mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}] = \int f(\mathbf{z})\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z},$$

and by its credible interval, equation (2.11), at the levels 0.05 and 0.95.

Specialized form. The question is equations (2.10) and (2.11) with $f(\mathbf{z}) = \ln h_u^{\mathrm{raw}}$, the logarithm of the dashboard's index, and $p(\mathbf{z} \mid \mathbf{y})$ the Laplace normal. The approach is a derived quantity of each window's entries and known inputs. By the first two forward equations (5.3), $T_{d,u} - T_s(t) = (T_r - T_s(t))\,(1 - G_{u,t}/C_u)$, and with the hall balance,

$$A_{u,t} = T_{d,u} - T_s(t) = \frac{Q_t + Q_o(t)}{G_{1,t} + G_{2,t}} \left(1 - \frac{G_{u,t}}{C_u}\right), \qquad h_u^{\mathrm{raw}} = \frac{A_{u,0}(\mathbf{z})}{A_{u,18}(\mathbf{z})}. \tag{5.18}$$

The supply temperature cancels, because the approach is measured from it; the heat does not. On the logarithm, $\ln A_{u,t} = \ln(Q_t + Q_o(t)) - \ln(G_{1,t} + G_{2,t}) + \ln(1 - G_{u,t}/C_u)$, and with $\partial G_{u,t}/\partial g_{u,t} = G_{u,t}$ its slopes on the window's log-derates are

$$\frac{\partial \ln A_{u,t}}{\partial g_{u,t}} = -\frac{G_{u,t}}{G_{1,t} + G_{2,t}} - \frac{G_{u,t}/C_u}{1 - G_{u,t}/C_u}, \qquad \frac{\partial \ln A_{u,t}}{\partial g_{v,t}} = -\frac{G_{v,t}}{G_{1,t} + G_{2,t}}, \quad v \neq u, \tag{5.19}$$

the first term of the own slope through the return air, the second through the coil's own air side. The gradient of $\ln h_u^{\mathrm{raw}} = \ln A_{u,0} - \ln A_{u,18}$ on the joint state has these slopes in the columns of window 0, the same with the sign changed in the columns of window 18, and $1/(Q_t + Q_o(t))$ with the matching signs in the load columns. One choice enters here: the *unaged counterfactual* sets today's log-derates to the commissioned ones, $g_{u,18} = g_{u,0}$, at today's load and inputs. This

is what splits the dashboard's number into the part load and season make and the rest; it is used for nothing else, and the duty a coil carries under control is Chapter 6's. With equal conductances in both windows, the last two factors of equation (5.18) cancel in the ratio, and

$$h_u^{\mathrm{raw}}\big|_{g_{\cdot,18}=g_{\cdot,0}} = \frac{Q_0 + Q_o(0)}{Q_{18} + Q_o(18)}, \tag{5.20}$$

the ratio of the heat to remove, the same for both coils and independent of the supply and of the coils themselves.

Building blocks. The pair $(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ of Section 5.2.3; the approach (5.18) and its slopes (5.19); the gradient of $\ln h_u^{\mathrm{raw}}$, which for coil 1 at $\hat{\mathbf{z}}$ is

$$\mathbf{g}_{\ln h_1^{\mathrm{raw}}} = \big(0.00844,\ -2.66532,\ -0.49170,\ -0.00735,\ 2.56364,\ 0.47207,\ 0\big)^{\mathsf{T}},$$

the load entries in kW^{-1}; the unaged value (5.20).

The dashboard's index from the state. Table 5.6 collects the numbers. Each row is a quantity, for coil 1 and coil 2. At the MAP state the approaches are 2.7918 and 3.5281 K for coil 1 at windows 0 and 18, and 2.9885 and 4.2881 K for coil 2. Carried through its gradient, the dashboard's index has median 0.7913 and band [0.6524, 0.9597] for coil 1, and median 0.6969 and band [0.5869, 0.8276] for coil 2, beside the readings' 0.7943 and 0.7063; the full posterior gives 0.7914 and 0.6964. Coils that had not aged would show 0.8708 on the tile, with band [0.8276, 0.9163], for both coils, and the full posterior agrees, 0.8708. By equation (5.20) it is the heat ratio at the MAP state, 118.4739/136.0452: the heat to remove rose by 14.8 %. Load and season alone are a 12.9 % "loss" on the dashboard: 59.1 % of coil 1's $\ln h^{\mathrm{raw}}$ and 38.3 % of coil 2's.

Table 5.6: The dashboard's raw-performance index taken apart, from the joint state at its MAP value; bands are 5–95 %.

Quantity [unit]	Coil 1	Coil 2
Approach read, window 0 and window 18 [K]	2.86, 3.60	3.08, 4.36
Dashboard index from the readings [–]	0.7943	0.7063
Approach at the MAP state, window 0 and window 18 [K]	2.7918, 3.5281	2.9885, 4.2881
Dashboard index from the state, median [–]	0.7913	0.6969
its band [–]	0.6524 to 0.9597	0.5869 to 0.8276
Approach of unaged coils today [K]	3.2059	3.4317
Dashboard index of unaged coils, median [–]	0.8708	0.8708
its band [–]	0.8276 to 0.9163	0.8276 to 0.9163
Share of $\ln h_u^{\mathrm{raw}}$ from load and season [%]	59.1	38.3
The rest of $\ln h_u^{\mathrm{raw}}$ [–]	−0.09578	−0.22279
The coil's own $\ln h_{u,18}$; the other coil's [–]	−0.01870; −0.09734	−0.09734; −0.01870
Readings taken as exact through the physics [–]	0.9803	0.9070
Health index, median [–]	0.9815	0.9072

The remainder of the dashboard's index. By equation (5.19) at window 18, the approach's slope on its own coil is −2.5636 for coil 1, −0.5279 through the return air and −2.0357 through the coil's own air side, and −1.9698 for coil 2; on the other coil the slopes are −0.4721 and −0.5279. The gradient of coil 1's dashboard index shows what follows: its entries in the columns of $g_{1,0}$ and $g_{1,18}$, −2.66532 and 2.56364, are the index's own gradient (0, −1, 0, 0, 1, 0, 0)

magnified, and its entries in the columns of $g_{2,0}$ and $g_{2,18}$, -0.49170 and 0.47207, are coil 2's index gradient $(0,\ 0,\ -1,\ 0,\ 0,\ 1,\ 0)$ at about half its size. The dashboard magnifies a coil's own change 2.0 to 2.6 times and charges it with about half of the other coil's. Linearized with window 18's slopes, coil 1's rest splits into its own change magnified, $-(-2.5636)(-0.01870) = -0.0479$, and coil 2's change at 0.47 of its size, $-(-0.4721)(-0.09734) = -0.0460$: a sum of -0.0939 against the rest's -0.09578. The two parts are of about the same size, and the first is coil 1's own change of -0.01870 magnified, not that change itself.

The readings taken as exact. Taking today's and the commissioning readings as exact and solving the air side of equation (5.2) for the conductance, $G_{u,\text{test}} = C_u\,(y_{T_r} - y_{T_{d,u}})/(y_{T_r} - T_s)$, gives an index of 0.9803 for coil 1 and 0.9070 for coil 2, within 0.0012 and 0.0002 of the posterior medians. The physics, not the probability, takes the load and the season out: the readings-as-exact index already uses the known inputs. What it lacks is the band, and with it the answer to the second question.

Figure 5.5 sets the four readings of each coil on one axis of the health index. Each row is a value; the bars run from $h = 1$ to it, so their length is the loss the value reports. For each coil, from the top: the dashboard's index from the readings, the value coils that had not aged would show, the readings taken as exact, and the health index with its 5–95 % band as an error bar; the dotted line is the stated truth. Each value is printed beside its bar, the health index's at the left end of its band. The first bar of each coil is the longest by far; the second, which no aging produced, is longer than the fourth.

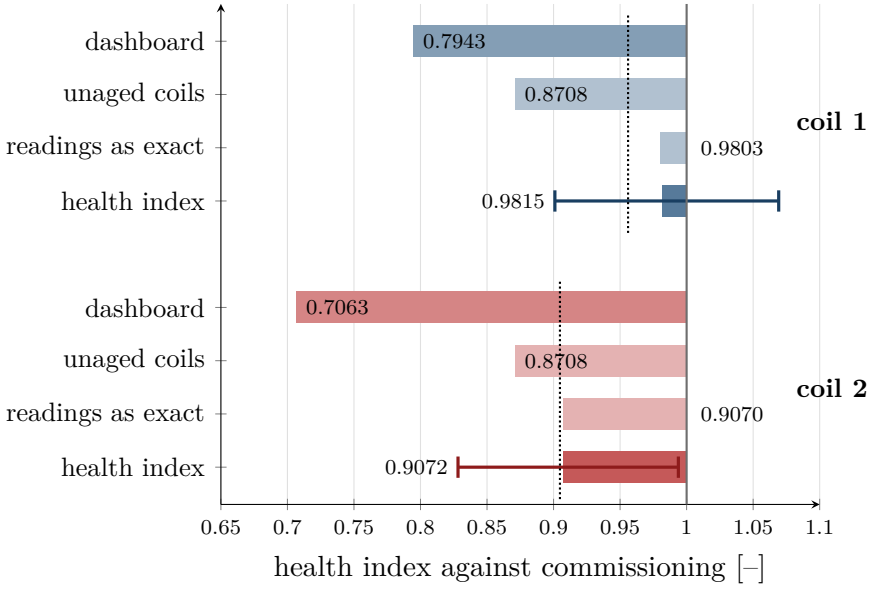


Figure 5.5: The dashboard's index against the health index, per coil. Bars run from $h = 1$ (gray line) to each value: the dashboard's raw-performance index from the readings; the value of coils that had not aged, equation (5.20); the readings taken as exact through the physics; and the health index, its median with the 5–95 % band as an error bar. Dotted: the stated truth of each coil's index, a check marker (Section 5.4.5).

5.3.4 The capacity scenario today

Asset manager's question. Do the coils still carry the hall's capacity scenario today? The record states the scenario: rack 3 at most 27°C at an IT load of 150 kW and a chilled-water supply of 15.5°C.

General form. A prediction under a threshold is the posterior probability that the requirement holds, equation (2.12),

$$\Pr(f(\mathbf{z}) \geq f^* \mid \mathbf{y}) = \int \mathbf{1}\big[f(\mathbf{z}) \geq f^*\big]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}.$$

Specialized form. The question is equation (2.12) with $f(\mathbf{z}) = G_{1,18} + G_{2,18} = \bar{G}_1 \mathrm{e}^{z_5} + \bar{G}_2 \mathrm{e}^{z_6}$, the hall's total conductance today, $f^* = G^*$, and $p(\mathbf{z} \mid \mathbf{y})$ the Laplace normal. Its gradient, by $\partial G_{u,t}/\partial g_{u,t} = G_{u,t}$, is

$$\mathbf{g}_G = (0,\ 0,\ 0,\ 0,\ G_{1,18},\ G_{2,18},\ 0)^{\mathsf{T}}, \qquad \Pr(G_{1,18} + G_{2,18} \geq G^* \mid \mathbf{y}) = \Phi\left(\frac{f(\hat{\mathbf{z}}) - G^*}{\sqrt{\mathbf{g}_G^{\mathsf{T}} \boldsymbol{\Sigma}\, \mathbf{g}_G}}\right). \quad (5.21)$$

One choice enters here: $G^* = 10.761\,\mathrm{kW/K}$, the total conductance the scenario needs, is the record's stated requirement, taken from the capacity study as a stated input, as a planning date is, and not a quantity of this posterior. This is what keeps the answer to today's margin on today's evidence, and it fails if the study behind G^* is wrong; giving the requirement a band of its own belongs to the crossing time of Chapter 14 and the sensor selection of Chapter 23.

Building blocks. The pair $(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$; today's conductances of Section 5.2.3, which are the entries of $\mathbf{g}_G$; the requirement G^*; Φ.

The margin today. With $G_{1,18} = 6.7059$ and $G_{2,18} = 5.9963\,\mathrm{kW/K}$ in $\mathbf{g}_G$, equation (5.21) gives $G_{1,18} + G_{2,18} = 12.7022\,\mathrm{kW/K}$ with standard deviation $0.2408\,\mathrm{kW/K}$, so the requirement lies $(12.7022 - 10.7610)/0.2408 = 8.06$ standard deviations below today's total, and

$$\Pr(G_{1,18} + G_{2,18} < G^* \mid \mathbf{y}) = \Phi(-8.06) = 3.8 \times 10^{-16}.$$

None of the 200 000 draws of the full posterior falls below G^*, and their median total is $12.7000\,\mathrm{kW/K}$. Against its rating of $14\,\mathrm{kW/K}$ the hall stands at 0.9073; the requirement is 0.7686. The margin is $1.9412\,\mathrm{kW/K}$, 15.3 % of today's total, and since commissioning the total has fallen by $0.7396\,\mathrm{kW/K}$, from $13.4418\,\mathrm{kW/K}$. The answer stops at the margin: how long it lasts needs the degradation rate of Chapter 10 and the crossing time of Chapter 14.

5.4 The checks (step 5)

Each answer of Section 5.3 rests on approximations and inputs that this hall lets be tested against the calculation they stand for.

5.4.1 The full posterior against the Laplace answers

The reweighted draws of Section 5.2.4 have been set beside every answer of step 4; their Monte Carlo standard errors, by batch means, are 0.0001 and 0.0002 on the two medians, at most 0.0004 on the band ends, 0.00121 and 0.00063 on the two probabilities of a loss, and 0.000847 on the probability of the third question. The full posterior's means of $\ln h$ are -0.01854 and -0.09766, against the MAP values -0.01870 and -0.09734, and its standard deviations 0.05225 and 0.05561, against 0.05204 and 0.05540. Across the three probabilities the largest gap from the Laplace answer is 0.0006, and across the medians and band ends of h it is 0.0008. The probability gaps lie within the Monte Carlo standard errors: 0.63976 and 0.96007 against 0.64033 and 0.96056 with standard errors of 0.00121 and 0.00063, and 0.802758 against 0.802793 with 0.000847. Coil 1's mean of $\ln h$ lies about one standard error (0.00014) from the MAP

value and coil 2's about two. What the draws resolve is a slightly wider posterior: the standard deviations of $\ln h$, 0.05225 and 0.05561 against 0.05204 and 0.05540, with standard errors of 0.00009 and 0.00011, and coil 1's upper band end, 1.0699 against 1.0692, with 0.0002. The Laplace normal is the posterior to the precision the answers are reported in, and the logarithm of the index is close to normal, as the choice of Section 5.3.1 required.

5.4.2 The limit of exact rows and the curvature term

The fit runs on the reduced state, which equation (4.48) makes the limit $\sigma_p \to 0$ of the full state. The check fits the chapter's own full state of seventeen entries, the seven of equation (5.5) and each window's $(T_r, T_{d,1}, T_{d,2}, T_c, T_1)$, with the ten physical rows (5.2) at a common finite tolerance beside the readings and priors, and lets the tolerance shrink. Table 5.7 gives the result: each row is a tolerance, the columns the logarithm of each index with its standard deviation, and the last row the reduced state. At 0.1 the spreads are wider by the discrepancy the rows allow; at 0.01 and 0.001 the full state reproduces the reduced state to the fifth decimal. Each fit took five iterations.

Table 5.7: The limit of exact physical rows on the seventeen-entry full state of the two windows: the logarithm of each health index and its standard deviation at three tolerances, against the reduced state.

Tolerance σ_p [kW or K]	$\ln h_{1,18}$ [–]	sd [–]	$\ln h_{2,18}$ [–]	sd [–]
0.1	−0.01858	0.05219	−0.09722	0.05555
0.01	−0.01870	0.05204	−0.09734	0.05540
0.001	−0.01870	0.05204	−0.09734	0.05540
reduced state	−0.01870	0.05204	−0.09734	0.05540

The covariance (5.10) is the Gauss–Newton form, without the curvature term $\mathbf{C} = \mathbf{H} - \mathbf{J}^{\mathsf{T}}\mathbf{J}$ of equation (4.44). The full Hessian of the objective at $\hat{\mathbf{z}}$, computed by differences, gives the Newton covariance, in which the standard deviations of $\ln h$ are 0.05200 and 0.05534, against Gauss–Newton's 0.05204 and 0.05540, ratios of 0.9993 and 0.9990. The largest entry of $\mathbf{C}$ is 1.09×10^{-3} of the largest entry of $\mathbf{J}^{\mathsf{T}}\mathbf{J}$: the residuals are small, so the term they multiply is too. The Gauss–Newton covariance is the one reported.

5.4.3 Sensitivity to the priors

Table 5.8 repeats the fit under other priors. Each row is one prior; the columns give each coil's index, the standard deviation of its logarithm, the probability of a loss, and the shift of $\ln h$ from the chapter's value in units of its posterior standard deviation. The readings set the index: every row moves it by 0.14 standard deviations or less, except a rating prior three times tighter, which moves coil 2's by −0.31 and raises its probability of a loss to 0.9900. The rating prior put on today's entries as well, the other choice of Section 5.1.4, pulls both of today's entries toward the rating, and the two coils move in opposite directions. Coil 2's entry, the farther from the rating, moves from −0.1548 to −0.1464, its index from 0.9072 to 0.9146 and its probability of a loss from 0.9606 to 0.9500: the prior hides part of its loss. Coil 1's entry moves the other way, from −0.0429 to −0.0445, its index falls from 0.9815 to 0.9796 and its probability of a loss rises from 0.6403 to 0.6556. The readings fix the sum of today's two conductances better than their split, the correlation −0.4607 of Section 5.2.3, so a higher $g_{2,18}$ at the same readings is paid for by a lower $g_{1,18}$. The survey's width on β_1 does not move either index.

Table 5.8: The health indices under other priors. The shift is the change of $\ln h_{u,18}$ from the chapter's fit in units of its posterior standard deviation.

Prior	$h_{1,18}$ [–]	sd $\ln h_1$ [–]	P_1 [–]	shift [sd]	$h_{2,18}$ [–]	sd $\ln h_2$ [–]	P_2 [–]	shift [sd]
rating sd 0.05	0.9797	0.04605	0.6721	−0.04	0.8920	0.04912	0.9900	−0.31
rating sd 0.15 (the chapter's)	0.9815	0.05204	0.6403	0.00	0.9072	0.05540	0.9606	0.00
rating sd 0.30	0.9813	0.05300	0.6394	0.00	0.9097	0.05647	0.9532	+0.05
no rating prior	0.9812	0.05335	0.6393	−0.01	0.9106	0.05687	0.9503	+0.07
rating prior centered at −0.10	0.9850	0.05209	0.6140	+0.07	0.9109	0.05546	0.9538	+0.07
rating prior on today's entries too	0.9796	0.05137	0.6556	−0.04	0.9146	0.05431	0.9500	+0.14
survey sd 0.04 on β_1	0.9815	0.05203	0.6403	0.00	0.9072	0.05539	0.9606	0.00
survey sd 0.16 on β_1	0.9815	0.05204	0.6403	0.00	0.9072	0.05540	0.9606	0.00

5.4.4 Other forms of the state

Section 5.2.2 chose one state over two windows; the check sets the two ways of carrying the commissioning fit forward beside it, and then the trend rows and the racks the state leaves out. Window 0 fitted alone, on the state $(Q_0, g_{1,0}, g_{2,0}, \beta_1)$, gives $g_{1,0} = -0.02494$ with standard deviation 0.04000, $g_{2,0} = -0.05814$ with 0.04111, a correlation of −0.4009 between them, and $\beta_1 = 0.19989$ with 0.04711. Carried into a fit of window 18 as Chapter 2's marginal rows on $g_{1,0}$ and $g_{2,0}$, with the survey on β_1 again, it gives $h_{1,18} = 0.9826$ and $h_{2,18} = 0.9082$, standard deviations of $\ln h$ of 0.05316 and 0.05647, and $\mathrm{sd}(\ln h_2 - \ln h_1) = 0.08491$, against the joint state's 0.05204, 0.05540 and 0.09234. Each index is slightly too wide and the comparison of the two too narrow, 91.9 % of the joint state's. Carried as one row on $(g_{1,0}, g_{2,0}, \beta_1)$ with the full covariance of window 0's fit, it gives 0.9815 and 0.9072, with 0.05203, 0.05539 and 0.09235: the joint state, to the fourth decimal. The marginal rows drop the commissioning correlation, which by equation (5.17) adds to the variance of the comparison, and they drop what window 0's rack-1 reading said about β_1, which the fit of window 18 then knows from the survey and its own reading alone. The trend rows of Chapter 4 with a free rate per coil, at $\sigma_q = 0.005$ and as one marginal row over the 18 windows, move the seven entries' MAP state and covariance block by less than 10^{-6}, as Section 5.1.4 said. The nine-entry state that also carries β_2 and β_3 with their survey priors agrees with the seven-entry state to below 10^{-6} in every MAP entry and every entry of $\mathbf{\Sigma}$; both fractions keep the survey's spread 0.0800 and are correlated with no other entry, to below 10^{-6}. Their block is the prior's, as Section 5.1.2 stated.

5.4.5 The answers against the stated truth

The last check sets the answers against the stated truth of the record (Section 5.6.6). Table 5.9 collects the routes: each row is a route or a variation, the columns each coil's index, its band, the standard deviation of its logarithm and its probability of a loss, and the third question's probability or the standard deviation of the difference. Coil 1's true log-derate goes from −0.0200 to −0.0650, a true index of 0.9560; the posterior median lies +0.51 standard deviations from it, inside the band. Coil 2's goes from −0.0400 to −0.1400, a true index of 0.9048, and the median lies +0.05 standard deviations from it, also inside the band; that closeness is the luck of the record's draws, not the precision of the fit, whose band spans 0.8282 to 0.9938. The true difference of the logarithms is −0.0550, and the posterior's lies −0.26 standard deviations from it. The true total conductance today is 12.6450 kW/K, +0.24 standard deviations from the posterior's. The true loads are 110.00 and 128.96 kW, against the MAP's 110.47 and 130.17 kW, and the true β_1 is 0.22, against 0.2040. The dashboard's index at the true state is 0.7430 for coil 1 and 0.6813 for coil 2, and that of unaged coils at the true loads 0.8751: the dashboard's distortion is the plant's, not the sensors'.

Table 5.9: The routes to the health indices of window 18 and the stated truth. Bands are 5–95 %; P_u is the probability of a loss; the last column is the probability of the third question or, where marked, sd($\ln h_2 - \ln h_1$).

Route	$h_{1,18}$ [–]	band [–]	sd $\ln h_1$ [–]	P_1 [–]	$h_{2,18}$ [–]	band [–]	sd $\ln h_2$ [–]	P_2 [–]	third question [–]
Laplace, joint state (reported)	0.9815	0.9010–1.0692	0.05204	0.6403	0.9072	0.8282–0.9938	0.05540	0.9606	0.8028
Full posterior, reweighted draws	0.9815	0.9008–1.0699	0.05225	0.6398	0.9070	0.8279–0.9939	0.05561	0.9601	0.8028
Newton covariance	–	–	0.05200	–	–	–	0.05534	–	–
Full state, $\sigma_p = 0.001$	–	–	0.05204	–	–	–	0.05540	–	–
Commissioning as marginal rows	0.9826	–	0.05316	–	0.9082	–	0.05647	–	sd 0.08491
Commissioning as one full-covariance row	0.9815	–	0.05203	–	0.9072	–	0.05539	–	sd 0.09235
Joint state, for comparison	0.9815	–	0.05204	–	0.9072	–	0.05540	–	sd 0.09234
Stated truth	0.9560	–	–	–	0.9048	–	–	–	–

5.5 The report (step 6)

The dashboard's "lost 21 %" and "lost 29 %" are not condition. Coils that had not aged at all would read 0.87 on that tile, because the heat to remove rose 14.8 % since handover, and the approach magnifies a coil's own change 2.0 to 2.6 times and charges it with about half of the other's. Read as a parameter against the commissioning fit, coil 1 is at 0.98 (0.90 to 1.07) and no loss is demonstrated, $P = 0.64$; coil 2 is at 0.91 (0.83 to 0.99), with probability 0.96 that it has lost conductance; coil 2 has fallen more than coil 1 with probability 0.80; and the hall carries its capacity scenario with 8 standard deviations to spare.

5.5.1 The answers and their bands

Table 5.10 gives the answers in the order of Table 5.2. Each row is a question; the second column gives the reported value from the Laplace normal, the third the full posterior's value beside it, and the fourth the band: the 5–95 % band of an index or a ratio, and for a probability its range over the priors of Table 5.8.

The answers rest on the joint state's posterior, and the commissioning test carries more than half of each index's variance. The readings taken as exact give nearly the same medians, 0.9803 and 0.9070: the physics removes the load and the season, and the posterior adds the band, which is what separates coil 1's 0.64 from coil 2's 0.96.

5.5.2 The entry in the maintenance log

The condition assessment of window 18 is the first condition entry of the two-unit hall's maintenance log, after its window-0 commissioning test. It uses nothing later than window 18, and in the values of Table 5.10 it reads:

> *Window 18 (Jun Y2), condition assessment against the commissioning baseline (window 0).*
> Coil 1: health index 0.98 (5–95 % band 0.90 to 1.07); no loss demonstrated, $P(\text{loss}) = 0.64$.
> Coil 2: health index 0.91 (0.83 to 0.99); $P(\text{loss}) = 0.96$; fallen more than coil 1 with $P = 0.80$.
> Capacity scenario ($G^* = 10.76\,\text{kW/K}$): held, $G_1 + G_2 = 12.70 \pm 0.24\,\text{kW/K}$.
> Dashboard approach indices (0.79, 0.71): not a condition reading; load and season alone

Table 5.10: The report for the two-unit hall at window 18 (Jun Y2) against its commissioning test. Reported values from the Laplace normal of the joint state; bands are 5–95 %, or for a probability the range over the priors of Table 5.8.

Asset manager's question	Reported value	Full posterior	Band
How far has this coil's conductance fallen from its commissioned value?	coil 1: $h = 0.9815$ [–]; coil 2: $h = 0.9072$ [–]	0.9815; 0.9070 [–]	coil 1: 0.9010 to 1.0692; coil 2: 0.8282 to 0.9938 [–]
Has it lost any conductance at all, beyond what the readings' scatter allows?	coil 1: 0.640; coil 2: 0.961 [–]	0.640; 0.960 [–]	coil 1: 0.614 to 0.672; coil 2: 0.950 to 0.990 [–]
Which coil has fallen more, and how sure is that?	coil 2, probability 0.803 [–]; ratio $h_2/h_1 = 0.9244$ [–]	0.803 [–]	ratio 0.7941 to 1.0760 [–]
Is the dashboard's 21 % and 29 % the coils, or the load and the season?	dashboard index 0.7913 and 0.6969 [–]; unaged coils 0.8708 [–]	0.7914 and 0.6964; 0.8708 [–]	0.6524 to 0.9597 and 0.5869 to 0.8276; unaged 0.8276 to 0.9163 [–]
Do the coils still carry the hall's capacity scenario today?	yes: $G_1 + G_2 = 12.7022$ [kW/K], sd 0.2408 [kW/K]; probability short 3.8×10^{-16} [–]	median 12.7000 [kW/K]; no draw short	margin 1.9412 [kW/K], 8.06 sd above $G^* = 10.761$ [kW/K]

account for 0.87.
Action: the cleaning work orders raised from the dashboard are not supported for coil 1. Coil 2 is kept for the trend of later windows; a cleaning decision needs a decision level, not set here.
Next measurement: a reference discharge-air reading on unit 2 in this window.

The entry records what the band rests on as well as the values: the commissioning test and window 18, each fixing its window's log-derates to between 0.03476 and 0.04074, the earlier of them carrying more than half of every later index's variance. Which coil to clean, and when, is not settled by a probability of a loss; it needs the decision level that Chapter 19 calibrates against the costs of false and missed work orders.

Common mistake

Reading a raw-performance index as a health index. A ratio of approaches mixes the load, the season and both coils; the coil's condition is its conductance, read against the commissioning fit with its band.

5.6 The next measurement (step 7)

The report leaves coil 1's index consistent with no change and coil 2's loss anywhere from 0.62 to 17.18 %. A further measurement in today's window is one more row of the residual vector, and what it removes from each band is known before its value is.

5.6.1 The preposterior of a candidate row

A candidate measurement a is one row or more in today's window, with Jacobian lines $\mathbf{J}_a$ in the columns of the joint state, each the gradient of the quantity it reads over its standard deviation. Its preposterior covariance is equation (4.65),

$$\boldsymbol{\Sigma}_a = \left(\boldsymbol{\Sigma}^{-1} + \mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a\right)^{-1}, \qquad \mathrm{sd}_a(\ln h_{u,18}) = \sqrt{\mathbf{g}_{\ln h_u}^{\mathsf{T}}\,\boldsymbol{\Sigma}_a\,\mathbf{g}_{\ln h_u}}, \tag{5.22}$$

and the width it takes off each index is $1 - \mathrm{sd}_a/\mathrm{sd}$ of $\ln h_{u,18}$. The candidates open to the asset manager are of four kinds. A water-side heat meter on coil u, of standard deviation

$\sigma_W = 3.0\,\mathrm{kW}$, reads the duty $D_{u,18} = G_{u,18}(T_r - T_s)$; it is the permanent instrument at the place of the question. A reference discharge sensor of standard deviation 0.1 K reads $T_{d,u}$ beside the unit's own sensor. A second logged hour of today's five readings is a new load entry with five fresh rows on the same coils; it widens the state by one entry, and its preposterior is formed on the eight-entry state. And a second return-air sensor or a second load meter repeats a row already present. The last candidate is not a measurement: today's log-derates known exactly, the limit of a perfect reading of each coil today.

5.6.2 The candidates

Table 5.11 ranks the candidates. Each row is a measurement in today's window; the columns give, for each coil, the standard deviation of $\ln h_{u,18}$ after it by equation (5.22) and the share of the present 0.05204 and 0.05540 that it removes. The widths removed are compared as bars in Figure 5.6, two per candidate, ranked by coil 2's, with the limit of today's log-derates known exactly as dashed lines.

Table 5.11: Candidate measurements in today's window, from the posterior of the joint state; the width removed is taken from the present $\mathrm{sd}(\ln h_{1,18}) = 0.05204$ and $\mathrm{sd}(\ln h_{2,18}) = 0.05540$.

Candidate (today's window)	$\mathrm{sd}_a(\ln h_1)$ [–]	coil 1 [%]	$\mathrm{sd}_a(\ln h_2)$ [–]	coil 2 [%]
Water-side heat meter on coil 1, $\sigma_W = 3.0\,\mathrm{kW}$	0.04821	7.3	0.05375	3.0
Water-side heat meter on coil 2, $\sigma_W = 3.0\,\mathrm{kW}$	0.05055	2.8	0.05072	8.4
Water-side heat meters on both coils	0.04733	9.0	0.04976	10.2
Reference discharge sensor on unit 1, 0.1 K	0.04214	19.0	0.05326	3.9
Reference discharge sensor on unit 2, 0.1 K	0.05039	3.2	0.04416	20.3
A second logged hour of today's five readings	0.04605	11.5	0.04844	12.6
A second return-air sensor, 0.3 K	0.05194	0.2	0.05534	0.1
A second load meter, 3.0 kW	0.05157	0.9	0.05474	1.2
Today's log-derates known exactly (the floor)	0.03957	23.9	0.04070	26.5

The ranking follows each candidate row's slope on today's entries, its line in the columns of $g_{1,18}$ and $g_{2,18}$, while the budget of Section 5.3.1 sets the cap. A reference discharge sensor acts on its own coil's column, where the discharge rows of equation (5.12) are largest, and removes 19.0 % of coil 1's width or 20.3 % of coil 2's; a second logged hour takes 11.5 and 12.6 % off both; the heat meters take less; and a second return-air sensor or load meter removes almost nothing, because the rows already present fix what they read.

5.6.3 The commissioning floor and the heat meter

A measurement today barely narrows the commissioning entries, whose standard deviations, 0.03962 and 0.04074, are 76.1 % and 73.5 % of $\mathrm{sd}(\ln h)$: even today's log-derates known exactly, through the small covariance the shared β_1 leaves, move them only to 0.03957 and 0.04070, which are then the whole of $\mathrm{sd}(\ln h)$. That limit takes off 23.9 % and 26.5 %, and every candidate in the table is capped there. The commissioning test sets the precision of every later health index: it is the floor of the gold bars of Figure 5.3, and after handover it can only be carried, not improved.

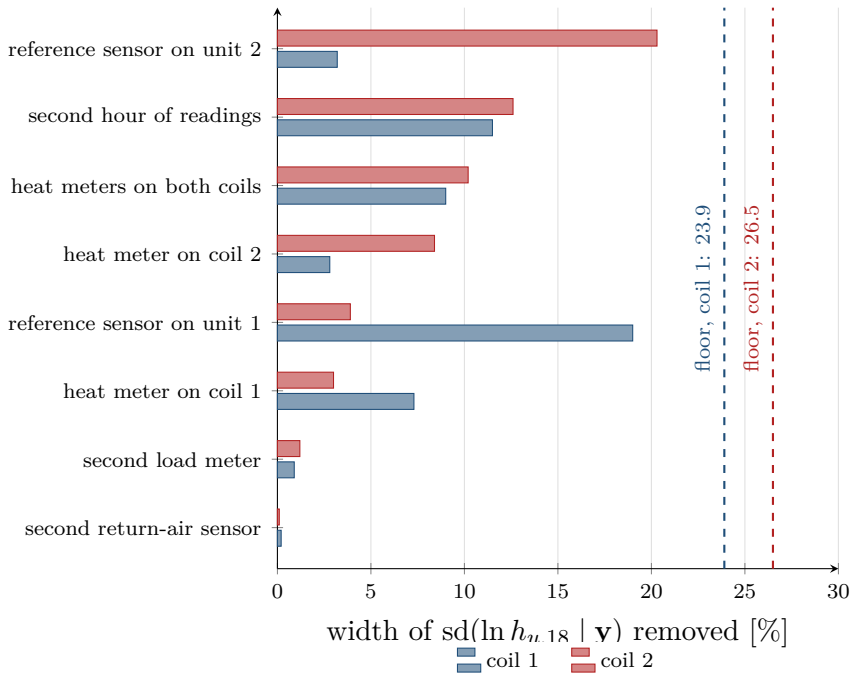


Figure 5.6: The candidate measurements of Table 5.11, ranked by the width they remove from coil 2's index, with coil 1's beside each, by the preposterior (5.22). Dashed: the width removed if today's log-derates were known exactly, the most any measurement in today's window can remove.

The heat meter is read by its row. The coil-1 meter reads $D_{1,18} = G_{1,18}(T_r - T_s)$, and its gradient on the joint state has the entries 0.5279 on Q_{18}, +33.9052 on $g_{1,18}$ and −33.9052 on $g_{2,18}$, in kW per unit entry, and zeros elsewhere; its Jacobian line is that gradient over $\sigma_W = 3.0\,\text{kW}$. The equal and opposite entries say that the meter sees the split of the heat between the coils, $G_1/(G_1 + G_2)$ of it to coil 1, and not the hall's total, and a 3 kW meter on coil 1's 71.82 kW reads that split weakly, with a line of 33.9052/3.0 in the column of $g_{1,18}$ against the discharge reading's −30.1496 on 0.3 K. With meters on both coils the standard deviation of $g_{1,18}$ falls from 0.03476 to 0.02719, and of $g_{2,18}$ from 0.03847 to 0.02976, but the index barely moves, from 0.05204 to 0.04733, because the commissioning entry is untouched. A permanent meter's value lies in the later windows it will read, which is Chapter 23's question.

5.6.4 The recommendation

For coil 2, the coil with the probable loss, the next measurement is a reference discharge-air reading on unit 2 in this window, which takes 20.3 % off its band; for coil 1 the matching reading on unit 1 takes 19.0 % off. A second logged hour of today's readings costs nothing and takes 11.5 to 12.6 % off both, provided its errors are independent of the first hour's; readings whose errors are correlated from one hour to the next are Chapter 24's question. A second return-air sensor or a second load meter is not worth requesting for this question. None of these reaches the floor, and none touches the part of the band the commissioning test set.

5.6.5 What the analysis showed

The dashboard said the coils had lost 20.6 and 29.4 % of their condition, at 0.79 and 0.71; read as a parameter against the commissioning fit, they stand at 0.98 and 0.91. The difference is not the probability. With the load an entry of each window and the season its known inputs, the physics takes both out of the parameter, and even the readings taken as exact through the

same physics land within 0.0012 and 0.0002 of the posterior medians. The dashboard's ratio cannot do this, because an approach is a figure of the whole window, of the heat to remove and of both coils, and it magnifies the coil's own change while borrowing half of the other's. What the probability adds is the band, and the band says that one of the two "losses" is within the scatter of the two windows: coil 1's index is consistent with no change, and coil 2's is a loss more likely than not by a wide margin, but of uncertain size. The commissioning test is a permanent part of every later index, carrying more than half of its variance, and no instrument added after handover removes that part: it sets the floor of every condition assessment the asset will have.

5.6.6 Limits of the model

A constructed record. The condition record is generated from a stated truth, with the stated sensor errors, and step 5 compares the answers with that truth (Section 5.4.5). It is not field data, and the agreement is a check of the method on its own model, not of the model against a real hall.

Steady state, pinned airflow, no control. The rows (5.2) hold within a window only once the hall has settled, and the discharge readings inform the coils only while nothing holds the discharge at a set point. Under an active discharge controller the coil's loss shows in its valve and its duty, not in its discharge temperature, and the duty is Chapter 6's question.

Unbiased sensors. A drift in a discharge sensor would move its coil's log-derate, and the rows of this chapter would carry it into the index as aging. Telling the two apart is the comparison of Chapter 7.

The shared fraction and the known inputs. The fraction β_1 is taken as unchanged since handover, and the chilled-water supply and the other heat as known. A rearranged rack, a supply sensor out of calibration or a wrong schedule of the other heat would each put its error into the log-derates.

The requirement as stated. The capacity requirement G^* is taken from the record, not from this posterior, and the margin of Section 5.3.4 is only as good as the capacity study behind it.

Two windows. Two windows say nothing about how fast the loss accrued or whether it was steady; the degradation rate is Chapter 10's question, its acceleration Chapter 11's, and a change point Chapter 12's.

The next chapter. Chapter 6 asks a second question of the same fit: whether a coil that still meets its set point is working harder to do so.

5.7 What to retain

Definition 5.1 (Health index from a physical parameter)**.** The health index of a parameter against commissioning is its change since window 0, $h_{u,t} = \mathrm{e}^{g_{u,t}-g_{u,0}}$, read from one posterior of a state that carries both windows: $\widehat{\ln h} = \hat{g}_{u,t} - \hat{g}_{u,0}$ and $\mathrm{Var}(\ln h \mid \mathbf{y}) = \Sigma_{aa} + \Sigma_{bb} - 2\Sigma_{ab}$, with a and b the positions of $g_{u,0}$ and $g_{u,t}$ in the state (for coil 1 here, $\Sigma_{22} + \Sigma_{55} - 2\Sigma_{25}$), the median $\mathrm{e}^{\widehat{\ln h}}$ and the band the image of the normal quantiles of $\ln h$. The known inputs of each window enter its physical rows and leave the parameter; a raw-performance index, a ratio of

performance figures such as approaches, is a derived quantity of every entry and input of both windows, and it is not a health index. The commissioning entry's variance is a floor that no later measurement removes.

Three conclusions carry beyond the hall. A health index is a parameter of the physics read in two windows, and the physics, given the inputs of each window, is what makes the two windows comparable. The commissioning test is an estimate with an error, and that error belongs to every later index; carrying it in the same state as today's window keeps the covariances that a carried marginal row drops. And a dashboard's raw-performance index can be read as a quantity of the same state, which says how much of its number is load, season and the other coil.

Table 5.12 lists the terms of the chapter with their expressions, units and the subsection that defines each.

Table 5.12: The terms of the health index and where each is defined. Units for the two-unit hall.

Term	Expression	Units	Meaning	Defined in
Health index against commissioning	$h_{u,t} = \mathrm{e}^{g_{u,t} - g_{u,0}}$, eq. (5.1)	–	the parameter's change since window 0	5.1.1
Derate factor	$d_{u,t} = \mathrm{e}^{g_{u,t}}$	–	the index against rating	5.1.1
Known input	$T_s(t)$, $Q_o(t)$ in eq. (5.2)	°C, kW	pinned per window; takes the season out	5.1.2
Joint state of two windows	eq. (5.5)	kW, –	both windows in one posterior	5.1.2
Rows and posterior				
Sparsity pattern	eq. (5.6)	–	the windows meet only through β_1	5.1.4
Residual vector	eq. (5.8), thirteen rows	–	readings of both windows and the priors	5.2.1
Block structure	eq. (5.11)	mixed	two window blocks and β_1	5.2.2
Reweighted draws	eq. (5.13), $N_{\mathrm{eff}} = 1/\sum_j \varpi_j^2$	–	the full posterior beside the Laplace normal	5.2.4
Questions and answers				
Variance of the index	eq. (5.14)	–	a difference's variance	5.3.1
Probability of a loss	eq. (5.16)	–	any loss beyond the scatter	5.3.2
Approach temperature	$A_{u,t}$, eq. (5.18)	K	a derived quantity of the window	5.3.3
Raw-performance index	$h_u^{\mathrm{raw}} = A_{u,0}/A_{u,18}$	–	the dashboard's tile	5.3.3
Unaged counterfactual	eq. (5.20)	–	the heat ratio; load and season alone	5.3.3
Commissioning floor	$\mathrm{sd}(g_{u,0} \mid \mathbf{y})$ against $\mathrm{sd}(\ln h)$	–	what no later measurement removes	5.6.3
Preposterior of a candidate	eq. (5.22)	–	the band after a measurement	5.6.1

5.8 Exercises

Exercise 5.1. *Formulation.* Write the question "has coil 2 lost more than 5 %?" in the form of Table 5.2 and evaluate it from the fit of this chapter; do the same for coil 1. (A prediction under a threshold, equation (2.12) with $f = -\ln h_{2,18}$ and $f^* = -\ln 0.95$: $\Pr(\ln h_{2,18} < \ln 0.95 \mid \mathbf{y}) = \Phi\big((\ln 0.95 + 0.09734)/0.05540\big) = 0.7971$; for coil 1, $\Phi\big((\ln 0.95 + 0.01870)/0.05204\big) = 0.2655$.)

Exercise 5.2. *Derivation.* Show that the dashboard's index of coils that have not aged is

the ratio of the heat to remove, independent of the chilled-water supply and of the coils, and evaluate it. (By equation (5.18), $A_{u,t} = (Q_t + Q_o(t))\,(1 - G_{u,t}/C_u)/(G_{1,t} + G_{2,t})$, which holds no T_s; with $G_{u,18} = G_{u,0}$ the last two factors are equal in both windows and cancel in $A_{u,0}/A_{u,18}$, leaving $(Q_0 + Q_o(0))/(Q_{18} + Q_o(18)) = 118.4739/136.0452 = 0.8708$.)

Exercise 5.3. *Derivation.* Derive $\partial \ln A_u/\partial g_u$ and $\partial \ln A_u/\partial g_v$ from the forward equations (5.3), and evaluate them for coil 1 at window 18. ($\ln A_u = \ln(Q+Q_o) - \ln(G_1+G_2) + \ln(1-G_u/C_u)$ and $\partial G_u/\partial g_u = G_u$, so $\partial \ln A_u/\partial g_u = -G_u/(G_1+G_2) - (G_u/C_u)/(1-G_u/C_u)$ and $\partial \ln A_u/\partial g_v = -G_v/(G_1+G_2)$; for coil 1 at window 18, $-0.5279 - 2.0357 = -2.5636$ and -0.4721.)

Exercise 5.4. *Application.* Carry the commissioning fit forward as Chapter 2's marginal rows on $g_{1,0}$ and $g_{2,0}$, and recompute $\mathrm{sd}(\ln h_2 - \ln h_1)$. Say why it is narrower than the joint state's. (The difference D_h takes $g_{1,0}$ and $g_{2,0}$ with opposite signs, so their covariance enters $\mathrm{Var}(D_h)$ as $-2\,\mathrm{Cov}(g_{1,0}, g_{2,0})$, a positive term for the negative correlation -0.40 of window 0's fit; the marginal rows make the two entries independent and drop that term: 0.08491 against 0.09234.)

Exercise 5.5. *Application.* A reference discharge sensor on unit 2 reads 17.20°C (standard deviation 0.1 K) in today's window, against the fitted 17.2275°C, whose posterior standard deviation is 0.2700 K. Update $h_{2,18}$ by the correction form and by a refit. (The row's gradient is that of $T_{d,2}$ today, $(0,\ 0,\ 0,\ 0.0315,\ -2.2638,\ -8.4465,\ 0)$; by the rank-one update (4.66) with that gradient and the innovation $17.20 - 17.2275$, the gain on $\ln h_2$ is -0.11619 per K, $\ln h_2$ moves from -0.09734 to -0.09415 and its standard deviation from 0.05540 to 0.04416: $h_{2,18}$ goes from 0.9072 to 0.9101, band $[0.8464,\ 0.9787]$, and $P(\text{loss}) = 0.9835$. The refit with the row gives $\ln h_2 = -0.09415$, standard deviation 0.04413, and $h_{2,18} = 0.9101$.)

Chapter 6

Hidden Degradation Under Control

6.1 The held set point and the effort behind it

Window 18 is June of year 2, eighteen months after the two-unit hall was handed over. The condition assessment of Chapter 5 put coil 2 at 0.9072 of its commissioned conductance and kept it for the trend. Today's dashboard shows the hall at work. The set-point tile reads the cold aisle, which the mixing of the two discharge streams makes the mean of the two discharge readings, the units' capacity rates being equal: 16.92°C against 16.97°C at handover, −0.05 K. Unit 1's discharge reads 16.54°C (16.86°C at handover), unit 2's 17.30°C (17.08°C) and rack 1's inlet 18.26°C (18.08°C): the set point is met. The supply tile shows the chilled water at 12.94°C, 1.06 K colder than at handover, which the asset manager reads as the chiller's winter reset. The hall's heat by the meter, the IT load and the other heat together, has grown from 118.92 to 136.87 kW, by 15.1 %, and the even split puts each unit at 68.43 kW against 59.46 kW, 8.97 kW more, in step with the load. The return air less the supply has widened from 8.71 to 10.54 K, by 1.83 K, which the asset manager puts down to the load. The conclusion the dashboard invites is that the units hold their set point, that they carry more only because the IT load grew, and that coil 2's loss costs nothing yet.

The tiles cannot say how much of the colder water is spent on the coils rather than on the load, how much harder each unit works than the commissioned coils would at today's load and today's water, which unit carries the heat that a fouled coil no longer passes, or how well two windows eighteen months apart know any of this. The question of this chapter is

how the effort a unit spends to hold its set point is read against the commissioned model at today's conditions, with its band, and where that effort goes.

It is answered from the fit of Chapter 5, unchanged, read at a scenario node: a copy of today's physical rows that carries the commissioned coils.

Jargon: hidden degradation, set point and controlled variable, manipulated variable, commissioned model at today's conditions, excess approach, supply penalty, duty, duty shift

Hidden degradation is a loss of condition that a reading held at its set point does not show, because something else gives way to hold it. A *set point* is the value a reading is to be kept at; the reading so held is the *controlled variable*, here the cold aisle. The *manipulated variable* is the input that moves to hold it, here the chilled-water supply. The *commissioned model at today's conditions* is the plant's rows at today's load and today's inputs with the coils' commissioned conductances; in the graph it is a *scenario node*, a copy of today's rows with no reading of its own. The *excess approach* of a unit is its approach today less the commissioned model's. The *supply penalty* is how much colder the chilled water runs, at today's heat, than the commissioned coils would need to hold the same cold aisle. The *duty* of a unit is the heat its coil passes to the chilled water, and the *duty shift* is its duty today less the commissioned model's at the same heat.

Key idea

A unit can hold its set point and still work harder. The extra effort shows in what the hall spends to hold the set point, the chilled-water supply and each unit's duty, and not in the reading at the set point. That effort is the excess of today's state over a scenario node: today's rows, carrying the commissioned coils, at today's load and inputs. It is read from the one posterior that carries both windows, so its band holds the commissioning test's error as well as today's.

6.1.1 Effort at a held set point

Chapter 5 read the coils' parameters, each against its commissioned value. A coil whose conductance has fallen still delivers the air it is asked for, as long as something else gives way: a colder supply, a wider approach, or the other unit's duty. A reading held at its set point is the sum of everything that moves it, the season, the load and the coils, and it says nothing of any one of them. The effort shows elsewhere, in what the hall spends to hold the reading and in how the heat divides between the units.

To read it, the plant is asked what the commissioned coils would do with today's load and today's water. That answer is a function of the state like any other: the commissioning test's coil entries, set into today's rows. Today's performance less that answer is the effort the coils' loss costs, and both terms are functions of one posterior that carries the commissioning test and today, so the difference carries the commissioning test's error as well as today's, and the covariance between the two windows.

> **Principle 6.1.** *Hidden degradation is the excess of a performance quantity over the commissioned model's value at the same load and inputs. It is a derived quantity of a posterior that carries both windows, and its variance takes each window's block of* $\boldsymbol{\Sigma}$ *and the cross term between them.*

6.1.2 The hall and the scenario node (step 0)

The plant is the two-unit hall of Section 5.1.2, unchanged: the physical rows (5.2) in each window, their forward equations (5.3), the reduced state of Section 4.2.8 and the joint state (5.5) of the commissioning test and window 18. A second subscript names the window where it is needed: $T_{r,18}$ is the return air of window 18, $T_{c,0}$ the cold aisle of window 0. Two things are added to the graph, a flow and a node.

The duty. The duty of unit u in window t is the heat its coil passes to the chilled water, the flow along the coil edge of the graph,

$$D_{u,t} = G_{u,t}\,(T_{r,t} - T_s(t)) = C_u\,(T_{r,t} - T_{d,u,t}), \qquad D_{1,t} + D_{2,t} = Q_t + Q_o(t), \tag{6.1}$$

the second form by the air side of unit u in equation (5.2) and the sum by its hall balance. On the reduced state the duty is a derived quantity: by the first forward equation, $D_{u,t} = (Q_t + Q_o(t))\,G_{u,t}/(G_{1,t} + G_{2,t})$, the heat to remove at the coil's share of the total conductance.

The scenario node. The question is asked at a scenario node: a copy of window 18's physical rows at today's load entry Q_{18} and today's known inputs $T_s(18)$ and $Q_o(18)$, carrying the commissioned log-derates $g_{1,0}$ and $g_{2,0}$ in place of $g_{1,18}$ and $g_{2,18}$. It has no reading and no prior: it is the form in which *Reliability Engineering by Derivation*, Chapter 20, carries a condition that was never observed. Its quantities are marked with a check accent, $\check{X}$ for the value of X at the scenario node ($\check{T}_r$, $\check{D}_u$, $\check{A}_u$), a mark that cannot be read as a degree sign. Written with the duty as the flow of each coil edge, its six rows are

$$\begin{aligned}
\check{r}_{\mathrm{h}} &= \big[\check{D}_1 + \check{D}_2 - Q_{18} - Q_o(18)\big]/\sigma_p && \text{hall balance,}\\
\check{r}_{\mathrm{w},u} &= \big[\check{D}_u - G_{u,0}\,(\check{T}_r - T_s(18))\big]/\sigma_p && \text{coil of unit } u = 1, 2,\\
\check{r}_{\mathrm{a},u} &= \big[C_u\,(\check{T}_r - \check{T}_{d,u}) - \check{D}_u\big]/\sigma_p && \text{air side of unit } u = 1, 2,\\
\check{r}_{\mathrm{m}} &= \big[(C_1 + C_2)\,\check{T}_c - C_1\check{T}_{d,1} - C_2\check{T}_{d,2}\big]/\sigma_p && \text{cold-aisle mixing,}
\end{aligned} \tag{6.2}$$

in kW, with $G_{u,0} = \bar{G}_u\,\mathrm{e}^{g_{u,0}}$. The hall balance with the two coil rows is the hall balance of equation (5.2), and each coil row with its air side is that unit's air side there, so the node's rows are window 18's rows of Chapter 5 with the duty made explicit and the commissioned coils in place of today's. The node has no rack-1 row, because no question reads its rack. Solved in the order of equation (5.3), the rows give

$$\begin{gathered}
\check{T}_r = T_s(18) + \frac{Q_{18} + Q_o(18)}{G_{1,0} + G_{2,0}}, \qquad \check{D}_u = G_{u,0}\,(\check{T}_r - T_s(18)),\\
\check{A}_u = \check{T}_{d,u} - T_s(18) = \check{T}_r - T_s(18) - \frac{\check{D}_u}{C_u}, \qquad \check{T}_c = \check{T}_r - \frac{Q_{18} + Q_o(18)}{C_1 + C_2},
\end{gathered} \tag{6.3}$$

the last by the air balance (5.4), which holds at the node as in every window. On the reduced state the node is Chapter 5's unaged counterfactual, window 18 with $g_{u,18} = g_{u,0}$, which split the dashboard's number in equation (5.20); here it is the question's own object.

The excesses. Today's state set against the node gives the three quantities the chapter asks for,

$$\begin{aligned}
\eta_{A,u} &= A_{u,18} - \check{A}_u && \text{excess approach of unit } u \text{ [K],}\\
\eta_{D,u} &= D_{u,18} - \check{D}_u && \text{duty shift of unit } u \text{ [kW],}\\
\eta_s &= T_{r,18} - \check{T}_r && \text{supply penalty [K],}\\
&= (Q_{18} + Q_o(18))\Big[\frac{1}{G_{1,18} + G_{2,18}} - \frac{1}{G_{1,0} + G_{2,0}}\Big],
\end{aligned} \tag{6.4}$$

the last by the first forward equation in window 18 and at the node. Each is a function of $(g_{1,0}, g_{2,0}, Q_{18}, g_{1,18}, g_{2,18})$: the commissioning test's coils and today's window.

Two identities. The rows tie the excesses together at every point of the state, not by substitution. The node and window 18 carry the same heat, $Q_{18} + Q_o(18)$, so by the hall balance (6.1) the duties sum to the same total in both; and by the air side of each unit, $A_u = T_r - T_s - D_u/C_u$ in both, at the same supply. Hence

$$\eta_{D,1} + \eta_{D,2} = 0, \qquad \eta_{A,u} = \eta_s - \frac{\eta_{D,u}}{C_u}, \quad u = 1, 2. \tag{6.5}$$

A duty that one unit sheds, the other carries. Each unit's excess approach is the hall's supply penalty, which reaches both units through the return air they share, less its own duty shift over its capacity rate. A third relation follows from the air balance (5.4): $T_r - T_c = (Q + Q_o)/(C_1 + C_2)$ does not depend on the coils, so the cold aisle's excess equals the return's, $T_{c,18} - \check{T}_c = \eta_s$. Section 6.4.1 checks the three relations on every draw of the full posterior.

The supply penalty and the set point. By the forward equations (5.3) and the air balance (5.4) the cold aisle of a window is

$$T_c = T_s + (Q + Q_o)\Big[\frac{1}{G_1 + G_2} - \frac{1}{C_1 + C_2}\Big], \tag{6.6}$$

so holding a cold-aisle set point T_c^{sp} at today's heat needs the supply

$$T_s = T_c^{\mathrm{sp}} - (Q_{18} + Q_o(18))\Big[\frac{1}{G_1 + G_2} - \frac{1}{C_1 + C_2}\Big]. \tag{6.7}$$

Set at the commissioned coils and at today's, the two supplies differ by $(Q_{18} + Q_o(18))\,[1/(G_{1,18} + G_{2,18}) - 1/(G_{1,0} + G_{2,0})]$, which is η_s of equation (6.4): the set point's value cancels. The supply penalty is what the coils' loss costs the hall's manipulated variable, whatever the set point, and the chapter needs no set-point number.

The full state. The fit runs on the reduced state. Its full state, the one whose limit of exact rows the reduced state is, carries the seven entries of equation (5.5); each window's return air, two duties, two discharges, cold aisle and rack-1 inlet; and the node's return air, two duties, two discharges and cold aisle: twenty-seven entries under twenty physical rows, seven in each window (the hall balance on the duties, the two coil rows, the two air sides, the mixing and the rack-1 inlet) and the six of equation (6.2). On it every excess is the difference of two entries: $\eta_{A,u} = T_{d,u,18} - \check{T}_{d,u}$, $\eta_{D,u} = D_{u,18} - \check{D}_u$ and $\eta_s = T_{r,18} - \check{T}_r$. The duty and the approach of the drill question are then state entries, read against the node's.

Modeling choices at the graph. The choices of Section 5.1.2 carry over: steady state within the window, the airflow pinned at the units' fixed fan speed, the supply and the other heat as known inputs, and one conductance per coil. Two are new. The record has no discharge controller: the units run at a fixed fan speed and full water flow, no loop holds the cold aisle, and what holds it near its handover value in June is the colder winter supply. This is what lets the discharge and return readings inform the coils, as in Chapter 5, and it puts the supply, the hall's only manipulated variable, into the rows as a known input, whatever sets it. It fails under an active discharge controller, which would throttle each unit's water valve to hold $T_{d,u}$: the discharge would then read its set point, the return air would read $T_c^{\mathrm{sp}} + (Q + Q_o)/(C_1 + C_2)$ whatever the coils, and the coil's loss would show only in the valve position and the water flow, which no row of this record reads. The coil row is the duty to the water, with no water-side losses between the coil and the loop: this is what lets $D_{u,t}$ answer which unit carries the heat, and it fails if a unit's piping gains or loses heat to the hall.

The graph. Figure 6.1 sets the rows of window 18 and the scenario node side by side. On the left is window 18 as in Figure 5.1: the IT load Q_{18} (green) and the other heat feed the return air, the return air feeds the two coils (green, each with its free log-derate), which pass heat to the pinned chilled water and send their discharge air to the cold aisle, and the cold aisle feeds rack 1's inlet; the dashed arrows leave for the five readings, with their values. On the right, inside the dashed frame, is the scenario node: a dashed copy of the return air, the two units and the cold aisle, fed by the same load entry Q_{18} (the long arrow across the top) and the same known inputs, with the coils carrying the commissioned entries $g_{1,0}$ and $g_{2,0}$ of the small window-0 node above it. No dashed arrow leaves the frame: the node has no reading. The red double arrows between the two copies mark the excesses: η_s at the return air and again at the cold aisle, $\eta_{D,u}$ at the coil edges and $\eta_{A,u}$ at the discharges.

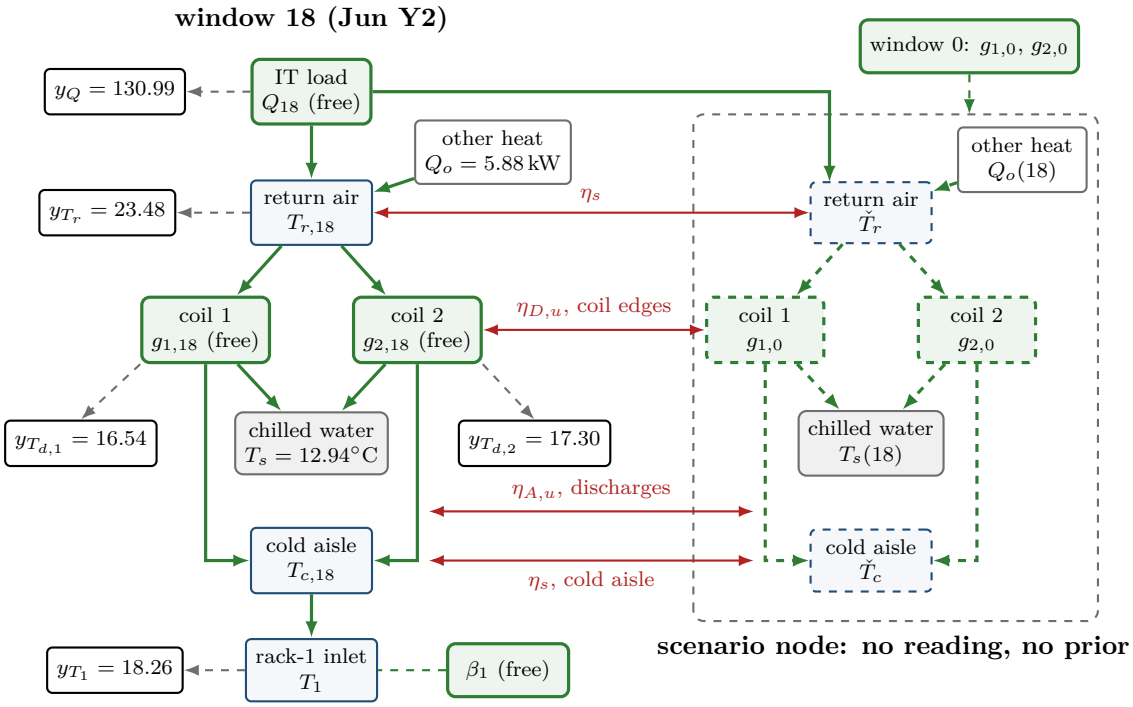


Figure 6.1: The rows of window 18 (left) and the scenario node (right). Free (green): the IT load Q_{18}, today's log-derates $g_{1,18}$ and $g_{2,18}$, the recirculation fraction β_1, and the commissioning entries $g_{1,0}$ and $g_{2,0}$ of window 0, drawn as one small node. Blue: temperatures fixed by the physical rows (5.2). Pinned (gray): the chilled-water supply $T_s(18)$; known input (white): the other heat $Q_o(18)$; also pinned, not drawn, the ratings $\bar{G}_u = 7.0\,\text{kW/K}$ and the capacity rates $C_u = 10.0\,\text{kW/K}$. Observed (dashed gray): the five readings of window 18, in °C and kW; window 0's readings and the three prior rows of Figure 5.1 are not redrawn. Scenario (dashed frame): window 18's rows (6.2) at the same load entry and inputs with the commissioned coils, and no reading. Red: the excesses of equation (6.4).

6.1.3 The asset manager's questions (step 1)

The asset manager asks four questions of the two windows, and step 1 writes each as a probability over the joint state (5.5), with the unknown on the left of the bar and the ten readings and the three prior rows on the right. The third asks why the set-point tile has not moved, and it needs the cold aisle's change written as a sum. By equation (6.6) in window 0 and in window

18, with the excess (6.4),

$$\begin{gathered} T_{c,18} - T_{c,0} = s_{\text{season}} + s_{\text{load}} + \eta_s, \\ s_{\text{season}} = \left[T_s(18) - T_s(0)\right] + \left[Q_o(18) - Q_o(0)\right]\zeta_0, \qquad s_{\text{load}} = \left(Q_{18} - Q_0\right)\zeta_0, \\ \zeta_0 = \frac{1}{G_{1,0} + G_{2,0}} - \frac{1}{C_1 + C_2}, \end{gathered} \tag{6.8}$$

with ζ_0 the cold aisle's rise per kW of heat at the commissioned coils, in K/kW. The season's term is the change of the known inputs, the load's the change of the IT load, both at the commissioned coils, and the coils' term is the supply penalty; the three sum to the change at every point of the state.

Table 6.1 lists the four questions. The first column is the question in the asset manager's words, the first of them the drill question; the second the general form of its category from Chapter 2, for a quantity $f(\mathbf{z})$ derived from the state and a level f^* of it; the third its mathematical form for the hall; the fourth its category; the fifth the computation of Section 2.3.6 that answers it; and the last the subsection where it is answered. "The unit" in the drill question is each unit in turn, because the state carries both; unit 2 is the unit Chapter 5's log flagged. Three further questions are not rows. What the extra effort costs in chiller energy and in reais needs a chiller model the plant does not have, and it is an expected consequence that a later part of the book prices. When the effort stops hiding the loss, the set point lost at a summer supply, is the crossing of Chapter 14. Which measurement would narrow the answers asks nothing new of today's evidence; it is the next measurement for these four questions, Section 6.6.

Table 6.1: The questions asked of the two-unit hall at window 18 about the effort behind its held set point. $\mathbf{y}$ is the ten readings and the three prior rows; $\eta_{A,u}$, η_s and $\eta_{D,u}$ the excesses of equation (6.4); s_{season} and s_{load} the terms of equation (6.8).

Asset manager's question	General form	Mathematical form	Category	Computation	Answered in
The unit still meets its set-point; is it working harder to do so?	$\Pr(f(\mathbf{z}) > f^* \mid \mathbf{y})$; moments; credible interval, (2.12), (2.10), (2.11)	$\Pr(\eta_{A,u} > 0 \mid \mathbf{y})$ and $p(\eta_{A,u} \mid \mathbf{y})$, $\eta_{A,u} = A_{u,18}(\mathbf{z}) - \check{A}_u(\mathbf{z})$, $u = 1, 2$: median and 5–95 % band	prediction under a threshold	posterior	6.3.1
How much colder must the chilled water run, because of the coils?	$\Pr(f(\mathbf{z}) \le f^* \mid \mathbf{y})$; moments; credible interval, (2.9)–(2.11)	$p(\eta_s \mid \mathbf{y})$, $\eta_s = T_{r,18} - \check{T}_r$; $\Pr(\eta_s > 0 \mid \mathbf{y})$	estimation	posterior	6.3.2
Why does the cold aisle still read as it did at handover?	moments and credible intervals of derived quantities, (2.10), (2.11)	$T_{c,18} - T_{c,0} = s_{\text{season}} + s_{\text{load}} + \eta_s$, each term's $p(\cdot \mid \mathbf{y})$	estimation	posterior	6.3.3
Which unit carries the heat the fouled coil no longer passes?	$\Pr(f(\mathbf{z}) > f^* \mid \mathbf{y})$, (2.12)	$p(\eta_{D,1} \mid \mathbf{y})$, $\eta_{D,1} = -\eta_{D,2}$; $\Pr(\eta_{D,1} > 0 \mid \mathbf{y})$	prediction under a threshold	posterior	6.3.4

6.1.4 The evidence as rows (step 2)

The evidence is that of Section 5.1.4, its thirteen rows unchanged: the five readings of each window, the rating prior on the two commissioning log-derates and the survey prior on β_1. No row is added. The scenario node carries no reading and no prior, and nothing reads a duty: the hall has no water-side heat meter.

Which row acts on which entry. The Jacobian has the pattern (5.6). By equations (6.4) and (6.8) the questions' quantities depend on the entries below, with $*$ for an entry that can be nonzero and the columns in the order $(Q_0, g_{1,0}, g_{2,0}, Q_{18}, g_{1,18}, g_{2,18}, \beta_1)$ of equation (5.5):

$$\begin{aligned} \mathbf{g}_{\eta_{A,u}}^{\mathsf{T}} = \mathbf{g}_{\eta_s}^{\mathsf{T}} = \mathbf{g}_{\eta_{D,1}}^{\mathsf{T}} &= (0,\ *,\ *,\ *,\ *,\ *,\ 0), \\ \mathbf{g}_{T_{c,18}-T_{c,0}}^{\mathsf{T}} &= (*,\ *,\ *,\ *,\ *,\ *,\ 0), \end{aligned} \tag{6.9}$$

the last with an entry in the column of Q_0 because the split's load term carries the commissioning load. These are not rows; they are the patterns of the answers. Two facts are read here, before any number. Every excess takes the commissioning coil entries and today's window, so its variance holds a block from each window and the cross term between them, and by the block structure (5.11) that cross term runs through β_1 alone. And no gradient touches β_1: its covariances with the other entries are multiplied by zero, so β_1 is used by the fit, through the rack-1 rows, and by no answer.

Modeling choices at the rows. Three choices make the node a reference for today's window. The node shares the load entry Q_{18}, not the meter's reading: "the same conditions" is the same state entry, and this is what makes the load's error cancel in every excess to first order; read at the meter's 130.99 kW instead, the node would carry the meter's 3.0 kW error with nothing to cancel it. The choice fails only where the asset manager's "same conditions" means a load other than today's, such as a planned one, which is then a stated input of the node. The commissioned model is the commissioning fit's coils, $g_{1,0}$ and $g_{2,0}$, not the rating: this is what the drill's "commissioned model" asks for, and it makes the reference carry the commissioning test's own error; it fails if the commissioning test did not represent the coils as built, and Section 6.3.1 gives the excess against rating in one line, as Chapter 5 gave the derate factor. And a coil's conductance is taken not to depend on the load, the choice of *Reliability Engineering by Derivation*, Chapter 20, for its scenario node: this is what lets the node carry $g_{u,0}$ at a metered heat 15.1 % above the commissioning test's, and it fails if the coil's conductance moves with its water flow or its air-side temperatures, which two windows cannot test.

6.2 The posterior of the state (step 3)

All four questions of Table 6.1 are read from one posterior, that of the joint state (5.5) under the thirteen rows of Section 5.1.4, and the scenario node is read from it at its MAP state.

6.2.1 The posterior read unchanged

Nothing is fitted. The scenario node adds no row: it has no reading and no prior, so the residual vector (5.8), its Jacobian, the information matrix $\mathbf{J}^{\mathsf{T}}\mathbf{J}$ of the normalized rows ($\mathbf{J}^{\mathsf{T}}\mathbf{W}\mathbf{J}$ with unnormalized rows and their weights) and its inverse are those of Section 5.2.3. The posterior is the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ with the MAP state of Table 5.4 and the covariance (5.10), at $2\mathcal{L}(\hat{\mathbf{z}}) = 0.9849$. This is the one fit of the condition record that this part of the book reads: a second question of the same two windows is a second function of the same posterior.

The block the answers use. By the patterns (6.9) every excess takes the five entries $(g_{1,0}, g_{2,0}, Q_{18}, g_{1,18}, g_{2,18})$, and the split's load term also Q_0. The block $\boldsymbol{\Sigma}_{(5)}$ of $\boldsymbol{\Sigma}$ on those five entries, in that order, is

$$\boldsymbol{\Sigma}_{(5)} = \begin{bmatrix} 1.5696\times10^{-3} & -6.8900\times10^{-4} & 5.8158\times10^{-4} & 3.4901\times10^{-5} & 3.5380\times10^{-5} \\ -6.8900\times10^{-4} & 1.6598\times10^{-3} & 5.8124\times10^{-4} & 3.4880\times10^{-5} & 3.5359\times10^{-5} \\ 5.8158\times10^{-4} & 5.8124\times10^{-4} & 7.7031 & 2.9026\times10^{-2} & 3.5372\times10^{-2} \\ 3.4901\times10^{-5} & 3.4880\times10^{-5} & 2.9026\times10^{-2} & 1.2080\times10^{-3} & -6.1601\times10^{-4} \\ 3.5380\times10^{-5} & 3.5359\times10^{-5} & 3.5372\times10^{-2} & -6.1601\times10^{-4} & 1.4797\times10^{-3} \end{bmatrix}, \tag{6.10}$$

the load entry in kW and kW^2 and the others dimensionless. Its correlations read by sign, each fixed by the rows (Section 5.2.3). The two log-derates of one window are correlated negatively, -0.4269 at window 0 and -0.4607 at window 18: the return air knows the sum of the conductances better than their split. Today's load is correlated positively with today's log-derates, $+0.3009$ and $+0.3313$: a larger load needs larger conductances to give the same return air. And the two windows are correlated by $+0.0226$ to $+0.0253$ between their coil entries, and by $+0.0051$ and $+0.0053$ between today's load and the commissioning coils, through β_1 alone.

The full posterior. The full posterior beside the Laplace normal is that of Section 5.2.4: the same $N_{\mathrm{d}} = 200\,000$ draws from the seeded generator, with the same weights (5.13) and the effective sample size 197 321. Each draw is pushed through the exact forward equations of both windows and of the node (6.3), so every answer of step 4 has a full-posterior value from the same draws, with a Monte Carlo standard error by batch means.

6.2.2 The scenario node at the MAP state

The node's quantities at $\hat{\mathbf{z}}$ follow from equations (5.3) and (6.3). Table 6.2 sets them beside the two windows: each row is a quantity, the columns its value in window 0, in window 18 and at the scenario node, and the excess of window 18 over the node. The heat to remove is 118.4739 kW at window 0 and 136.0452 kW at window 18, and the total conductance is 13.4418 kW/K at window 0 and 12.7022 kW/K at window 18. The return air, the cold aisle and the return approach carry the same excess, $+0.5893$ K, the supply penalty, as the third relation of Section 6.1.2 requires. The two duty excesses are equal and opposite: unit 1's share of the coil heat is 0.50830 at the commissioned coils and 0.52793 today. By equation (6.7), the commissioned coils would hold today's cold aisle at today's heat with the supply at 13.5287°C, against today's 12.94°C.

Table 6.2: The two windows and the scenario node at the MAP state; the excess is window 18 less the node.

Quantity [unit]	Window 0 (commissioning)	Window 18 (Jun Y2)	Scenario node	Excess
Return air T_r [°C]	22.8139	23.6497	23.0604	+0.5893
Cold aisle T_c [°C]	16.8902	16.8475	16.2581	+0.5893
Return approach $A_r = T_r - T_s$ [K]	8.8139	10.7104	10.1211	+0.5893
Approach of unit 1, A_1 [K]	2.7918	3.5281	3.2059	+0.3222
Approach of unit 2, A_2 [K]	2.9885	4.2881	3.4317	+0.8564
Duty of unit 1, D_1 [kW]	60.2203	71.8226	69.1517	+2.6709
Duty of unit 2, D_2 [kW]	58.2537	64.2226	66.8935	−2.6709

The gradients at the MAP state. Table 6.3 gives the gradient of each question's quantity at $\hat{\mathbf{z}}$; each row is a quantity, each column an entry of the joint state. Each excess has zeros in the columns of Q_0 and β_1, as the patterns (6.9) say, and opposite signs on the two windows' coil entries. For the two approaches and the supply penalty a larger commissioned conductance raises the commissioned model's performance and so the excess, and a larger conductance today lowers it; the duty shift's signs follow the split of the heat instead, as the end of this paragraph says. The supply penalty's coil entries have the closed forms $(Q_{18} + Q_o(18))\, G_{u,0}/(G_{1,0} + G_{2,0})^2$ on the commissioning entries and $-(Q_{18} + Q_o(18))\, G_{u,18}/(G_{1,18} + G_{2,18})^2$ on today's, 5.1445, 4.9765, -5.6544 and -5.0560. Its slope on today's load is $\partial\eta_s/\partial Q_{18} = 1/(G_{1,18} + G_{2,18}) - 1/(G_{1,0} + G_{2,0}) = 0.004332$ K/kW: the node and today's window carry the same load entry, and the difference removes it to first order. With the covariance (6.10), that slope's own term is 1.45×10^{-4} K^2, 0.18 % of the variance of η_s. The duty shift's gradient is antisymmetric across

the two coils of each window, ∓ 34.00193 at commissioning and ± 33.90516 today: it reads the split of the heat between the coils, not its total.

Table 6.3: The gradients of the questions' quantities at the MAP state, by entry of the joint state (5.5). The column headers carry [–] because the unit of each cell is the row quantity's unit per unit of the column's entry: K/kW in the load columns of the K-quantities and kW/kW for $\eta_{D,1}$; K, or kW for $\eta_{D,1}$, in the log-derate and β_1 columns.

Quantity [unit]	Q_0 [–]	$g_{1,0}$ [–]	$g_{2,0}$ [–]	Q_{18} [–]	$g_{1,18}$ [–]	$g_{2,18}$ [–]	β_1 [–]
$\eta_{A,1}$ [K]	0	8.54473	1.57634	0.00237	−9.04487	−1.66552	0
$\eta_{A,2}$ [K]	0	1.74434	8.37673	0.00630	−2.26384	−8.44655	0
η_s [K]	0	5.14454	4.97653	0.00433	−5.65436	−5.05603	0
$\eta_{D,1}$ [kW]	0	−34.00193	34.00193	0.01963	33.90516	−33.90516	0
$T_{c,18} - T_{c,0}$ [K]	−0.02439	4.48008	4.33378	0.02873	−5.65436	−5.05603	0
s_{season} [K]	0	0.08022	0.07760	0	0	0	0
s_{load} [K]	−0.02439	−0.74467	−0.72036	0.02439	0	0	0

6.3 The answers (step 4)

Every answer is read from the pair $(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ of Section 6.2.1, and that value is the one reported. Each quantity is carried by its gradient of Table 6.3, with variance $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$ by the delta method (4.53) and a normal band. The full posterior of the same draws stands beside each value, a median beside a median.

6.3.1 The approach of each unit against the commissioned model

Asset manager's question. The unit still meets its set-point; is it working harder to do so? It is asked of each unit in turn, unit 2 first, because Chapter 5's log flagged its coil.

General form. A prediction under a threshold is a tail of a derived quantity under the posterior, equation (2.12),

$$\Pr(f(\mathbf{z}) \geq f^* \mid \mathbf{y}) = \int \mathbf{1}\big[f(\mathbf{z}) \geq f^*\big]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z},$$

and the size of the quantity is summarized by its moments, equation (2.10), and its credible interval, equation (2.11), at the levels 0.05 and 0.95, the 5–95 % band.

Specialized form. The question is equation (2.12), with equations (2.10) and (2.11) for its size, with $\mathbf{z}$ the joint state (5.5), $f(\mathbf{z}) = \eta_{A,u}$ of equation (6.4), the event $\eta_{A,u} > 0$ at the level $f^* = 0$, and $p(\mathbf{z} \mid \mathbf{y})$ the Laplace normal of Section 6.2.1. On the full state of Section 6.1.2 the excess is the difference of two entries, $T_{d,u,18} - \breve{T}_{d,u}$, the drill's approach as a state entry; on the reduced state it is a derived quantity with the gradient $\mathbf{g}_\eta$ of Table 6.3. For any of the excesses η, the delta method gives

$$\hat{\eta} = \eta(\hat{\mathbf{z}}), \qquad \mathrm{Var}(\eta \mid \mathbf{y}) = \mathbf{g}_\eta^{\mathsf{T}} \boldsymbol{\Sigma}\, \mathbf{g}_\eta, \qquad \Pr(\eta > 0 \mid \mathbf{y}) = \Phi\!\left(\frac{\hat{\eta}}{\mathrm{sd}(\eta \mid \mathbf{y})}\right), \tag{6.11}$$

with the band $\hat{\eta} \pm 1.6449\, \mathrm{sd}(\eta \mid \mathbf{y})$. Split by the blocks of $\boldsymbol{\Sigma}$, the variance is the commissioning block, today's block and twice the cross term between them,

$$\mathbf{g}_\eta^{\mathsf{T}} \boldsymbol{\Sigma}\, \mathbf{g}_\eta = \mathbf{g}^{(0)\mathsf{T}} \boldsymbol{\Sigma}_{0,0}\, \mathbf{g}^{(0)} + \mathbf{g}^{(18)\mathsf{T}} \boldsymbol{\Sigma}_{18,18}\, \mathbf{g}^{(18)} + 2\, \mathbf{g}^{(0)\mathsf{T}} \boldsymbol{\Sigma}_{0,18}\, \mathbf{g}^{(18)}, \tag{6.12}$$

with $\mathbf{g}^{(0)}$ and $\mathbf{g}^{(18)}$ the parts of $\mathbf{g}_\eta$ in the columns of window 0, $(Q_0, g_{1,0}, g_{2,0})$, and of window 18, $(Q_{18}, g_{1,18}, g_{2,18})$, and $\mathbf{\Sigma}_{0,0}$, $\mathbf{\Sigma}_{18,18}$ and $\mathbf{\Sigma}_{0,18}$ the blocks of Section 5.2.3; the column of β_1, where every gradient is zero, drops out. One choice enters here: "working harder" is read as the posterior probability of any excess, the level $f^* = 0$. This is what makes the question one the record can answer without a decision rule, and it is not a decision: an excess large enough to act on needs a decision level, Chapter 19's question, and the assurance level κ of equation (2.12) is not set here.

Building blocks. The pair $(\hat{\mathbf{z}}, \mathbf{\Sigma})$ of Section 6.2.1, with the block (6.10); the gradients of Table 6.3; the moments and tail (6.11) and the budget (6.12); the normal distribution function Φ; the air-side identity (6.5).

Unit 2. By equation (6.11), $\hat{\eta}_{A,2} = 0.8564\,\mathrm{K}$ with variance $1.7745 \times 10^{-1}\,\mathrm{K}^2$, standard deviation $0.4213\,\mathrm{K}$ and band $[0.1635,\ 1.5493]\,\mathrm{K}$, and

$$\Pr(\eta_{A,2} > 0 \mid \mathbf{y}) = \Phi(2.0330) = 0.979.$$

The full posterior gives a median of $0.8591\,\mathrm{K}$, a band of $[0.1566,\ 1.5425]\,\mathrm{K}$ and a probability of 0.97753. Unit 2's approach, $4.2881\,\mathrm{K}$, is 25.0 % above the commissioned model's $3.4317\,\mathrm{K}$: unit 2 works harder in approach, not in duty (Section 6.3.4).

Unit 1. $\hat{\eta}_{A,1} = 0.3222\,\mathrm{K}$ with standard deviation $0.4189\,\mathrm{K}$ and band $[-0.3668,\ 1.0113]\,\mathrm{K}$, and $\Pr(\eta_{A,1} > 0 \mid \mathbf{y}) = \Phi(0.7692) = 0.779$; the full posterior gives $0.3221\,\mathrm{K}$, $[-0.3745,\ 1.0066]\,\mathrm{K}$ and 0.77784. The approach is 10.1 % above the commissioned model's, and the excess is not demonstrated. The two excesses are correlated by -0.1137.

Where each excess comes from. The air-side identity (6.5) splits each excess into the hall's supply penalty and the unit's own duty shift. For unit 2 it is $0.5893\,\mathrm{K}$ from the hall plus $0.2671\,\mathrm{K}$ from its duty shift, because it passes less heat than the commissioned model; for unit 1 it is $0.5893\,\mathrm{K}$ from the hall less $0.2671\,\mathrm{K}$, because it carries more. Unit 1's approach has widened through the return air it shares with unit 2, while its own coil's loss is not demonstrated: Chapter 5 gave it a probability of a loss of 0.6403.

The budget of the variance. By equation (6.12), unit 2's variance is 57.0 % the commissioning block ($g_{2,0}$ 65.6 %, $g_{1,0}$ 2.7 %, their cross term −11.3 %), 47.3 % today's block ($g_{2,18}$ 59.5 %, $g_{1,18}$ 3.5 %, Q_{18} 0.2 %, the cross terms −15.9 %) and −4.3 % the term between the windows. Unit 1's mirrors it: 57.1 %, 47.2 % and −4.3 %. The cross terms within each window are negative because the gradient has the same sign on both coils of a window and the coils of a window are anticorrelated, the correlations -0.4269 and -0.4607 of the covariance (6.10). The term between the windows is negative because the gradient has opposite signs on the two windows' entries and the windows are correlated positively, through β_1. More than half of each excess's variance is the commissioning test's.

Against rating, and the readings taken as exact. Read against the rating instead of the commissioning fit, the rated coils' approach at today's conditions is $2.9153\,\mathrm{K}$ for both units, and the excesses would be 0.6129 and $1.3729\,\mathrm{K}$. Taken as exact, the readings give the commissioned model at today's heat without a fit: at fixed conductances each approach is proportional to the heat to remove, by equation (5.18), so the commissioning approaches read from the record, scaled by the metered heat ratio 1.1509, give 3.2917 and $3.5449\,\mathrm{K}$ against today's 3.60 and $4.36\,\mathrm{K}$: excesses of 0.3090 and $0.8158\,\mathrm{K}$.

The two posteriors are plotted in Figure 6.2: the horizontal axis is the excess approach, the vertical axis its posterior density; the shaded areas are the 5–95 % bands, the solid gray line marks zero, no excess, the dashed lines the readings taken as exact, and the dotted lines the stated truth of the record, check markers for Section 6.4.4. Unit 2's density lies almost wholly to the right of zero; unit 1's straddles it. Each readings-as-exact value lies close to its density's peak.

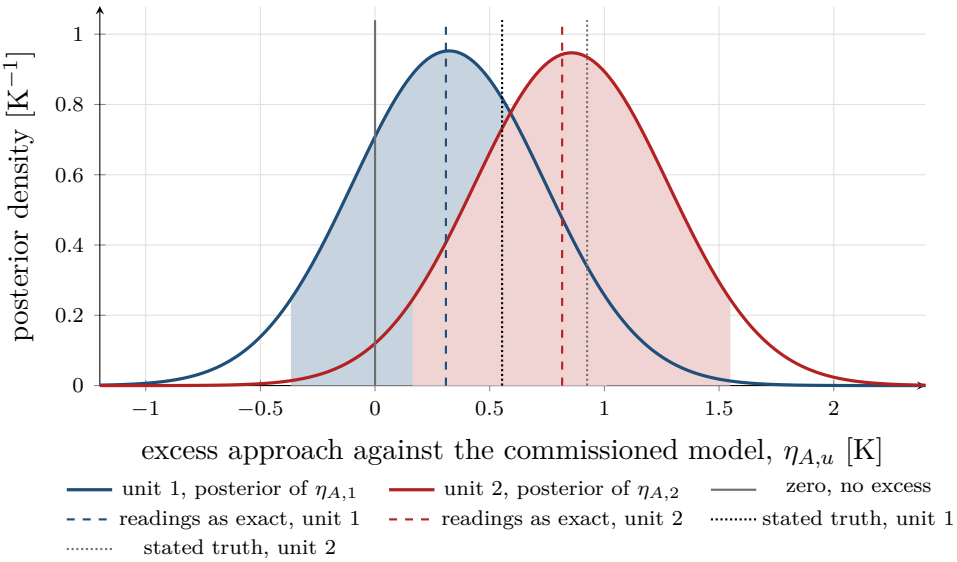


Figure 6.2: The posterior densities of the excess approach of each unit, the normals of equation (6.11): unit 1 with $\eta_{A,1} \sim \mathcal{N}(0.3222,\ 0.4189^2)$, unit 2 with $\eta_{A,2} \sim \mathcal{N}(0.8564,\ 0.4213^2)$, in K. Shaded: the 5–95 % bands. Solid gray: no excess. Dashed: the readings taken as exact. Dotted: the stated truth of the record, a check marker (Section 6.4.4).

6.3.2 The supply penalty

Asset manager's question. How much colder must the chilled water run, because of the coils? It asks what the coils' loss costs the hall's manipulated variable.

General form. An estimation question about a derived quantity asks for its posterior distribution, equation (2.9),

$$\Pr(f(\mathbf{z}) \le f^* \mid \mathbf{y}) = \int \mathbf{1}\left[f(\mathbf{z}) \le f^*\right] p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z},$$

summarized by its moments, equation (2.10), and by its credible interval, equation (2.11), at the levels 0.05 and 0.95.

Specialized form. The question is equations (2.9), (2.10) and (2.11) with $f(\mathbf{z}) = \eta_s$ of equation (6.4), whose meaning is the cancellation of the set point in equation (6.7), and $p(\mathbf{z} \mid \mathbf{y})$ the Laplace normal; on the full state of Section 6.1.2 it is $T_{r,18} - \check{T}_r$. Its value, variance and probability of being positive are equation (6.11) with $\eta = \eta_s$. The return approach that the dashboard put down to the load is a function of the state too: by the first forward equation, $A_{r,t} = T_{r,t} - T_s(t) = (Q_t + Q_o(t))/(G_{1,t} + G_{2,t})$, and its change since commissioning splits exactly into the IT load, the other heat and the coils,

$$A_{r,18} - A_{r,0} = \frac{Q_{18} - Q_0}{G_{1,0} + G_{2,0}} + \frac{Q_o(18) - Q_o(0)}{G_{1,0} + G_{2,0}} + \eta_s, \tag{6.13}$$

the first two terms at the commissioned coils.

Building blocks. The pair $(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ with the block (6.10); the gradient of η_s in Table 6.3; equations (6.11) and (6.12); the split (6.13).

The penalty. By equation (6.11), $\hat{\eta}_s = 0.5893\,\mathrm{K}$ with standard deviation $0.2797\,\mathrm{K}$ and band $[0.1293,\ 1.0493]\,\mathrm{K}$, and

$$\Pr(\eta_s > 0 \mid \mathbf{y}) = \Phi(2.1073) = 0.982;$$

the full posterior gives a median of $0.5905\,\mathrm{K}$, a band of $[0.1237,\ 1.0435]\,\mathrm{K}$ and a probability of 0.98090. The chilled water runs $0.5893\,\mathrm{K}$ colder than the commissioned coils would need. It is the best-determined excess, and the budget (6.12) says why: η_s depends on the sum of the conductances, its gradient has the same sign on both coils of each window, and the coils of a window are anticorrelated, so the cross terms within the windows, $-45.1\,\%$ at commissioning and $-48.8\,\%$ today, remove most of each block. The commissioning block is $60.6\,\%$ of the variance ($g_{1,0}$ $53.1\,\%$, $g_{2,0}$ $52.6\,\%$), today's $49.1\,\%$ ($g_{1,18}$ $49.4\,\%$, $g_{2,18}$ $48.4\,\%$, Q_{18} $0.2\,\%$), and the term between the windows $-9.7\,\%$. The readings know the hall's total conductance better than its split, and the supply penalty asks for the total.

The budgets of η_s and of $\eta_{A,2}$ are taken apart term by term in Figure 6.3. Each row is one block of the variance (6.12), the commissioning block, today's block and the term between the windows, for η_s in the upper three rows and for $\eta_{A,2}$ in the lower three; the bar to the right of zero stacks the block's diagonal terms, one segment per entry, the bar to the left its cross terms, and the diamond marks the block's net share. The horizontal axis is in percent of the quantity's variance. For η_s the left bars cancel nearly half of the right ones; for $\eta_{A,2}$ they cancel a sixth to a quarter, because its gradient weights one coil far more than the other.

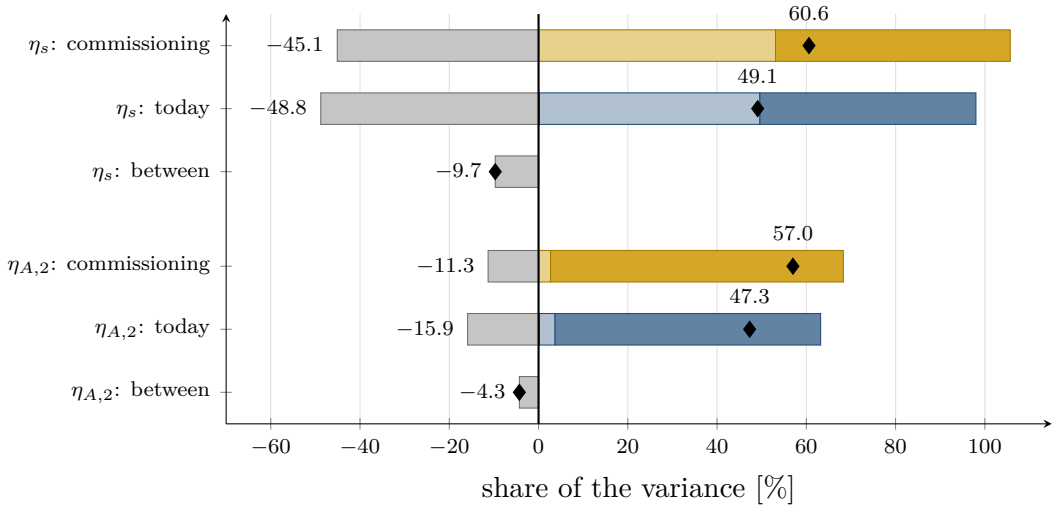


Figure 6.3: The variance budgets (6.12) of the supply penalty η_s (upper three rows) and of unit 2's excess approach $\eta_{A,2}$ (lower three), in percent of each variance. Right of zero: the diagonal terms of the block, stacked, coil 1's entry light and coil 2's dark, gold for the commissioning entries, blue for today's, green for today's load. Left of zero (gray): the cross terms. Diamond: the block's net share; the three net shares of each quantity sum to 100 %.

The return approach since handover. The dashboard put the $1.83\,\mathrm{K}$ rise of the return air over the supply down to the load. By equation (6.13) at $\hat{\mathbf{z}}$ the rise is $1.8965\,\mathrm{K}$ (standard deviation $0.2524\,\mathrm{K}$): the IT load's $1.4650\,\mathrm{K}$ ($77.2\,\%$), the other heat's $-0.1578\,\mathrm{K}$ ($-8.3\,\%$) and

the coils' 0.5893 K (31.1 %). The asset manager's reading assigns the whole rise to the load; about a third of it is the coils.

6.3.3 The cold aisle held at its handover value

Asset manager's question. Why does the cold aisle still read as it did at handover? It asks what the set-point tile's −0.05 K is made of.

General form. An estimation question about derived quantities is summarized by their moments, equation (2.10),

$$\mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}] = \int f(\mathbf{z})\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z},$$

and the variance beside it, and by their credible intervals, equation (2.11), at the levels 0.05 and 0.95.

Specialized form. The question is equations (2.10) and (2.11) for each term of the split (6.8) and for their sum, with $p(\mathbf{z} \mid \mathbf{y})$ the Laplace normal. Each term's variance is $\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}$ with its gradient of Table 6.3, and two terms a and b have the covariance $\mathbf{g}_a^\mathsf{T}\boldsymbol{\Sigma}\,\mathbf{g}_b$; the sum's variance is the sum of the three variances and twice the three covariances. One choice enters here: the season's term is the change of the known inputs, T_s and Q_o, at the commissioned coils. This is what makes the split exact for one pair of windows, with every term a quantity of the state, and it is not a seasonal model: the season across the whole record is Chapter 13's question.

Building blocks. The pair $(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$; the split (6.8); the gradients of Table 6.3; the covariances $\mathbf{g}_a^\mathsf{T}\boldsymbol{\Sigma}\,\mathbf{g}_b$.

The split. Table 6.4 gives the three terms and their sum: each row is a term, with its value, its standard deviation and its band. The cold aisle of the fit is 16.8902°C at window 0 and 16.8475°C at window 18, a change of −0.0427 K with standard deviation 0.2186 K.

Table 6.4: The cold aisle's change since commissioning split by equation (6.8) into season, IT load and coils, at the MAP state; bands are 5–95 %.

Term	Value [K]	sd [K]	Band [K]
Season, s_{season}: supply −1.0607 K (known), other heat −0.0517 K	−1.1124	0.0034	[−1.1180, −1.1068]
IT load, $s_{\text{load}} = (Q_{18} - Q_0)\,\zeta_0$	+0.4804	0.1116	[+0.2969, +0.6639]
Coils, η_s	+0.5893	0.2797	[+0.1293, +1.0493]
Change of the cold aisle, $T_{c,18} - T_{c,0}$	−0.0427	0.2186	[−0.4022, +0.3168]

The set-point tile's −0.05 K is the season's −1.1124 K against the load's and the coils' +1.0697 K together. The probability that the cold aisle has warmed at all is $\Pr(T_{c,18} > T_{c,0} \mid \mathbf{y}) = 0.423$, and 0.42401 by the full posterior: the tile is consistent with no change, while the coils' term is positive with probability 0.982. The covariances read by sign. The load's term and the coils' are correlated by −0.7018 (covariance $-2.1892 \times 10^{-2}\,\mathrm{K}^2$): both carry the commissioned total $G_{1,0} + G_{2,0}$, and a larger total shrinks the load's term, through ζ_0, and enlarges the coils', through η_s. The season's term is nearly fixed, with standard deviation 0.0034 K, because the coils enter it only through ζ_0 on the small change of the other heat; its covariances with the load's and the coils' terms are -2.3372×10^{-4} and $6.7961 \times 10^{-4}\,\mathrm{K}^2$.

Figure 6.4 shows the split as a waterfall. The vertical axis is the cold aisle in °C; the first bar is the handover value of the fit, each of the next three moves the level by one term of the

split, and the last bar is today's value. The error bars on the load's and the coils' steps are their marginal 5–95 % bands, drawn from each step's start. The dashed line marks the scenario node's cold aisle, which the season's and the load's steps reach and the coils' step leaves. The season's step falls by about as much as the other two rise together.

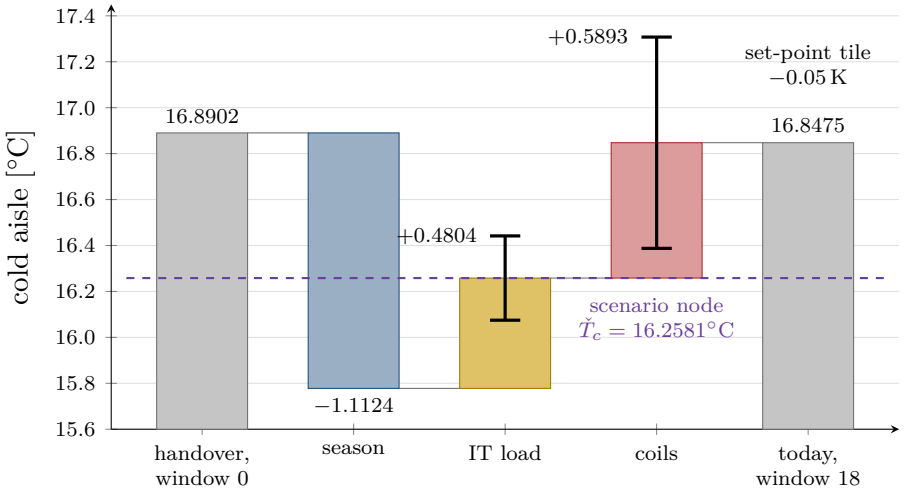


Figure 6.4: The cold aisle's change since commissioning as a waterfall of the split (6.8) at the MAP state: from the handover value 16.8902°C through the season (−1.1124 K), the IT load (+0.4804 K) and the coils (+0.5893 K, the supply penalty) to today's 16.8475°C. Error bars: the marginal 5–95 % bands of the load's and the coils' terms. Dashed: the scenario node's cold aisle. The set-point tile's change from the readings is noted above today's bar.

6.3.4 The duty each unit carries

Asset manager's question. Which unit carries the heat the fouled coil no longer passes? It asks where the effort goes, unit by unit.

General form. A prediction under a threshold is a tail of a derived quantity under the posterior, equation (2.12),

$$\Pr(f(\mathbf{z}) \geq f^* \mid \mathbf{y}) = \int \mathbf{1}\big[f(\mathbf{z}) \geq f^*\big]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}.$$

Specialized form. The question is equation (2.12) with $f(\mathbf{z}) = \eta_{D,1}$, $f^* = 0$ and $p(\mathbf{z} \mid \mathbf{y})$ the Laplace normal. By the duty (6.1) in window 18 and at the node,

$$\eta_{D,1} = (Q_{18} + Q_o(18))\Big[\frac{G_{1,18}}{G_{1,18} + G_{2,18}} - \frac{G_{1,0}}{G_{1,0} + G_{2,0}}\Big], \qquad \eta_{D,2} = -\eta_{D,1}, \tag{6.14}$$

the second by the identity (6.5). The heat to remove is positive, and with equal ratings the share $G_1/(G_1 + G_2) = 1/(1 + \mathrm{e}^{g_2 - g_1})$ falls as $g_2 - g_1$ rises, so

$$\eta_{D,1} > 0 \iff g_{2,18} - g_{1,18} < g_{2,0} - g_{1,0} \iff \ln h_{2,18} < \ln h_{1,18}, \tag{6.15}$$

by the index (5.1): the event is Chapter 5's third question, whose linear form is the difference of equation (5.17). The share moves toward unit 1 exactly when coil 2 has fallen more. One choice enters here: the probability is reported as Chapter 5's, read through that linear form, because the two are one event and one event takes one value across the two chapters; the gradient of $\eta_{D,1}$, a linearization of a nonlinear function of the same event, differs from it only in the fourth decimal (Exercise 6.4). The size of the shift is read by equation (6.11) with $\eta = \eta_{D,1}$.

Building blocks. The pair $(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$; the gradient of $\eta_{D,1}$ in Table 6.3; equations (6.11) and (6.12); the event equivalence (6.15) and the variance of equation (5.17).

The shift. By equation (6.11), $\hat{\eta}_{D,1} = 2.6709\,\text{kW}$ with standard deviation 3.1348 kW and band $[-2.4855,\ 7.8272]\,\text{kW}$, and $\eta_{D,2} = -2.6709\,\text{kW}$ with the same standard deviation, correlated with it by -1.0000. By equation (6.15) and Chapter 5's linear form,

$$\Pr(\eta_{D,1} > 0 \mid \mathbf{y}) = \Pr(\ln h_{2,18} < \ln h_{1,18} \mid \mathbf{y}) = 0.803,$$

and on the full posterior the two events are the same draw by draw, with the probability 0.802758 of Chapter 5; the full posterior's median shift is 2.6821 kW and its band $[-2.4844,\ 7.8636]\,\text{kW}$. Today's duties are 71.8226 and 64.2226 kW; the commissioned model at today's heat gives 69.1517 and 66.8935 kW, and the even split 68.0226 kW. Unit 2, the unit Chapter 5 flagged, carries less, not more: it works harder in approach, not in duty, with a wider approach for less heat. Unit 1's duty has grown by 11.6023 kW since commissioning: the heat's growth at its commissioned share, 8.9315 kW, and the shift, 2.6709 kW. The even split's 8.97 kW is close to unit 1's load term and blind to the shift.

The budget of the shift. By equation (6.12) the commissioning block is 54.2 % of the variance ($g_{1,0}$ 18.5 %, $g_{2,0}$ 19.5 %, their cross term +16.2 %), today's block 45.8 % ($g_{1,18}$ 14.1 %, $g_{2,18}$ 17.3 %, the cross terms +14.3 %) and the term between the windows 0.0 %. The gradient is antisymmetric across the coils of each window, so within each window it meets the negative covariance with opposite signs and the cross terms add: the record knows the split worse than the sum. Between the windows the four covariances are nearly equal and positive, and the antisymmetric gradient cancels them.

The duties are compared in Figure 6.5. The horizontal axis is the duty in kW; for each unit, from the top, the bars give the commissioning duty, the commissioned model at today's heat, and today's duty, whose error bar is the 5–95 % band of the duty shift added to the commissioned model's value. The dashed line is the even split and the dotted marks the stated truth of today's duties. Unit 1's today bar ends to the right of its commissioned model's; unit 2's ends to the left of it, and of the even split.

6.4 The checks (step 5)

Each answer of Section 6.3 rests on approximations and choices that this hall lets be tested against the calculation they stand for.

6.4.1 The full posterior against the Laplace answers

The reweighted draws have been set beside every answer of step 4. Their Monte Carlo standard errors, by batch means, are at most 0.0011 K on the medians (0.0089 kW for the duty shift), at most 0.0023 K on the band ends (0.0177 kW), and at most 0.00112 on the probabilities. Against them, the draws resolve four gaps from the Laplace answers. The probability of a supply penalty is 0.9809 against 0.9825 (standard error 0.0004), and that of unit 2's excess approach 0.9775 against 0.9790 (0.0004). Both ends of the bands of the three excesses in K lie to the left of the Laplace ends, by up to 0.0077 K, against standard errors of at most 0.0023 K. And the median of $\eta_{A,2}$ lies 0.0027 K to the right (0.0010). The draws do not resolve the gap in unit 1's probability, 0.77784 against 0.77911 with a standard error of 0.00112, or in the median of the duty shift, 0.0112 against 0.0089 kW; the gap at the duty shift's upper band end, 0.0363 kW against a standard error of 0.0175 kW, is resolved. The largest gaps are 0.0016 in a probability and 0.0077 K at a band end: the Laplace normal is the posterior to within these gaps, and

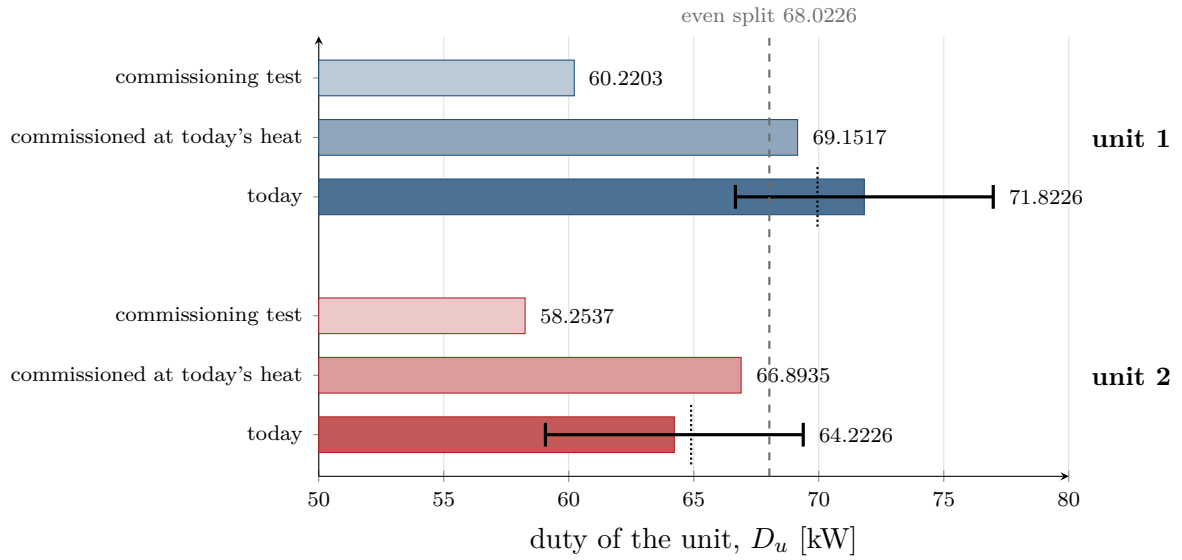


Figure 6.5: The duty of each unit, equation (6.1), at the MAP state: at the commissioning test, at the scenario node (the commissioned model at today's heat), and today. Error bars on today's bars: the 5–95 % band of the duty shift (6.14), $[-2.4855,\ 7.8272]$ kW for unit 1 and its negative for unit 2, added to the commissioned model's value. Dashed: the even split of today's heat to remove. Dotted: the stated truth of today's duties, a check marker (Section 6.4.4).

the full posterior beside it is the same model without the Gaussian approximation, up to its Monte Carlo error. On every draw the identities hold to round-off: the largest $|\eta_{D,1} + \eta_{D,2}|$ is 9.9×10^{-14} kW, the largest departure from the air-side identity (6.5) 4.3×10^{-15} K and from the split (6.8) 1.4×10^{-14} K; and the event $\eta_{D,1} > 0$ is $\ln h_{2,18} < \ln h_{1,18}$ on every draw, as equation (6.15) says.

6.4.2 The limit of exact rows and the curvature term

The answers are read on the reduced state, which equation (4.48) makes the limit $\sigma_p \to 0$ of a full state. The check fits the chapter's own full state of Section 6.1.2, twenty-seven entries with the duties, the discharges and the scenario node among them, under its twenty physical rows at a common finite tolerance beside the readings and the priors, and lets the tolerance shrink. On it each excess is read as the difference of two entries. Table 6.5 gives the result: each row is a tolerance, the columns each excess with its standard deviation, and the last row the reduced state. At 0.1 the spreads are wider by the discrepancy the rows allow; at 0.01 and 0.001 the full state reproduces the reduced state to the fourth and fifth decimals. Each fit took five iterations.

Table 6.5: The limit of exact physical rows on the twenty-seven-entry full state that carries the duties and the scenario node: each excess as a difference of two entries, with its standard deviation, at three tolerances, against the reduced state.

σ_p [kW or K]	$\eta_{A,1}$ [K]	sd [K]	$\eta_{A,2}$ [K]	sd [K]	η_s [K]	sd [K]	$\eta_{D,1}$ [kW]	sd [kW]
0.1	0.32110	0.42133	0.85533	0.42366	0.58808	0.28297	2.67067	3.13895
0.01	0.32222	0.41895	0.85640	0.42128	0.58931	0.27969	2.67087	3.13488
0.001	0.32223	0.41893	0.85641	0.42125	0.58932	0.27965	2.67087	3.13484
reduced state	0.32223	0.41893	0.85641	0.42125	0.58932	0.27965	2.67087	3.13484

The covariance (5.10) is the Gauss–Newton form, without the curvature term $\mathbf{C}$ of equa-

tion (4.44). Chapter 5's Newton covariance, from the full Hessian of the objective at $\hat{\mathbf{z}}$, gives the standard deviations 0.41872, 0.42088 and 0.27962 K and 3.13123 kW for the four excesses, against Gauss–Newton's 0.41893, 0.42125, 0.27965 and 3.13484: ratios of 0.9995, 0.9991, 0.9999 and 0.9988. The Gauss–Newton covariance is the one reported.

6.4.3 Sensitivity to the priors

Table 6.6 repeats the fit under the priors of Table 5.8 and reads the four excesses from each refit. Each row is one prior; each cell gives the excess, its probability of being positive by its gradient, and its shift from the chapter's value in units of its posterior standard deviation. Every row moves every answer by 0.34 standard deviations or less: the readings set the answers.

Table 6.6: The four excesses under other priors: in each cell the value, the probability of a positive excess by the quantity's gradient, and the shift from the chapter's value in units of its posterior standard deviation. Values in K, and in kW for $\eta_{D,1}$; probabilities and shifts dimensionless.

Prior	$\eta_{A,1}$ [K; –; sd]	$\eta_{A,2}$ [K; –; sd]	η_s [K; –; sd]	$\eta_{D,1}$ [kW; –; sd]
rating sd 0.05	0.3639; 0.8337; +0.10	1.0011; 0.9961; +0.34	0.6825; 0.9958; +0.33	3.1859; 0.8790; +0.16
rating sd 0.15 (the chapter's)	0.3222; 0.7791; 0.00	0.8564; 0.9790; 0.00	0.5893; 0.9825; 0.00	2.6709; 0.8029; 0.00
rating sd 0.30	0.3199; 0.7738; −0.01	0.8344; 0.9742; −0.05	0.5772; 0.9796; −0.04	2.5724; 0.7888; −0.03
no rating prior	0.3194; 0.7721; −0.01	0.8265; 0.9723; −0.07	0.5729; 0.9785; −0.06	2.5355; 0.7836; −0.04
rating prior centered at −0.10	0.2852; 0.7515; −0.09	0.8167; 0.9734; −0.09	0.5509; 0.9751; −0.14	2.6575; 0.8014; 0.00
rating prior on today's entries too	0.3263; 0.7845; +0.01	0.7943; 0.9719; −0.15	0.5603; 0.9780; −0.10	2.3400; 0.7763; −0.11
survey sd 0.04 on β_1	0.3223; 0.7792; 0.00	0.8565; 0.9790; 0.00	0.5894; 0.9825; 0.00	2.6709; 0.8029; 0.00
survey sd 0.16 on β_1	0.3222; 0.7791; 0.00	0.8564; 0.9790; 0.00	0.5893; 0.9824; 0.00	2.6709; 0.8029; 0.00

The priors move the units differently, and the rows say why. A rating prior three times tighter pulls the commissioning entries up toward the rating; that raises the commissioned model's conductance and so every excess, and it moves unit 2's most, +0.34 standard deviations, because $g_{2,0}$ lies farther from the rating than $g_{1,0}$. A rating prior centered at −0.10 pulls them down and lowers every excess, by 0.09 to 0.14 standard deviations, except the duty shift, which depends on the commissioned split rather than the total. The rating prior put on today's entries too, the other choice of Section 5.1.4, pulls $g_{2,18}$ up from −0.1548 to −0.1464 and hides part of unit 2's loss: −0.15 standard deviations on its excess approach and −0.11 on the duty shift. It moves unit 1's excess the other way, +0.01: $g_{1,18}$ goes down from −0.0429 to −0.0445, through the anticorrelation of the two coils within today's window, while today's total rises from 12.7022 to 12.7417 kW/K. The survey's width on β_1 moves nothing, as the zero columns of the patterns (6.9) require.

6.4.4 The answers against the stated truth

The last check sets the answers against the stated truth of the record (Section 6.6.5). Table 6.7 collects the routes: each row is a route, each column one excess, with its value, its band in brackets and its probability of being positive where the route gives one. The true excess approaches are 0.5540 and 0.9246 K; the posterior lies 0.55 standard deviations below the first and 0.16 below the second. The true supply penalty is 0.7393 K, 0.54 standard deviations above the posterior's, and the true duty shift 1.8529 kW, 0.26 standard deviations below it. All four lie inside their bands. The split's terms lie inside theirs too: the season's true −1.1107 K (0.50

standard deviations above the posterior's), the load's 0.4476 K (0.29 below), and the true change of the cold aisle is +0.0761 K, against the posterior's −0.0427 K. The true heat to remove is 118.0000 and 134.8426 kW, the true duties today 69.9484 and 64.8942 kW, and the supply the true commissioned coils would need to hold the true cold aisle is 13.6786°C.

Table 6.7: The routes to the four excesses of window 18 and the stated truth. In each cell the value, the 5–95 % band in brackets and the probability of a positive excess, or the standard deviation where marked; values in K, and in kW for $\eta_{D,1}$.

Route	$\eta_{A,1}$ [K; –]	$\eta_{A,2}$ [K; –]	η_s [K; –]	$\eta_{D,1}$ [kW; –]
Laplace on the reduced state (reported)	0.3222 [−0.3668, 1.0113]; 0.7791	0.8564 [0.1635, 1.5493]; 0.9790	0.5893 [0.1293, 1.0493]; 0.9825	2.6709 [−2.4855, 7.8272]; 0.8028, P by the linear form (6.15)
Full posterior, reweighted draws	0.3221 [−0.3745, 1.0066]; 0.7778	0.8591 [0.1566, 1.5425]; 0.9775	0.5905 [0.1237, 1.0435]; 0.9809	2.6821 [−2.4844, 7.8636]; 0.8028
Full state, $\sigma_p = 10^{-3}$ Newton covariance	0.32223, sd 0.41893 sd 0.41872	0.85641, sd 0.42125 sd 0.42088	0.58932, sd 0.27965 sd 0.27962	2.67087, sd 3.13484 sd 3.13123
Readings taken as exact	0.3090	0.8158	0.5161	2.6527
Stated truth	0.5540	0.9246	0.7393	1.8529

6.5 The report (step 6)

The set-point tile's −0.05 K is not "nothing has changed". The season's −1.11 K cancels the IT load's +0.48 K and the coils' +0.59 K. At today's heat the commissioned coils would hold the same cold aisle with the water at 13.53°C instead of 12.94°C: the coils' loss spends 0.59 K of the supply (0.13 to 1.05 K), with probability 0.98. Unit 2 works harder in approach, not in duty: its approach is 0.86 K (0.16 to 1.55 K) above the commissioned model's, with probability 0.98, while it passes 2.7 kW less. Unit 1 carries the 2.7 kW (−2.5 to 7.8 kW, with probability 0.80, the same event as Chapter 5's "coil 2 has fallen more"), and its own excess approach, 0.32 K with probability 0.78, is not demonstrated. The even split's 15.1 % more per unit is the load, and it is blind to the shift.

6.5.1 The answers and their bands

Table 6.8 gives the answers in the order of Table 6.1. Each row is a question; the second column gives the reported value from the Laplace normal, the third the full posterior's value beside it, and the fourth the 5–95 % band.

6.5.2 The entry in the maintenance log

The assessment of window 18 is the second entry of the two-unit hall's maintenance log, after Chapter 5's. It uses nothing later than window 18, and in the values of Table 6.8 it reads:

> *Window 18 (Jun Y2), effort at the set point against the commissioned model at today's load and supply.*
> Cold aisle 16.92°C, as at handover (−0.05 K): held by the winter supply, not a condition reading.
> Supply penalty 0.59 K (5–95 % band 0.13 to 1.05 K), $P = 0.98$: the chilled water runs 0.59 K colder than the commissioned coils need.
> Unit 2: approach 0.86 K above the commissioned model (0.16 to 1.55 K), $P = 0.98$; duty 2.7 kW below it.

Table 6.8: The report for the two-unit hall at window 18 (Jun Y2): the effort behind the held set point, against the commissioned model at today's conditions. Reported values from the Laplace normal of the joint state; bands are 5–95 %.

Asset manager's question	Reported value	Full posterior	Band
The unit still meets its set-point; is it working harder to do so?	unit 2: $\eta_{A,2} = 0.8564$ [K], probability 0.979 [–]; unit 1: $\eta_{A,1} = 0.3222$ [K], probability 0.779 [–]	0.8591 [K], 0.978 [–]; 0.3221 [K], 0.778 [–]	unit 2: 0.1635 to 1.5493 [K]; unit 1: −0.3668 to 1.0113 [K]
How much colder must the chilled water run, because of the coils?	$\eta_s = 0.5893$ [K], probability 0.982 [–]; supply the commissioned coils need 13.5287 [°C]	0.5905 [K], 0.981 [–]	0.1293 to 1.0493 [K]
Why does the cold aisle still read as it did at handover?	change −0.0427 [K] = season −1.1124 + IT load 0.4804 + coils 0.5893 [K]	change −0.0416 [K]; IT load 0.4791 [K]	change −0.4022 to 0.3168 [K]; IT load 0.2969 to 0.6639 [K]
Which unit carries the heat the fouled coil no longer passes?	unit 1: $\eta_{D,1} = 2.6709$ [kW], probability 0.803 [–]	2.6821 [kW], 0.803 [–]	−2.4855 to 7.8272 [kW]

Unit 1: approach 0.32 K above (−0.37 to 1.01 K), not demonstrated ($P = 0.78$); carries the 2.7 kW (−2.5 to 7.8 kW), $P = 0.80$.
Action: none on condition from this entry (no decision level is set). The supply penalty joins the health indices as a monitored quantity, and the set-point tile is removed from the condition review.
Next measurement: a reference discharge sensor on unit 2, read in this window (Jun Y2).

The entry records what the effort is measured against as well as its size: the commissioning fit's coils at today's load entry and supply. Whether a supply penalty of this size warrants cleaning is not settled by its probability; it needs the decision level that Chapter 19 calibrates against the costs of false and missed work orders, and the price of the chiller energy the penalty spends, which a later part of the book carries.

Common mistake

Reading a held set point as condition. A controlled variable at its set point is the sum of everything that moves it, the season, the load and the coils; the effort of a fouled coil shows in the manipulated variable and in the duties, read against the commissioned model at today's conditions.

6.6 The next measurement (step 7)

The report leaves unit 1's excess approach not demonstrated and the duty shift's band running from −2.4855 to 7.8272 kW. A further measurement in today's window is one more row of the residual vector, and what it removes from each band is known before its value is.

6.6.1 The candidates

The preposterior is equation (5.22) with the candidate rows of Section 5.6.1 in today's window, applied to the four excesses: for a candidate a with Jacobian lines $\mathbf{J}_a$, $\boldsymbol{\Sigma}_a = (\boldsymbol{\Sigma}^{-1} + \mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a)^{-1}$ and $\mathrm{sd}_a(\eta) = \sqrt{\mathbf{g}_\eta^{\mathsf{T}}\boldsymbol{\Sigma}_a\,\mathbf{g}_\eta}$, and the width it takes off is $1 - \mathrm{sd}_a/\mathrm{sd}$. The candidates are a water-side heat meter on either coil, of standard deviation $\sigma_W = 3.0$ kW, which reads that unit's duty $D_{u,18}$; meters on both coils, the permanent instrument at the place of the question; a reference discharge sensor of 0.1 K on either unit; a second logged hour of today's five readings, a new

load entry with five fresh rows on the same coils, formed on the eight-entry state; and a second return-air sensor or load meter, which repeats a row already present. Two limits that are not measurements close the list: today's window known exactly, and the commissioning entries known exactly, which is not open today.

Table 6.9 ranks the candidates. Each row is a candidate; the columns give the share of the present standard deviation of each answer that it removes. The shares for unit 2's excess approach and for the supply penalty are set out as bars in Figure 6.6, two per candidate, ranked by unit 2's, with the limit of today's window known exactly as dashed lines.

Table 6.9: Candidate measurements in today's window, by the preposterior (5.22): the width each takes off the present standard deviations 0.41893 and 0.42125 K ($\eta_{A,1}$, $\eta_{A,2}$), 0.27965 K (η_s) and 3.13484 kW ($\eta_{D,1}$).

Candidate (today's window)	$\eta_{A,1}$ [%]	$\eta_{A,2}$ [%]	η_s [%]	$\eta_{D,1}$ [%]
Water-side heat meter on coil 1, $\sigma_W = 3.0$ kW	6.9	1.6	1.0	6.8
Water-side heat meter on coil 2, $\sigma_W = 3.0$ kW	1.7	6.9	1.0	7.0
Heat meters on both coils (the permanent instrument)	7.7	7.7	2.3	12.2
Reference discharge sensor on unit 1, 0.1 K	19.8	1.1	5.9	13.3
Reference discharge sensor on unit 2, 0.1 K	1.1	19.7	5.9	13.4
A second logged hour of today's five readings	12.0	12.0	11.7	12.2
A second return-air sensor, 0.3 K	0.3	0.3	0.7	0.0
A second load meter, 3.0 kW	1.5	1.6	3.6	0.0
Today's window known exactly (the floor)	24.6	24.7	22.6	26.4
The commissioning entries known exactly (not open today)	31.4	31.4	30.2	32.3

A reference discharge sensor acts on its own coil's column and takes 19.8 % off unit 1's excess approach or 19.7 % off unit 2's, 5.9 % off the supply penalty and 13.3 to 13.4 % off the duty shift. A second logged hour takes 11.7 to 12.2 % off all four excesses. A heat meter takes 6.8 to 7.0 % off the duty shift, and meters on both coils 12.2 %, but they take only 1.0 to 2.3 % off the supply penalty; a second return-air sensor or load meter removes almost nothing, because the rows already present fix what it reads.

6.6.2 The floor and the duty meter

No reading today takes more than 22.6 to 26.4 % off any answer: that is the width removed if today's window were known exactly, and every candidate is capped there, because every excess sets today against the commissioning test. The commissioning entries known exactly would take 30.2 to 32.3 % off. That test cannot be repeated, and its share of every band is carried, not improved.

The duty meter is read by its row. The meter on coil 1 reads $D_{1,18} = G_{1,18}(T_{r,18} - T_s(18))$, and its gradient on the joint state has the entries 0.5279 on Q_{18}, +33.9052 on $g_{1,18}$ and −33.9052 on $g_{2,18}$, in kW per unit entry, and none on the commissioning entries. It reads today's duty, whose posterior standard deviation is 2.5349 kW, but not the commissioned model's, whose standard deviation is 2.7050 kW and whose correlation with today's is 0.2855; the duty shift is the difference of the two. So meters on both coils take only 12.2 % off the duty shift and 2.3 % off the supply penalty. A permanent meter's value lies in the later windows it will read, and in a commissioning test that had one, which is Chapter 23's question. Under an active discharge controller it would be the instrument that still reads where the heat goes, and beside a water-flow or valve-position reading it would keep the coil observable.

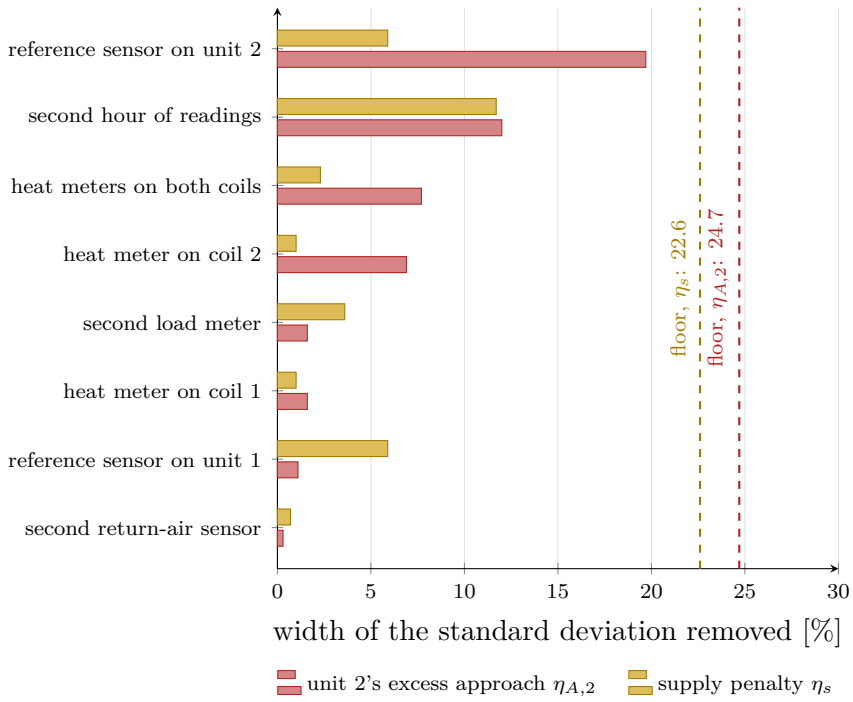


Figure 6.6: The candidate measurements of Table 6.9, ranked by the width they remove from unit 2's excess approach, with the supply penalty's beside each, by the preposterior (5.22). Dashed: the width removed if today's window were known exactly, the most any measurement in today's window can remove.

6.6.3 The recommendation

For the lead question the next measurement is a reference discharge sensor on unit 2 in this window, which takes 19.7 % off unit 2's excess approach and 13.4 % off the duty shift; the matching sensor on unit 1 takes 19.8 % off unit 1's. A second logged hour of today's readings costs nothing and takes 11.7 to 12.2 % off all four excesses, provided its errors are independent of the first hour's; readings whose errors are correlated from one hour to the next are Chapter 24's question. Heat meters are not worth requesting for today's answers: they read today's duties, not the commissioned model's, and their case rests on the windows still to come.

6.6.4 What the analysis showed

The dashboard said the set point was met and that the load explained the rest: the cold aisle within 0.05 K of handover and each unit 15.1 % busier. Read against the commissioned model at today's conditions, the same window shows a supply penalty of 0.59 K that the winter supply hides, an approach on unit 2 0.86 K wider than its commissioned coil would need, and a duty shift of 2.7 kW onto unit 1, the unit whose own loss is not demonstrated. The physics with the readings taken as exact gives nearly the same values, 0.52, 0.31 and 0.82 K and 2.65 kW, so it is the physics, the commissioned coils carried to today's load and water, that finds the effort; the probability adds the bands and says which excesses are demonstrated. The supply penalty and unit 2's approach are; unit 1's approach is not; the duty shift is likely and not settled. A held set point is a balance of the season, the load and the coils, and the effort of a fouled coil spreads to the water and to the other unit. The commissioning test is again more than half of every band, and no measurement taken today removes that half.

6.6.5 Limits of the model

A constructed record. The condition record is generated from a stated truth, with the stated sensor errors, and step 5 compares the answers with that truth (Section 6.4.4). It is not field data, and the agreement is a check of the method on its own model, not of the model against a real hall.

Steady state, pinned airflow, no discharge controller. The rows hold within a window only once the hall has settled, and the discharge readings inform the coils only while nothing holds the discharge at a set point. Under an active discharge controller each unit's valve would throttle the water to hold its discharge, the discharge would read its set point and the return air the set point plus the air's rise, whatever the coils; the coil's loss would then show only in the valve position and the water flow. A record with that controller needs a row on one of them, and this record has neither.

The scenario node's coils. The node carries the commissioned conductances at a heat higher than the commissioning test's, which holds only if a coil's conductance does not depend on its load. A conductance that rose with the water flow or the air-side temperatures would make part of each excess an artifact of the load.

The coil row and the known inputs. The duty is the heat the coil takes from the air, with no losses in the unit's piping. The supply and the other heat are known inputs: a supply sensor out of calibration would go into the supply penalty, and into the log-derates with it.

No chiller model. The penalty is stated in K, not in kWh or reais. Its cost needs the chiller's efficiency against its supply temperature, which the plant does not carry, and the expected consequence of a later part of the book.

Two windows. Two windows say nothing about how fast the effort grew; the degradation rate is Chapter 10's question, and the season across the whole record, which a single pair of windows splits only for itself, is Chapter 13's.

The next chapter. A discharge reading that has moved since handover is either the coil or its sensor; Chapter 7 tells them apart.

6.7 What to retain

Definition 6.1 (Hidden degradation as an excess over the commissioned model)**.** Hidden degradation is the excess $\eta = f(\text{today}) - f(\text{scenario node})$ of a performance quantity f (an approach, a duty, the supply a set point needs) over its value at a scenario node: a copy of today's physical rows that carries the commissioned parameters at today's load entry and inputs, with no reading and no prior. The node adds no row, so the excess is read from the posterior that carries both windows, with $\mathrm{Var}(\eta \mid \mathbf{y}) = \mathbf{g}_\eta^\mathsf{T} \boldsymbol{\Sigma}\, \mathbf{g}_\eta$: each window's block and the cross term between them. For the two-unit hall the duty shifts sum to zero, $\eta_{D,1} + \eta_{D,2} = 0$, each excess approach is the supply penalty less the duty shift over the capacity rate, $\eta_{A,u} = \eta_s - \eta_{D,u}/C_u$, and the supply penalty η_s does not depend on the set point.

Three conclusions carry beyond the hall. A reading held at its set point is not a condition reading: it is a sum of the season, the load and the asset, and the asset's share of the effort is read against the commissioned model at today's conditions, as an excess with a band. The commissioned model is a scenario node of the same graph, sharing today's load entry so that

the load's error cancels, and the excess is a derived quantity of the one posterior, never a second fit. And the conservation rows say where the effort goes: a duty one unit sheds the other carries, and a coil's loss widens the approach of the unit beside it through the air they share.

Table 6.10 lists the terms of the chapter with their expressions, units and the subsection that defines each.

Table 6.10: The terms of hidden degradation and where each is defined. Units for the two-unit hall.

Term	Expression	Units	Meaning	Defined in
Hidden degradation	excess over the scenario node, eq. (6.4)	K, kW	the effort a held set point hides	6.1.1
Duty	$D_{u,t} = G_{u,t}(T_{r,t} - T_s(t))$, eq. (6.1)	kW	the heat a coil passes to the water	6.1.2
Scenario node	eqs. (6.2), (6.3)	–	today's rows with the commissioned coils; no reading	6.1.2
Excess approach	$\eta_{A,u} = A_{u,18} - \check{A}_u$	K	how much harder a unit works	6.1.2
Supply penalty	$\eta_s = T_{r,18} - \check{T}_r$, eq. (6.7)	K	the supply the coils' loss costs; the set point cancels	6.1.2
Duty shift	$\eta_{D,u} = D_{u,18} - \check{D}_u$, zero sum	kW	the heat moved between the units	6.1.2
Air-side identity	$\eta_{A,u} = \eta_s - \eta_{D,u}/C_u$, eq. (6.5)	K	the hall's share and the unit's own	6.1.2
Split of a held reading	eq. (6.8)	K	season, load and coils	6.1.3
Variance budget of an excess	eq. (6.12)	–	two window blocks and their cross term	6.3.1
Floor of an excess	today's window known exactly	–	the commissioning test's share of every band	6.6.2

6.8 Exercises

Exercise 6.1. *Formulation.* Write the question "has unit 2's approach risen more than 0.5 K beyond the commissioned model?" in the form of Table 6.1 and evaluate it from the fit of this chapter; do the same for unit 1. (A prediction under a threshold, equation (2.12) with $f = \eta_{A,2}$ and $f^* = 0.5\,\mathrm{K}$: $\Pr(\eta_{A,2} > 0.5 \mid \mathbf{y}) = \Phi\big((0.8564 - 0.5)/0.4213\big) = 0.8012$; for unit 1, $\Phi\big((0.3222 - 0.5)/0.4189\big) = 0.3357$.)

Exercise 6.2. *Derivation.* Derive the air-side identity $\eta_{A,u} = \eta_s - \eta_{D,u}/C_u$ and evaluate it for unit 2. (By equation (6.1), $T_{d,u} = T_r - D_u/C_u$, so $A_u = T_{d,u} - T_s = T_r - T_s - D_u/C_u$ in window 18 and at the node, at the same supply $T_s(18)$; their difference is $(T_{r,18} - \check{T}_r) - (D_{u,18} - \check{D}_u)/C_u = \eta_s - \eta_{D,u}/C_u$. For unit 2, $0.5893 - (-2.6709)/10.0 = 0.8564\,\mathrm{K}$.)

Exercise 6.3. *Derivation.* Show that $T_r - T_c$ does not depend on the coils, so that the set point cancels in the supply penalty, and compute the supply the commissioned coils need to hold today's cold aisle. (The two air sides summed with the mixing row and the hall balance give the air balance (5.4), $(C_1 + C_2)(T_r - T_c) = Q + Q_o$, which holds no conductance; at window 18, $(Q_{18} + Q_o(18))/(C_1 + C_2) = 6.8023\,\mathrm{K}$ both today and at the node. Hence $T_c = T_s + (Q + Q_o)/(G_1 + G_2) - 6.8023\,\mathrm{K}$, and the supply for any T_c^{sp} differs between the two sets of coils by η_s, equation (6.7): the commissioned coils need 13.5287°C.)

Exercise 6.4. *Application.* Show that $\eta_{D,1} > 0$ is the event $\ln h_{1,18} > \ln h_{2,18}$, and compare its probability through the gradient of $\eta_{D,1}$, through the linear form, and by the draws. (By

equation (6.14) the sign of $\eta_{D,1}$ is that of the change of unit 1's share, and the share $1/(1+\mathrm{e}^{g_2-g_1})$ falls as g_2-g_1 rises, so $\eta_{D,1}>0$ exactly when $g_{2,18}-g_{1,18}<g_{2,0}-g_{1,0}$, equation (6.15). Through the gradient of $\eta_{D,1}$ the probability is 0.8029; through Chapter 5's linear form, equation (5.17), 0.8028; the full posterior gives 0.8028 for both.)

Exercise 6.5. *Application.* A water-side heat meter on unit 1 reads 74.00 kW (standard deviation 3.0 kW) in today's window, against the fitted 71.8226 kW, whose posterior standard deviation is 2.5349 kW. Update the duty shift by the correction form and by a refit. (The row's gradient is that of $D_{1,18}$, (0, 0, 0, 0.5279, 33.9052, −33.9052, 0); by the rank-one update (4.66) with that gradient and the innovation $74.00-71.8226$, the gain on $\eta_{D,1}$ is 0.28965 per kW, and $\eta_{D,1}$ moves from 2.6709 to 3.3015 kW, its standard deviation from 3.1348 to 2.9211 kW, and its probability of being positive, by its gradient, from 0.8029 to 0.8708. The refit with the row gives $\eta_{D,1}=3.3015$ kW, standard deviation 2.9284 kW and probability 0.8702; the supply penalty moves to 0.5679 K.)

Chapter 7

Distinguishing Degradation from Sensor Drift

7.1 A rise in the approach with two explanations

The hall is two years and a month past its handover. In June of year 2, window 18, the condition assessment of Chapter 5 found coil 1 at a health index of 0.98 against its commissioning test, with no loss demonstrated. Today is January of year 3, window 25, summer in the southern hemisphere where the hall stands. The condition-monitoring dashboard reads unit 1's approach, its discharge air less the chilled-water supply, at 4.36 K, against 2.86 K at commissioning and 3.60 K in June: a rise of 1.50 K since commissioning, or 52.5 %. The health-index tile, the commissioning approach over today's, puts coil 1 at 0.66, "lost 34.4 %", now worse than coil 2's 0.70, "lost 30.0 %", the reverse of June's order. The physics gives the same reversal. Taken as exact, today's readings and the commissioning readings put coil 1 at 0.903 and coil 2 at 0.933, and the fit of Chapter 5's drawing on today's window puts coil 1 at 0.915 (0.840 to 0.996), with probability 0.957 that it has lost conductance. The maintenance planner reads both as fouling and wants a cleaning work order for coil 1.

The controls technician points out that unit 1's discharge sensor has not been checked since handover. A sensor that reads high raises the approach as a fouled coil does: a discharge reading 1 K high moves coil 1's readings-as-exact index by -0.1344, to 0.7686. The dashboard cannot say which of the two raised the approach, or whether both did. Nor can it say how much of the 1.50 K is either, when the load, the season and coil 2 move the approach too. The question of this chapter is

whether the rise of a coil's approach is the coil or its sensor, read as the evidence of drawings of the plant that free a derate on the coil, a bias on the sensor, both or neither; how much of the rise each drawing assigns to the coil; and which check would settle it.

It is answered with the method of Chapter 5, on the same joint state. What changes is the computation: the answer is a ranking of four drawings by their evidence, equations (4.18) and (4.20), not a band inside one drawing.

Jargon: sensor drift, candidate family, nested drawings, Occam factor, model averaging, bad-data detection, leverage, redundancy number, normalized residual, Savage–Dickey ratio, side-by-side calibration check

Sensor drift is the change of a sensor's offset since it was last calibrated, here since the commissioning test; a standing offset is the part already present at calibration. A *candidate family* is the set of drawings of the plant that free or pin each entry a question names. Drawings are *nested* when each is the largest of them with some of its entries pinned at fixed values. The *Occam factor* is the part of a log Bayes factor that charges a drawing for an entry it frees: the gain in fit less the log Bayes factor. *Model averaging* reports a quantity over the family as the mixture of its posteriors under the drawings, weighted by their posterior probabilities. *Bad-data detection* (the residual test) asks whether the rows of a window misfit, at the fit, by more than their errors allow. A row's *leverage* is the share of its own reading that its fitted value carries, and its *redundancy number* the rest, the variance of its residual at the fit; its *normalized residual* is the residual over the square root of the redundancy number. The *Savage–Dickey ratio* gives the Bayes factor of a nested pin as the posterior density of the freed entry at the pinned value over its prior density there. A *side-by-side calibration check* reads a reference thermometer beside the logged sensor, at the sensor's own position, so that their difference reads the sensor's offset and nothing else.

Key idea

A rise in a coil's approach that its own discharge sensor reports is explained equally by fouling of the coil and by a drift of the sensor, because both move that reading. The two drawings part only in what they predict for the readings that do not pass through that sensor: the return air, which sees the hall's total conductance, the other unit's discharge and the load. The Bayes factor is set by those readings. When they cannot separate the drawings, the answer is the drawings' probabilities, the drawing in which both act, and the check that would separate them.

7.1.1 A rise that two drawings explain

Chapter 5 read the health index with every sensor taken as unbiased: each reading row compared the fitted temperature with the reading and nothing else, equation (5.7). A drift breaks that assumption in one window and not the other. Write unit 1's discharge reading in window t as the air's temperature plus the sensor's offset o_t plus its random error. An offset that does not change, $o_t = o_0$, enters both windows' rows. When both windows run at the same load and the same supply, and their entries carry the same priors, the commissioning entry and today's move with it by the same amount, and the health index, their difference, cancels it: Chapter 3 shows this for a commissioning test and a later condition test read by one sensor (Section 3.5.7). Here neither condition holds. The rating prior holds the commissioning entries and not today's, so even at the same inputs a small part of a standing offset stays in the index; and the load and the season have moved, so the slopes of the two windows' discharge rows differ. The drift, the change $o_t - o_0$ since commissioning, does not cancel at all: it enters today's row alone, and a fit that takes the sensor as unbiased reads it as a change of the coil. Step 2 gives the sizes of the two.

So the question is a structure question: which drawing of the plant produced the record, one in which coil 1 changed, one in which its sensor drifted, one with both, or one with neither. Its answer is a ratio of evidences, the hypothesis comparison of Chapter 4, run over a family of candidates as a diagnosis, Chapter 2.

Principle 7.1. *A sensor drift and a loss of conductance are told apart by the readings that see one of them and not the other, never by the reading they share. The evidence of the drawings measures how far those readings tell them apart.*

7.1.2 The two-unit hall with a suspect sensor (step 0)

The plant is the two-unit hall of Chapter 5, Section 5.1.2, cited here and not derived again: in each window the five physical rows (5.2) fix the return air, the two discharge temperatures, the cold aisle and rack 1's inlet, and the forward equations (5.3) give them at that window's known inputs. Racks 2 and 3 are not drawn, and Chapter 5's check that their block separates carries over. The modeling choices of the graph are that chapter's paragraph: steady state within the window, the airflow pinned at the units' fixed fan speed with nothing controlling the discharge air, the chilled-water supply and the other heat as known inputs, and one effective conductance per coil. The choices this chapter adds are named at the rows, in step 2.

The record. Table 7.1 gives the two windows the chapter reads, in the layout of Table 5.1: the two known inputs of each window and its five readings, with the readings' standard deviations in the last column. January is summer, so both known inputs differ from commissioning: the supply is 1.45 K warmer and the other heat 2.90 kW higher, and the meter reads a load 23.36 kW higher. By the meter the heat to remove has risen from 118.92 to 145.18 kW, a ratio of 1.2208. Load and season therefore enter the 1.50 K before any coil or sensor does.

Table 7.1: The two windows of the condition record read in this chapter: the commissioning test and today.

Quantity [unit]	Window 0 (commissioning)	Window 25 (Jan Y3)	sd [K or kW]
Chilled-water supply $T_s(t)$, known input [°C]	14.00	15.45	–
Other heat $Q_o(t)$, known input [kW]	8.00	10.90	–
Return air y_{T_r} [°C]	22.71	26.53	0.3
Discharge air of unit 1, $y_{T_{d,1}}$ [°C]	16.86	19.81	0.3
Discharge air of unit 2, $y_{T_{d,2}}$ [°C]	17.08	19.85	0.3
Hall load meter y_Q [kW]	110.92	134.28	3.0
Rack-1 inlet y_{T_1} [°C]	18.08	21.58	0.3

What the question adds to the graph. Two places of the graph are under suspicion, and each is given an entry that a drawing may free or pin. Coil 1's node of window 25 carries its change since commissioning, $\ln h_{1,25} = g_{1,25} - g_{1,0}$, the logarithm of its health index (5.1); pinned at zero, the coil is unchanged. The dashed reading edge from unit 1's discharge air to its sensor in window 25 carries a bias node $b_{d,1}$, the drift of that sensor since commissioning, an instance of the bias b_j of a reading; pinned at zero, the sensor is sound. The family of drawings is generated at the place of the question, one coil node and one sensor edge, each freed or pinned, which gives the four drawings of Table 7.2. Each row is a drawing; the columns give the state of the derate term $\ln h_{1,25}$ and of the bias term $b_{d,1}$ in it, with the prior a freed term carries, and what the drawing says of the plant.

Each of M_{neither}, M_{foul} and M_{drift} is M_{both} with one or two of its entries pinned at zero: the family is nested, and M_{both} contains the other three.

The graph. The rows of the two windows are set side by side in Figure 7.1, window 0 on the left and window 25 on the right, in the layout of Figure 5.1. In each copy the IT load and the

Table 7.2: The candidate family at the place of the question: four nested drawings of the two-unit hall at window 25. Coil 2's change and both loads are free in every drawing.

Drawing [–]	Derate term $\ln h_{1,25}$ [–]	Bias term $b_{d,1}$ [K]	What it says [–]
M_{neither}	pinned at 0	pinned at 0	coil 1 unchanged, sensor sound
M_{foul}	free, $\mathcal{N}(0,\ 0.10^2)$	pinned at 0	coil 1 fouled, sensor sound
M_{drift}	pinned at 0	free, $\mathcal{N}(0,\ 1.0^2)$	coil 1 unchanged, sensor drifted
M_{both}	free, $\mathcal{N}(0,\ 0.10^2)$	free, $\mathcal{N}(0,\ 1.0^2)$	both

other heat feed the return air, the return air feeds the two coils, which pass heat to the pinned chilled water and send their discharge air to the cold aisle, and the cold aisle feeds rack 1's inlet; dashed arrows leave each copy for its five readings, with their values. In window 25 coil 1's node carries $\ln h_{1,25}$ and is labeled with the drawings that free it, and unit 1's dashed discharge edge passes through the bias node $b_{d,1}$, labeled the same way. The shared recirculation fraction β_1 sits below the two copies, linked to both rack-1 inlets, with its survey prior; the rating prior is attached to window 0's coils only.

The joint state. The largest drawing, M_{both}, carries eight entries,

$$\mathbf{z} = \left(Q_0,\ g_{1,0},\ g_{2,0},\ Q_{25},\ \ln h_{1,25},\ \ln h_{2,25},\ \beta_1,\ b_{d,1}\right)^{\mathsf{T}}, \tag{7.1}$$

z_1 to z_8 in that order, the loads in kW, the bias in K and the others dimensionless. Today's coils are carried by their changes since commissioning, so that $g_{u,25} = g_{u,0} + \ln h_{u,25}$ in the forward equations of window 25. This is a linear change of coordinates of Chapter 5's state (5.5), which leaves the posterior unchanged, and it makes "coil 1 unchanged" the pin $\ln h_{1,25} = 0$ of one entry. M_{foul} and M_{drift} drop one entry each, $b_{d,1}$ and $\ln h_{1,25}$, and M_{neither} drops both.

7.1.3 The asset manager's questions (step 1)

The asset manager asks five questions of the two windows, and step 1 writes each over the drawings and their states, with the unknown on the left of the bar and the ten readings and the priors on the right. Table 7.3 lists them, in the columns of Table 5.2: the question in the asset manager's words, the drill question first; the general form of its category from Chapter 2; its mathematical form for the hall; its category; the computation of Section 2.3.6 that answers it; and the subsection where it is answered. The first question is the book's first whose computation is the evidence. The questions are about coil 1 and its sensor; coil 2's change is free in every drawing and is reported in one line. Which check would settle the comparison asks nothing new of today's evidence; it is the next measurement for these five questions, Section 7.6.

7.1.4 The evidence as rows and the drawings (step 2)

Step 2 attaches the evidence of Table 7.1 to the graph as rows on the joint state (7.1). Each window contributes Chapter 5's five reading rows, equation (5.7), with standard deviations 0.3 K for the four temperatures and 3.0 kW for the load meter, with one change: the discharge row of unit 1 in window 25 adds the drift to the air's temperature,

$$r_{d,1,25}(\mathbf{z}) = \left(T_{d,1}(\mathbf{z}_{25}) + b_{d,1} - y_{T_{d,1},25}\right)/0.3, \tag{7.2}$$

with $\mathbf{z}_{25} = (Q_{25},\ g_{1,0} + \ln h_{1,25},\ g_{2,0} + \ln h_{2,25},\ \beta_1)$ the quantities window 25's forward equations take. The priors are those of Chapter 5, the rating prior $g_{u,0} \sim \mathcal{N}(0,\ 0.15^2)$ on each commissioning entry and the survey prior $\beta_1 \sim \mathcal{N}(0.20,\ 0.08^2)$, and two more in the drawings that free

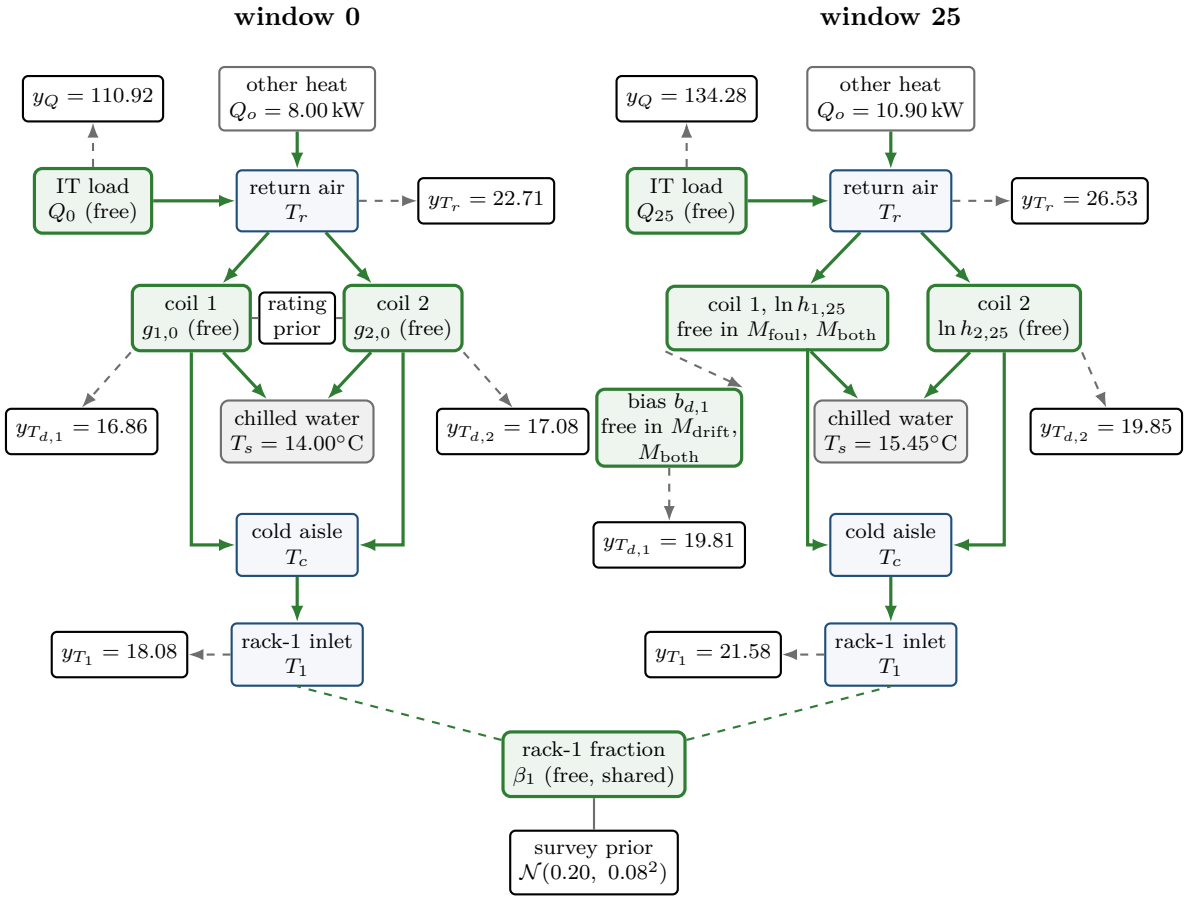


Figure 7.1: The rows of the two-unit hall in window 0 (left) and window 25 (right), with the two suspects of the question. Free (green) in every drawing: each window's IT load Q_t, the commissioned log-derates $g_{1,0}$, $g_{2,0}$, coil 2's change $\ln h_{2,25}$, and the recirculation fraction β_1 of rack 1, one entry shared by both windows (dashed green links). Free in some drawings (green, labeled): coil 1's change $\ln h_{1,25}$, free in M_{foul} and M_{both} and pinned at zero in M_{neither} and M_{drift}; the drift $b_{d,1}$ of unit 1's discharge sensor, on its reading edge of window 25, free in M_{drift} and M_{both} and pinned at zero in the others. Blue: temperatures fixed by the physical rows (5.2). Known inputs of each window, pinned: the chilled-water supply $T_s(t)$ (gray) and the other heat $Q_o(t)$ (white); also pinned, not drawn, the ratings $\bar{G}_u = 7.0\,\mathrm{kW/K}$ and the capacity rates $C_u = 10.0\,\mathrm{kW/K}$. Observed (dashed): the return air, the two discharge temperatures and rack 1's inlet in °C, and the load meter in kW. Priors (solid gray): the rating prior on window 0's coils, and the survey prior on β_1.

Table 7.3: The questions asked of the two-unit hall at window 25 against its commissioning test. $\mathbf{y}$ is the ten readings and the prior rows of the drawing; $M \in \{M_{\text{neither}}, M_{\text{foul}}, M_{\text{drift}}, M_{\text{both}}\}$ the drawings of Table 7.2; $A_{1,t}$ unit 1's approach in window t; r_k^{N} the normalized residual of row k.

Asset manager's question	General form	Mathematical form	Category	Computation	Answered in
The approach temperature has risen 1.5 K; is the exchanger fouled or is the sensor drifting?	$B_{AB} = p(\mathbf{y} \mid M_A)/p(\mathbf{y} \mid M_B)$ and the family's ranking, (2.14), (2.15)	$B(M_{\text{foul}} : M_{\text{drift}})$; $\Pr(M \mid \mathbf{y})$ over the four drawings	hypothesis comparison	evidence	7.3.1
How much of the 1.5 K is in question at all, once the load, the season and the other coil are taken out?	moments of a derived quantity, (2.10)	$A_{1,25} - A_{1,0}$ split into heat, coil 2, coil 1, sensor and reading errors, under each drawing	estimation	posterior	7.3.2
Could it be both, and how much of each?	moments; credible interval, (2.10), (2.11)	$p(\ln h_{1,25}, b_{d,1} \mid \mathbf{y}, M_{\text{both}})$: medians, 5–95 % bands, correlation	estimation	posterior	7.3.3
What is coil 1's health index once the drift is allowed for, and has it lost conductance?	median and credible interval, (2.11); the loss as its threshold reading, (2.12)	$p(h_{1,25} \mid \mathbf{y}, M_{\text{both}})$: median and band; its threshold reading $\Pr(\ln h_{1,25} < 0 \mid \mathbf{y}, M_{\text{both}})$; the model average beside it	estimation	posterior	7.3.4
Would the readings themselves have flagged a faulty sensor?	$\Pr(f(\mathbf{z}) \geq f^* \mid \mathbf{y})$ on the residual statistic, (2.12)	$\Pr(\chi^2_6 > 2\mathcal{L})$ under Chapter 5's drawing; the largest $\lvert r_k^{\text{N}} \rvert$	prediction under a threshold	posterior	7.3.5

the terms they act on: the derate-term prior row $\ln h_{1,25}/\sigma_{\ln h}$, with $\sigma_{\ln h} = 0.10$, and the bias prior row $b_{d,1}/\sigma_b$, with $\sigma_b = 1.0\,\mathrm{K}$. The loads and $\ln h_{2,25}$ have no prior row in any drawing. Their flat priors put the same constant into every evidence, so it cancels in every ratio. M_{both} has fifteen rows, M_{foul} and M_{drift} fourteen, and M_{neither} thirteen.

The one fit and its four changes. The chapters of a part read one fit of the condition record, and this chapter's is Chapter 5's joint state and thirteen rows, with four changes. The window: today is window 25 in place of window 18. The coordinates: today's coils are carried by their changes, equation (7.1); refitted in them, Chapter 5's window 18 gives $\ln h_{1,18} = -0.01870$ (standard deviation 0.05204), $\ln h_{2,18} = -0.09734$ (0.05540) and $2\mathcal{L} = 0.9849$, that chapter's numbers. A prior row on the derate term, in the drawings that free it: Chapter 5 put no prior row on today's entries, which is right for a posterior, but a Bayes factor between a drawing that frees an entry and one that pins it needs a proper prior on that entry, or the ratio is undefined. And the bias entry, in the drawings that free it. Chapter 5's own drawing, M_{foul} without the derate-term prior, reappears in step 4 as the drawing the asset manager ran.

Which row acts on which entry. On the reduced state a reading row sees the free entries through the row of $\mathbf{N}_{\mathrm{b}}$ of the temperature it reads, equation (4.47). Window 0's temperature rows depend on $(Q_0, g_{1,0}, g_{2,0})$, and rack 1's also on β_1. In the change coordinates window 25's temperatures depend on $g_{1,0} + \ln h_{1,25}$ and $g_{2,0} + \ln h_{2,25}$, so its temperature rows touch $(g_{1,0}, g_{2,0}, Q_{25}, \ln h_{1,25}, \ln h_{2,25})$, rack 1's also β_1, and unit 1's discharge row also $b_{d,1}$. With $*$ for an entry that can be nonzero and the columns in the order of equation (7.1), the 15×8 Jacobian of M_{both} has the pattern

$$\mathbf{J} = \begin{bmatrix} * & * & * & 0 & 0 & 0 & 0 & 0 \\ * & * & * & 0 & 0 & 0 & 0 & 0 \\ * & * & * & 0 & 0 & 0 & 0 & 0 \\ * & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ * & * & * & 0 & 0 & 0 & * & 0 \\ 0 & * & * & * & * & * & 0 & 0 \\ 0 & * & * & * & * & * & 0 & * \\ 0 & * & * & * & * & * & 0 & 0 \\ 0 & 0 & 0 & * & 0 & 0 & 0 & 0 \\ 0 & * & * & * & * & * & * & 0 \\ 0 & * & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & * & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & * & 0 \\ 0 & 0 & 0 & 0 & * & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & * \end{bmatrix} \begin{array}{c} \text{window 0: return air } y_{T_r} \\ \text{discharge } y_{T_{d,1}} \\ \text{discharge } y_{T_{d,2}} \\ \text{load meter } y_Q \\ \text{rack-1 inlet } y_{T_1} \\ \text{window 25: return air } y_{T_r} \\ \text{discharge } y_{T_{d,1}}\text{, with } b_{d,1} \\ \text{discharge } y_{T_{d,2}} \\ \text{load meter } y_Q \\ \text{rack-1 inlet } y_{T_1} \\ \text{rating prior on } g_{1,0} \\ \text{rating prior on } g_{2,0} \\ \text{survey prior on } \beta_1 \\ \text{derate-term prior on } \ln h_{1,25} \\ \text{bias prior on } b_{d,1} \end{array}, \tag{7.3}$$

$$\mathbf{g}_{\ln h_1}^{\mathsf{T}} = \mathbf{e}_5^{\mathsf{T}} = (0,\ 0,\ 0,\ 0,\ 1,\ 0,\ 0,\ 0), \qquad \mathbf{g}_b^{\mathsf{T}} = \mathbf{e}_8^{\mathsf{T}} = (0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 1).$$

The second line is not a row: it gives the gradients of the questions' two quantities, which in these coordinates are entries of the state. Two facts of the graph are read here, before any number. Window 25's discharge row of unit 1 has the same slope on $\ln h_{1,25}$ as on $g_{1,0}$, because it sees the two only through their sum, and its slope on $b_{d,1}$ is the constant $1/0.3$; the bias column has no other nonzero but its prior. That reading therefore cannot tell a lower $\ln h_{1,25}$ from a higher $b_{d,1}$: only the other rows of window 25, the return air, unit 2's discharge and rack 1's inlet, which touch $\ln h_{1,25}$ and not $b_{d,1}$, separate the two columns. And each nested drawing deletes a column and its prior row, so the four drawings share every other row.

Modeling choices at the rows. Five choices make these rows, each named once. The sensors are taken as calibrated at the commissioning test: the bias entry acts in window 25 only, so $b_{d,1}$ is the drift since commissioning, and it fails if a sensor was replaced or recalibrated between the windows, when the entry is the change across that event. A standing offset is then left out of the rows, and what it would do is small but not zero. Under Chapter 5's drawing, an offset of $+1\,\mathrm{K}$ on unit 1's discharge sensor in both windows moves $\ln h_{1,25}$ by -0.0118 when both windows run at window 0's inputs, because the rating prior holds the commissioning entry and not today's; without the rating prior it moves it by zero to the printed digits; and at the recorded windows 0 and 25 it moves it by $+0.0103$. A drift of $+1\,\mathrm{K}$ in window 25 alone moves it by -0.0995. One suspect sensor and one suspect coil, the place of the question: this keeps the family to four nested drawings, and it cannot explain a drift of the return-air sensor or of unit 2's, which the diagnosis over every sensor of *Operations and Control by Derivation*, Chapter 7, runs. The bias is constant within the window: one window gives one value of each reading, and a drift that grows from window to window would need trend rows on the bias, which this chapter does not write. The freed terms have proper priors, $\sigma_{\ln h} = 0.10$, a coil's change since handover within about 10 %, and $\sigma_b = 1.0\,\mathrm{K}$, a drift within about a kelvin: the evidence depends on these widths, and step 5 varies them, 0.05 to 0.20 and 0.5 to 2.0 K (Section 7.4.4). And the four drawings have equal prior probabilities, $\frac{1}{4}$ each: this is what lets the evidence alone rank them, and it fails where the asset manager's experience favors a sound sensor or a clean coil; Section 7.4.4 sets the alternative of 0.5 on M_{neither} beside it.

7.2 The posterior of the state (step 3)

All five questions of Table 7.3 are read from the posterior of the state under each of the four drawings, with the rows of step 2, and the first also from the drawings' evidences. Each drawing is fitted on its own rows, by the derivation of Chapter 5.

7.2.1 The rows and the normal equations

The rows of the reduced state are the evidence rows at the point of the set where the physical rows hold, equation (4.46), with each window's temperatures from equation (5.3) at that window's inputs. For M_{both} the residual vector stacks the two windows' reading rows and the five priors in the order of the pattern (7.3),

$$\mathbf{r}(\mathbf{z}) = \begin{bmatrix} \mathbf{r}_0(\mathbf{z}) \\ \mathbf{r}_{25}(\mathbf{z}) \\ g_{1,0}/0.15 \\ g_{2,0}/0.15 \\ (\beta_1 - 0.20)/0.08 \\ \ln h_{1,25}/\sigma_{\ln h} \\ b_{d,1}/\sigma_b \end{bmatrix}, \qquad \mathcal{L}(\mathbf{z}) = \tfrac{1}{2}\,\mathbf{r}(\mathbf{z})^{\mathsf{T}}\mathbf{r}(\mathbf{z}), \tag{7.4}$$

fifteen dimensionless rows on eight entries, with $\mathbf{r}_0$ window 0's five rows of equation (5.7) and $\mathbf{r}_{25}$ window 25's, whose unit-1 discharge row is equation (7.2). The other drawings drop the rows and columns of the terms they pin. The reduced Jacobian is $\mathbf{J} = \mathbf{J}_{\text{rows}}\mathbf{N}$ of equation (4.47), formed window by window as in Section 5.2.1; in window 25 a temperature row's slope on $\ln h_{u,25}$ equals its slope on $g_{u,0}$, by the change of coordinates, and the bias column is $1/0.3$ in the one row that carries it and $1/\sigma_b$ in its prior. Each Gauss–Newton iteration solves the normal equations (5.9) on the drawing's own rows, and at the minimum its Laplace covariance is $\mathbf{\Sigma} = (\mathbf{J}^{\mathsf{T}}\mathbf{J})^{-1}$, equation (5.10). With every row normalized the information matrix is $\mathbf{\Lambda} = \mathbf{J}^{\mathsf{T}}\mathbf{J}$; with unnormalized rows and their weights it is $\mathbf{J}^{\mathsf{T}}\mathbf{W}\mathbf{J}$.

The iterations. Each drawing starts from the state the record suggests: the meter's loads, the coils at their ratings and unchanged, the survey's fraction and no drift, for M_{both} $\mathbf{z}^{(0)} = (110.92,\ 0,\ 0,\ 134.28,\ 0,\ 0,\ 0.20,\ 0)^{\mathsf{T}}$. Table 7.4 runs the normal equations of M_{both} from it. Each row is an iteration, with the objective and the norm of its gradient $\mathbf{J}^{\mathsf{T}}\mathbf{r}$. The objective falls from 37.427034 to 1.638827 in the first step and to 1.608757 in the second, and the gradient is below 10^{-4} from the third. The other three drawings converge in six or seven iterations from the matching start.

Table 7.4: Gauss–Newton iterations of M_{both} on the joint state of windows 0 and 25, from the meter's loads, the rated and unchanged coils, the survey's fraction and no drift.

Iteration l [–]	Objective $\mathcal{L}(\mathbf{z}^{(l)})$ [–]	$\lVert\mathbf{J}^{\mathsf{T}}\mathbf{r}\rVert$ [–]
0	37.427034	6.428×10^{2}
1	1.638827	1.947×10^{1}
2	1.608757	3.250×10^{-2}
3	1.608757	$< 10^{-4}$
4	1.608757	$< 10^{-4}$
5	1.608757	$< 10^{-4}$
6	1.608757	$< 10^{-4}$

The Jacobian at the MAP state. Window 0's lines and the prior lines have the forms of equation (5.12). The five lines of window 25, with the columns of equation (7.1), are

$$\mathbf{J}_{25}(\hat{\mathbf{z}}) = \begin{bmatrix} 0 & -19.6184 & -18.1892 & 0.2621 & -19.6184 & -18.1892 & 0 & 0 \\ 0 & -31.6219 & -6.1857 & 0.0891 & -31.6219 & -6.1857 & 0 & 3.3333 \\ 0 & -7.6149 & -30.1927 & 0.1017 & -7.6149 & -30.1927 & 0 & 0 \\ 0 & 0 & 0 & 0.3333 & 0 & 0 & 0 & 0 \\ 0 & -19.6184 & -18.1892 & 0.1359 & -19.6184 & -18.1892 & 24.0413 & 0 \end{bmatrix} \begin{matrix} \text{return air} \\ \text{discharge 1} \\ \text{discharge 2,} \\ \text{load meter} \\ \text{rack-1 inlet} \end{matrix} \tag{7.5}$$

the load column in kW^{-1}, the bias column in K^{-1} and the others dimensionless. The columns of $g_{u,0}$ and $\ln h_{u,25}$ are equal in every line, as the change of coordinates makes them; the window-0 lines tell them apart. Unit 1's discharge line is -31.6219 on $\ln h_{1,25}$ and 3.3333 on $b_{d,1}$: per unit of the entry, -9.4866 K and 1 K of that reading. The return-air line is -19.6184 on $\ln h_{1,25}$ and zero on $b_{d,1}$, unit 2's discharge line -7.6149 and zero; these are the lines that separate the two terms. The ten reading rows have rank 8 of 8: the window resolves both terms even without their priors, though not by much.

7.2.2 The fits of the four drawings

Table 7.5 gives the four fits. Each row is a drawing; the columns give its numbers of rows and entries, twice its objective at the MAP state, and the three entries the questions read, each with its posterior standard deviation $\sqrt{\Sigma_{ii}}$, a dash where the drawing pins it.

Freeing either term lowers the misfit, from 6.3272 to 4.0287 with the derate and to 3.5821 with the drift, and freeing both lowers it only to 3.2175. The single-term drawings give the whole of the rise to their own term, M_{foul} a change of -0.07025 of coil 1 and M_{drift} a drift of 0.66038 K, while M_{both} shares it, -0.03579 and 0.46225 K, with wider spreads. Coil 2's change sits between -0.07072 and -0.10434 in every drawing. Chapter 5's own drawing, M_{foul} without the derate-term prior, has thirteen rows on seven entries; fitted on today's window it gives $\ln h_{1,25} = -0.08922$ (0.05193) and $2\mathcal{L} = 3.4019$; its fit is the one the asset manager ran.

The fitted temperatures. What the drawings disagree about, outside unit 1's discharge, is small. The fitted return air of window 25 is 26.9009°C under M_{foul} and 26.7356°C under

Table 7.5: The fits of the four drawings at window 25: the MAP values of the entries the questions read, with their posterior standard deviations in parentheses.

Drawing	Rows [–]	Entries [–]	$2\mathcal{L}(\hat{\mathbf{z}})$ [–]	$\ln h_{1,25}$ [–]	$b_{d,1}$ [K]	$\ln h_{2,25}$ [–]
M_{neither}	13	6	6.3272	–	–	−0.10434 (0.04630)
M_{foul}	14	7	4.0287	−0.07025 (0.04615)	–	−0.07072 (0.05156)
M_{drift}	14	7	3.5821	–	0.66038 (0.39779)	−0.10369 (0.04742)
M_{both}	15	8	3.2175	−0.03579 (0.05950)	0.46225 (0.51496)	−0.08631 (0.05526)

M_{drift}, less than a reading's standard deviation apart, and unit 2's discharge air 19.7789 against 19.9108°C. Unit 1's discharge air is 19.6517°C under M_{foul} and 19.0902°C under M_{drift}; under M_{drift} the sensor, reading the air plus its drift, is predicted to read 19.7506°C, against its 19.81. Both drawings fit unit 1's discharge reading; they differ in where its rise comes from.

The covariance of M_{both}. The other entries of M_{both}'s MAP state are $Q_0 = 110.39913\,\text{kW}$ (2.74530), $g_{1,0} = -0.02317$ (0.03816), $g_{2,0} = -0.04829$ (0.04015), $Q_{25} = 133.35015\,\text{kW}$ (2.83021) and $\beta_1 = 0.24289$ (0.03569). Its Laplace covariance (5.10) is written as $\boldsymbol{\Sigma} = \mathbf{D}\boldsymbol{\varrho}\mathbf{D}$, with $\mathbf{D}$ the diagonal of the posterior standard deviations and $\boldsymbol{\varrho}$ the correlation matrix, in the order of equation (7.1):

$$\mathbf{D} = \operatorname{diag}\big(2.74530,\ 0.03816,\ 0.04015,\ 2.83021,\ 0.05950,\ 0.05526,\ 0.03569,\ 0.51496\big),$$

$$\boldsymbol{\varrho} = \begin{bmatrix} 1 & 0.2870 & 0.3277 & 0.0205 & -0.1263 & -0.2636 & 0.0306 & 0.0507 \\ 0.2870 & 1 & -0.3947 & 0.0800 & -0.4077 & 0.1981 & 0.2103 & 0.2081 \\ 0.3277 & -0.3947 & 1 & -0.0144 & 0.2268 & -0.6912 & 0.1282 & -0.0181 \\ 0.0205 & 0.0800 & -0.0144 & 1 & 0.2685 & 0.1543 & 0.0946 & 0.2130 \\ -0.1263 & -0.4077 & 0.2268 & 0.2685 & 1 & -0.5204 & 0.2125 & 0.6414 \\ -0.2636 & 0.1981 & -0.6912 & 0.1543 & -0.5204 & 1 & -0.1099 & -0.3254 \\ 0.0306 & 0.2103 & 0.1282 & 0.0946 & 0.2125 & -0.1099 & 1 & 0.3335 \\ 0.0507 & 0.2081 & -0.0181 & 0.2130 & 0.6414 & -0.3254 & 0.3335 & 1 \end{bmatrix}, \tag{7.6}$$

the load entries of $\mathbf{D}$ in kW, the bias's in K and the others dimensionless. The block the questions read is $\Sigma_{55} = 3.5397 \times 10^{-3}$, $\Sigma_{58} = 1.9651 \times 10^{-2}\,\text{K}$ and $\Sigma_{88} = 2.6518 \times 10^{-1}\,\text{K}^2$. The correlations read by sign, and the rows fix each sign. $\ln h_{1,25}$ against $b_{d,1}$, +0.6414: unit 1's discharge reading fixes the combination $b_{d,1} - 9.4866 \ln h_{1,25}$, so a sensor that reads higher leaves less of the rise to the coil, and a larger drift goes with a smaller loss, a larger $\ln h_{1,25}$. This is the ridge of the two terms. $\ln h_{1,25}$ against $\ln h_{2,25}$, −0.5204: the return air sees the hall's total conductance, as in Chapter 5. $b_{d,1}$ against $\ln h_{2,25}$, −0.3254, through the same total: a larger drift goes with a larger $\ln h_{1,25}$, which the return air pays for with a smaller $\ln h_{2,25}$. $b_{d,1}$ against β_1, +0.3335, through the cold aisle that rack 1's inlet reads: a larger drift goes with a better coil 1 and a cooler cold aisle, which a larger fraction offsets at the same inlet reading. And $g_{1,0}$ against $\ln h_{1,25}$, −0.4077, with −0.6912 for coil 2: the change is a difference with the commissioned entry, so a higher commissioned value means a larger loss to the same value today.

7.2.3 The evidence of a drawing

One choice enters here: the evidence of each drawing is its Laplace evidence, the Gaussian integral of the objective about the drawing's own MAP state with its own curvature. This gives every evidence from the drawing's fit alone, and it is accurate only as far as each objective

is quadratic over its posterior's width, which the rows, nonlinear in the log-derates, make approximate; Section 7.4.1 sets importance sampling beside it.

The evidence of a drawing M is the density of the readings under it, averaged over its prior, equation (4.18). Its integrand is $\exp[-\mathcal{L}_M(\mathbf{z})]$ times the normalizing constants $(\sqrt{2\pi}\,\sigma_k)^{-1}$ of the drawing's rows with a proper density, the ten readings and its prior rows, times the constant density p_0 of the flat priors on Q_0, Q_{25} and $\ln h_{2,25}$. The Gaussian integral over the drawing's n_M entries about its MAP state $\hat{\mathbf{z}}_M$, with information $\mathbf{\Lambda}_M = \mathbf{J}_M^{\mathsf{T}}\mathbf{J}_M$, is the form of equation (4.19) on this drawing,

$$\ln p(\mathbf{y} \mid M) \approx \ln p_0 - \sum_{k\in M} \ln(\sqrt{2\pi}\,\sigma_k) - \mathcal{L}_M(\hat{\mathbf{z}}_M) + \frac{n_M}{2}\ln 2\pi - \frac{1}{2}\ln\det \mathbf{\Lambda}_M, \tag{7.7}$$

the height of the integrand at the MAP state times the volume of the posterior, which is $(2\pi)^{n_M/2}$ times $\det(\mathbf{\Sigma}_M)^{1/2}$. The constant p_0 and the ten reading rows' constants are common to the four drawings. A drawing that frees one more entry, with prior standard deviation σ, adds one prior row, with the constant $-\ln(\sqrt{2\pi}\,\sigma)$, and one dimension, which adds $\frac{1}{2}\ln 2\pi$; the $\sqrt{2\pi}$ cancels, and between two nested drawings

$$\ln B(M_{\text{big}} : M_{\text{small}}) = \left[\mathcal{L}_{\text{small}}(\hat{\mathbf{z}}_{\text{small}}) - \mathcal{L}_{\text{big}}(\hat{\mathbf{z}}_{\text{big}})\right] - \frac{1}{2}\ln\frac{\sigma^2 \det \mathbf{\Lambda}_{\text{big}}}{\det \mathbf{\Lambda}_{\text{small}}}, \tag{7.8}$$

with $B(M_A : M_B)$ the Bayes factor B_{AB} of equation (2.14). The first bracket is the gain in fit that the freed entry buys, half the drop in $2\mathcal{L}$. The second term is the Occam factor. For linear rows it is the logarithm of the ratio of the entry's prior width to its posterior width, so a wide prior on a term charges every drawing that frees it. Its units cancel: σ_b^2 is in K^2, and the determinant ratio in K^{-2}, because the bias column of $\mathbf{J}$ is in K^{-1}. Freeing two entries puts $\sigma_{\ln h}^2\sigma_b^2$ in place of σ^2. Two drawings that are not nested, such as M_{foul} and M_{drift}, compare by the difference of their ratios to the drawing both contain, $\ln B(M_A : M_B) = \ln B(M_A : M_{\text{neither}}) - \ln B(M_B : M_{\text{neither}})$. The posterior probability of each drawing is equation (2.47) with the prior probabilities of step 2, which, equal, cancel:

$$\Pr(M_j \mid \mathbf{y}) = \frac{p(\mathbf{y} \mid M_j)}{\sum_k p(\mathbf{y} \mid M_k)}, \qquad j \in \{0,\ \text{foul},\ \text{drift},\ \text{both}\}. \tag{7.9}$$

7.2.4 Model averaging and the full posterior

A quantity asked over the family, such as coil 1's change $\ln h_{1,25}$, which M_{neither} and M_{drift} pin at zero, has the mixture posterior of equation (3.39) with the probabilities (7.9) as weights,

$$p(f \mid \mathbf{y}) = \sum_j \Pr(M_j \mid \mathbf{y})\, p(f \mid \mathbf{y}, M_j), \tag{7.10}$$

a drawing that pins f contributing a point mass at the pinned value. This is model averaging, the form for a drawing that stays open. By the laws of total expectation and total variance, equation (3.34) with the drawing in place of the rate, with $\bar{f}_j$ and ς_j the mean and standard deviation of f under M_j ($\varsigma_j = 0$ where it is pinned) and $P_j = \Pr(M_j \mid \mathbf{y})$,

$$\mathbb{E}[f \mid \mathbf{y}] = \sum_j P_j\, \bar{f}_j, \qquad \mathrm{Var}(f \mid \mathbf{y}) = \sum_j P_j\, \varsigma_j^2 + \sum_j P_j\, (\bar{f}_j - \mathbb{E}[f \mid \mathbf{y}])^2, \tag{7.11}$$

the within-drawing part and the between-drawing part.

The full posterior of each drawing, without the Gaussian approximation, is computed by the reweighted draws of Section 5.2.4, with the drawing's Laplace normal as the proposal and the weights of equation (5.13). The same draws also estimate each evidence. With p_{L}

normalized, the mean of the weights before normalization is an unbiased estimate of the integral of $\exp[-\mathcal{L}_M]$, so that

$$\ln p(\mathbf{y} \mid M) \approx \ln p_0 - \sum_{k \in M} \ln(\sqrt{2\pi}\,\sigma_k) + \ln \left[\frac{1}{N} \sum_{j=1}^{N} \frac{\exp[-\mathcal{L}_M(\mathbf{z}^{(j)})]}{p_{\mathrm{L}}(\mathbf{z}^{(j)})} \right], \tag{7.12}$$

which is the Laplace evidence (7.7) when $\mathcal{L}_M$ is quadratic. For each drawing $N_{\mathrm{d}} = 200\,000$ draws from a seeded generator give effective sample sizes of 97.5 to 99.2 % of N_{d}, and the Monte Carlo standard errors come from 20 batches, as in Section 5.2.4; Section 7.4.1 states them.

7.3 The answers (step 4)

Every answer is read from the fits of Section 7.2.2 and the evidences of Section 7.2.3, and that value is the one reported. The full posterior of Section 7.2.4 stands beside each, a median beside a median and a probability beside a probability.

7.3.1 Fouling against drift

Asset manager's question. The approach temperature has risen 1.5 K; is the exchanger fouled or is the sensor drifting? The answer is a probability for each drawing of the family, and the Bayes factor of the two drawings the question names.

General form. A hypothesis comparison is the ratio of the evidences of two complete drawings fitted to the same evidence, equation (2.14),

$$B_{AB} = \frac{p(\mathbf{y} \mid M_A)}{p(\mathbf{y} \mid M_B)},$$

and run over a family of candidates it is a diagnosis, the ranking of equation (2.15),

$$\Pr(M_j \mid \mathbf{y}) \propto p(\mathbf{y} \mid M_j)\, \Pr(M_j).$$

Specialized form. The question is equation (2.14) with $M_A = M_{\mathrm{foul}}$ and $M_B = M_{\mathrm{drift}}$, and equation (2.15) over the four drawings of Table 7.2 with $\Pr(M_j) = \frac{1}{4}$, which is equation (7.9). Each evidence is the Laplace evidence (7.7) of the drawing's own fit. The two drawings the question names are not nested, so their ratio passes through M_{neither}, which both contain: by equation (7.8),

$$\ln B(M_{\mathrm{foul}} : M_{\mathrm{drift}}) = \ln B(M_{\mathrm{foul}} : M_{\mathrm{neither}}) - \ln B(M_{\mathrm{drift}} : M_{\mathrm{neither}}), \tag{7.13}$$

each term the gain in fit of its drawing less its Occam factor. Two sums of the probabilities answer the question in the asset manager's terms: coil 1 changed, $\Pr(M_{\mathrm{foul}} \mid \mathbf{y}) + \Pr(M_{\mathrm{both}} \mid \mathbf{y})$, and the sensor drifted, $\Pr(M_{\mathrm{drift}} \mid \mathbf{y}) + \Pr(M_{\mathrm{both}} \mid \mathbf{y})$.

Building blocks. Each drawing's objective at its MAP state and the logarithm of the determinant of its information (Section 7.2.2); the prior widths $\sigma_{\ln h}$ and σ_b in the Occam factor of equation (7.8); the prior probabilities of step 2.

Table 7.6: The evidence of each drawing by equation (7.7), up to the common constant of the flat priors, and the bookkeeping of equation (7.8) against M_{neither}. The determinants carry the units of the load and bias columns, which cancel in every ratio.

Drawing	$2\mathcal{L}(\hat{\mathbf{z}})$ [–]	$\ln\det\boldsymbol{\Lambda}_M$ [–]	$\ln p(\mathbf{y} \mid M)$ [–]	$\ln B$ vs M_{neither} [–]	gain [–]	Occam [–]	$\Pr(M \mid \mathbf{y})$ [–]
M_{neither}	6.3272	24.6860	−8.1846	0	–	–	0.1909
M_{foul}	4.0287	30.8221	−7.8009	+0.3838	1.1492	0.7655	0.2802
M_{drift}	3.5821	26.4593	−7.6987	+0.4859	1.3726	0.8867	0.3103
M_{both}	3.2175	32.1303	−8.0493	+0.1353	1.5548	1.4195	0.2186

The evidences. Table 7.6 evaluates equations (7.7) to (7.9). Each row is a drawing; the columns give twice its objective, $\ln\det\boldsymbol{\Lambda}_M$, its log evidence, its log Bayes factor against M_{neither} with that factor's two parts, the gain in fit and the Occam factor, and its posterior probability.

By equation (7.13), $\ln B(M_{\text{foul}} : M_{\text{drift}}) = 0.3838 - 0.4859 = -0.1021$, so

$$B(M_{\text{foul}} : M_{\text{drift}}) = 0.903.$$

The record does not prefer fouling to drift; it leans to drift, by a factor of 1.1075. Coil 1 changed, in M_{foul} or M_{both}, with probability 0.4988; the sensor drifted, in M_{drift} or M_{both}, with probability 0.5289; neither, with probability 0.1909. Each drawing's freed terms buy a gain in fit of 1.15 to 1.55 nats and pay an Occam factor of 0.77 to 1.42. M_{both} gains least per entry: against the drift alone, its derate term buys 0.1823 more gain in fit and pays 0.5329 more Occam factor, so $\ln B(M_{\text{both}} : M_{\text{drift}}) = -0.3506$. The record has room for one explanation of the rise, not two.

The comparison is near even for a reason the rows give. The two drawings put the return air at 26.9009 and 26.7356°C, against a return-air sensor of 0.3 K and a load known only to 2.83 kW, the posterior standard deviation of Q_{25} under M_{both}; unit 2's discharge they put closer still. Unit 1's discharge reading cannot vote, because both drawings fit it.

The posterior probabilities appear as bars in Figure 7.2, one per drawing, its height the posterior probability by the Laplace evidence; the dots are the probabilities by importance sampling, Section 7.4.1; the dashed line is the prior probability of each drawing; and the dotted outline marks the drawing of the stated truth, a check marker (Section 7.4.6). No bar is far from the dashed line: every posterior probability lies between 0.1909 and 0.3103, and the dots sit on the bar tops.

7.3.2 The part of the 1.5 K in question

Asset manager's question. How much of the 1.5 K is in question at all, once the load, the season and the other coil are taken out? The question asks how much of the dashboard's rise the drawings argue over.

General form. An estimation question about a derived quantity is summarized by its moments, equation (2.10),

$$\mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}] = \int f(\mathbf{z})\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}.$$

Specialized form. The question is equation (2.10) with f each of the parts of the rise of unit 1's approach reading, read at each drawing's MAP state, the first-order mean of a part. Unit 1's approach in window t, $A_{1,t}(\mathbf{z})$, is equation (5.18) at the window's known inputs. The rise the dashboard shows is split along a path of states at window 25's load and inputs, each

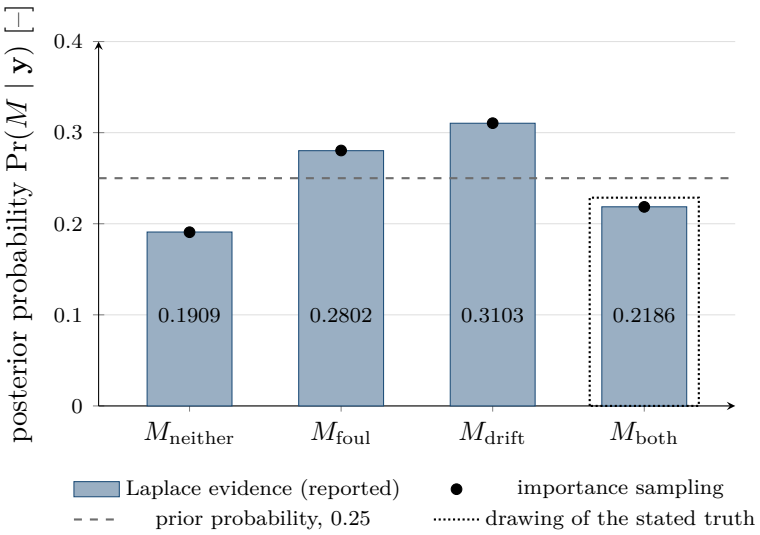


Figure 7.2: The posterior probabilities of the four drawings at window 25, equation (7.9), with equal prior probabilities. Bars: the Laplace evidence. Dots: importance sampling with each drawing's Laplace normal as the proposal. Dashed: the prior probability. Dotted outline: the drawing of the stated truth, a check marker (Section 7.4.6).

step changing one thing:

$$\begin{aligned}(y_{T_{d,1},25} - T_s(25)) - (y_{T_{d,1},0} - T_s(0)) = \; & \underbrace{A_{1,25}^{(0,0)} - \hat{A}_{1,0}}_{\text{heat (load and season)}} + \underbrace{A_{1,25}^{(0,2)} - A_{1,25}^{(0,0)}}_{\text{coil 2}} \\ & + \underbrace{A_{1,25}^{(1,2)} - A_{1,25}^{(0,2)}}_{\text{coil 1}} + \underbrace{b_{d,1}}_{\text{sensor}} + \epsilon_{\mathrm{r}},\end{aligned} \tag{7.14}$$

with $\hat{A}_{1,0}$ the fitted approach of window 0, $A_{1,25}^{(0,0)}$ the approach at today's load and inputs with both coils at their commissioned values $g_{u,0}$, $A_{1,25}^{(0,2)}$ the same with coil 2 at today's value $g_{2,0} + \ln h_{2,25}$, and $A_{1,25}^{(1,2)}$ with both coils at today's values. The remainder ϵ_{r} is the two readings' errors against their fits, $\epsilon_{\mathrm{r}} = (y_{T_{d,1},25} - T_s(25) - A_{1,25}^{(1,2)} - b_{d,1}) - (y_{T_{d,1},0} - T_s(0) - \hat{A}_{1,0})$, so the five terms add up to the dashboard's rise, 1.5011 K, which it shows as 1.50 K. One choice enters here: the order of the steps, heat, coil 2, coil 1, sensor. A different order moves the coil terms by second-order amounts, and the split is used only to take the dashboard's number apart, as Chapter 5's unaged counterfactual was. To first order the coil's and the sensor's terms together are the shift of unit 1's discharge reading that the two suspects make,

$$s_{d,1} = b_{d,1} + \frac{\partial T_{d,1}}{\partial \ln h_{1,25}} \ln h_{1,25}, \qquad \mathbf{g}_{s_{d,1}}^{\mathsf{T}} = \Big(0,\ 0,\ 0,\ 0,\ \frac{\partial T_{d,1}}{\partial \ln h_{1,25}},\ 0,\ 0,\ 1\Big), \tag{7.15}$$

with the slope that of the discharge line of equation (7.5) times 0.3, and its variance $\mathbf{g}_{s_{d,1}}^{\mathsf{T}} \boldsymbol{\Sigma}\, \mathbf{g}_{s_{d,1}}$.

Building blocks. Each drawing's MAP state (Section 7.2.2); the approach (5.18) through the forward equations (5.3) at each window's inputs; the two readings of Table 7.1; for $s_{d,1}$, the slope −9.4866 K of equation (7.5) and the block of $\boldsymbol{\Sigma}$ of equation (7.6).

The rise taken apart. Table 7.7 gives the split (7.14) under each drawing. Each row is a drawing; the columns are the five terms, then the part not in question, heat and coil 2 together, with its share of the rise.

Table 7.7: The rise of unit 1's approach reading since commissioning, 1.5011 K, taken apart under each drawing at its MAP state by equation (7.14).

Drawing	heat (load, season) [K]	coil 2 [K]	coil 1 [K]	sensor [K]	reading errors [K]	not in question [K]	share [%]
M_{neither}	+0.7096	+0.2019	0	0	+0.5896	0.9115	60.7
M_{foul}	+0.6062	+0.1200	+0.6577	0	+0.1173	0.7262	48.4
M_{drift}	+0.6348	+0.1784	0	+0.6604	+0.0275	0.8132	54.2
M_{both}	+0.6047	+0.1436	+0.3392	+0.4623	−0.0487	0.7483	49.9

About half of the 1.50 K is not in question. Load and season give 0.60 to 0.71 K and coil 2 0.12 to 0.20 K, together 0.7262 to 0.9115 K, 48.4 to 60.7 % of the rise across the drawings; under M_{both} it is 0.7483 K, 49.9 %. The drawings argue over the rest: 0.66 K under the single-term drawings, given wholly to the coil or wholly to the sensor, and 0.8015 K under M_{both}, shared. By equation (7.15), the combination unit 1's discharge reading fixes under M_{both} is $s_{d,1} = b_{d,1} - 9.4866 \ln h_{1,25}$, with mean 0.8017 K and standard deviation 0.4592 K, the same share by the gradient. Its standard deviation is smaller than either part's, 0.5150 K for the drift and 0.5644 K for the coil's term: the record knows the sum of the two suspects better than either. M_{neither} leaves 0.59 K to the reading errors, which is why it trails without being excluded. Over the family, the mixture (7.10) assigns 0.2584 K to coil 1 and 0.3060 K to the sensor.

Figure 7.3 stacks the split as one horizontal bar per drawing, and a fifth for the stated truth, each running along the horizontal axis through the five terms in fixed colors; the gray line is the dashboard's rise. The heat term is the longest segment of every drawing's bar, between 0.6047 and 0.7096 K; what changes from bar to bar is how the blue and gold segments share the rest, and how much is left to the purple one.

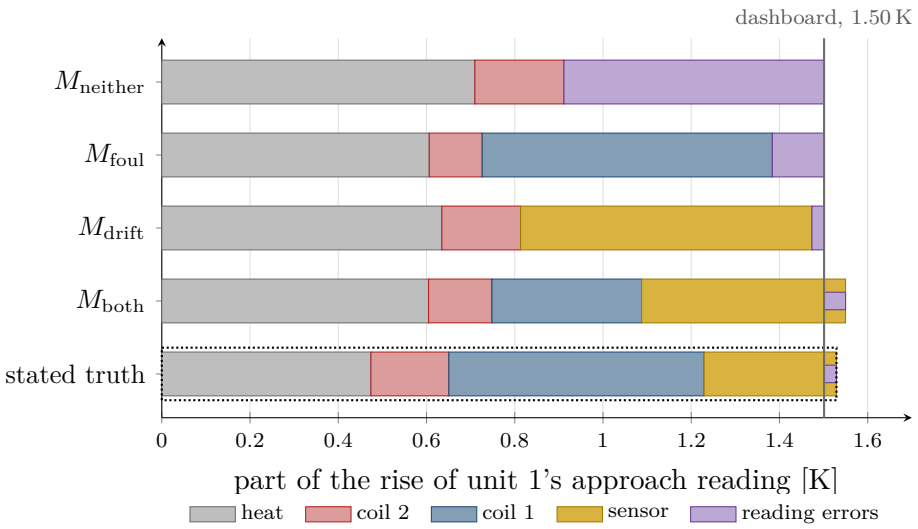


Figure 7.3: The rise of unit 1's approach reading taken apart, equation (7.14), under each drawing at its MAP state, and at the stated truth (dotted outline, a check marker, Section 7.4.6). Segments, from the left: heat (load and season), coil 2, coil 1, sensor, reading errors; a negative remainder is drawn as a thin bar back from the end of the sensor's term. Gray line: the rise the dashboard shows, 1.5011 K.

7.3.3 The two terms together

Asset manager's question. Could it be both, and how much of each? The question asks for the coil's change and the sensor's drift together, in the drawing in which both act.

General form. An estimation question is summarized by its moments, equation (2.10), and by a credible interval, equation (2.11), here at the levels 0.05 and 0.95:

$$\Pr(f_L \le f(\mathbf{z}) \le f_U \mid \mathbf{y}) = 0.90.$$

Specialized form. The question is equations (2.10) and (2.11) with f the pair $(\ln h_{1,25}, b_{d,1}) = (z_5, z_8)$ and $p(\mathbf{z} \mid \mathbf{y})$ the Laplace normal of M_{both}. Both are entries, so their gradients are the unit vectors $\mathbf{e}_5$ and $\mathbf{e}_8$ of the pattern (7.3), and their joint posterior is the normal with the means $(\hat{z}_5, \hat{z}_8)$ and the 2×2 block of $\mathbf{\Sigma}$,

$$\begin{bmatrix} \ln h_{1,25} \\ b_{d,1} \end{bmatrix} \Big| \, \mathbf{y}, M_{\text{both}} \sim \mathcal{N}\left(\begin{bmatrix} \hat{z}_5 \\ \hat{z}_8 \end{bmatrix}, \begin{bmatrix} \Sigma_{55} & \Sigma_{58} \\ \Sigma_{58} & \Sigma_{88} \end{bmatrix} \right), \tag{7.16}$$

each band the mean ± 1.6449 standard deviations and the probability of a positive drift $\Phi(\hat{z}_8/\sqrt{\Sigma_{88}})$. The block's shape is read from its standardized principal axes, the eigenvalues of the correlation matrix of the pair, $1 \pm \varrho$.

Building blocks. The MAP values and the block Σ_{55}, Σ_{58}, Σ_{88} of equation (7.6); the single-term fits of Table 7.5; the slopes of equation (7.5).

The joint posterior of the two terms. By equation (7.16), coil 1's change is $\ln h_{1,25} = -0.03579$ with standard deviation 0.05950, and the drift is $b_{d,1} = 0.4623\,\text{K}$ with standard deviation 0.5150 K, band $[-0.3848,\ 1.3093]\,\text{K}$, correlated by $+0.6414$. The probability of a positive drift is 0.8153, and of a drift above 0.5 K, 0.4708. The full posterior's values stand beside these in Table 7.13, and Section 7.4.1 states how far they move.

The ridge in numbers. Pinned one at a time, each term is fairly well known: $\ln h_{1,25}$ to 0.04615 under M_{foul}, and $b_{d,1}$ to 0.3978 K under M_{drift}. Freed together, they spread to 0.05950 and 0.5150 K. The standardized principal axes of the block have variances 0.3586 and 1.6414, standard deviations in the ratio 2.14: the posterior is a ridge along the line on which the discharge reading stays fixed. The priors hold the ridge in: with the bias prior flat the drift's standard deviation is 0.6019 K, and with both priors flat $\ln h_{1,25}$ has standard deviation 0.08264, $b_{d,1}$ has 0.6938 K and their correlation is 0.7854. The window alone resolves a drift to about ± 0.69 K. The slopes say why. Per unit of $\ln h_{1,25}$, unit 1's discharge reading moves by -9.4866 K, the return air by -5.8855 K and unit 2's discharge by -2.2845 K; per kelvin of drift the discharge reading moves by one kelvin and the others not at all. The rows that separate the terms have the smaller slopes, read by sensors of the same 0.3 K as the reading they would have to overrule.

Figure 7.4 maps the joint posterior. The horizontal axis is coil 1's change $\ln h_{1,25}$ and the vertical axis the drift $b_{d,1}$; the two ellipses are the contours at one and two standard deviations of M_{both}'s Laplace normal, tilted by the correlation $+0.6414$. The horizontal axis $b_{d,1} = 0$ is the drawing M_{foul}, with its 5–95 % band of $\ln h_{1,25}$ drawn on it; the vertical axis $\ln h_{1,25} = 0$ is M_{drift}, with its band of $b_{d,1}$; their crossing is M_{neither}. The dash-dotted line is $s_{d,1} = 0.80\,\text{K}$, equation (7.15), the combination the discharge reading fixes, and the cross is the stated truth. The ellipses lie along that line and reach both axes: each single-term drawing is a slice through the same ridge.

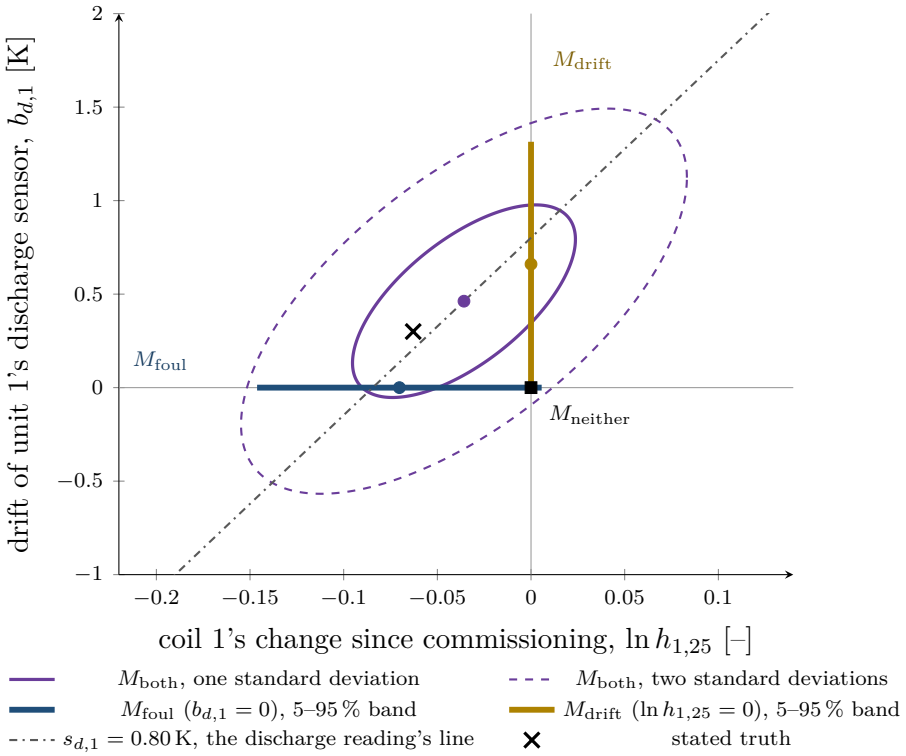


Figure 7.4: The joint posterior of coil 1's change and the drift of its discharge sensor under M_{both}, equation (7.16): means $(-0.03579,\ 0.46225\,\text{K})$, standard deviations 0.05950 and 0.5150 K, correlation +0.6414; contours drawn as in Figure 5.4. Bands: the 5–95 % bands of the single-term drawings on the axes that pin the other term; dots: the MAP values; square: M_{neither}. Dash-dotted: the line $b_{d,1} - 9.4866 \ln h_{1,25} = 0.80$ K. Cross: the stated truth, a check marker (Section 7.4.6).

7.3.4 Coil 1's health index with the drift allowed for

Asset manager's question. What is coil 1's health index once the drift is allowed for, and has it lost conductance? The index and its band, and the probability of a loss, read in the drawing that frees the drift.

General form. An estimation question about a derived quantity is summarized by its median and its credible interval, equation (2.11), at the levels 0.05 and 0.95,

$$\Pr(f_L \le f(\mathbf{z}) \le f_U \mid \mathbf{y}) = 0.90,$$

and the same posterior read at a level is its threshold reading, equation (2.12), here the probability of a loss.

Specialized form. The question is equation (2.11) with $f = \ln h_{1,25} = z_5$, read back to $h_{1,25}$ through the increasing map e^f, and its threshold reading, equation (2.12) with $f(\mathbf{z}) = -z_5$ and $f^* = 0$, the event of a loss, both under the Laplace normal of M_{both}. One choice enters here: the reported index is M_{both}'s, the drawing that allows for the drift without asserting either term. This is what keeps the index inside one drawing with a continuous posterior, and it holds the coil's term free even where the evidence leans to the drift alone; the model average over the family, the form for a drawing that stays open, stands beside it. Under M_{both} the index is an entry, its gradient $\mathbf{e}_5$, and as in equations (5.16) and (5.15),

$$\begin{aligned} \Pr(\ln h_{1,25} < 0 \mid \mathbf{y}, M_{\text{both}}) &= \Phi\!\left(\frac{-\hat{z}_5}{\sqrt{\Sigma_{55}}}\right), \\ \operatorname{med}(h_{1,25} \mid \mathbf{y}, M_{\text{both}}) = \mathrm{e}^{\hat{z}_5}, &\qquad \left[\mathrm{e}^{\hat{z}_5 - 1.6449\sqrt{\Sigma_{55}}},\ \mathrm{e}^{\hat{z}_5 + 1.6449\sqrt{\Sigma_{55}}}\right], \end{aligned} \tag{7.17}$$

the second the median and the 5–95 % band. Over the family, the mixture (7.10) of $\ln h_{1,25}$ has an atom at zero of weight $\Pr(M_{\text{neither}} \mid \mathbf{y}) + \Pr(M_{\text{drift}} \mid \mathbf{y})$ and two normal parts, so its probability of a loss is $\sum_{j\in\{\text{foul,both}\}} \Pr(M_j \mid \mathbf{y})\,\Phi(-\bar{f}_j/\varsigma_j)$, its moments are equation (7.11) and its quantiles are those of the mixture's distribution function.

Building blocks. $\hat{z}_5$ and Σ_{55} of M_{both} (Section 7.2.2); the posterior probabilities of Table 7.6; M_{foul}'s $\ln h_{1,25}$ and its standard deviation; the normal distribution function Φ.

The index with the drift allowed for. By equation (7.17), the reported index of coil 1 has median 0.9648 and band [0.8749, 1.0640], and

$$\Pr(\ln h_{1,25} < 0 \mid \mathbf{y}, M_{\text{both}}) = \Phi(0.03579/0.05950) = 0.7262.$$

The model average puts probability 0.5012 on $h_{1,25} = 1$, the mass of M_{neither} and M_{drift}, so its median sits on the atom, 1.0000; its band is [0.8837, 1.0166], its probability of a loss 0.4210, and the mean of $\ln h_{1,25}$ −0.02750, with variance 2.2765×10^{-3}, of which 1.3705×10^{-3} is within the drawings and 9.0606×10^{-4} between them, by equation (7.11). The full posterior gives the model average's probability of a loss as 0.4200 (Monte Carlo standard error 0.00030) and its band of $h_{1,25}$ as [0.8839, 1.0171], with the median on the atom. The drawings that take the sensor as sound read lower: M_{foul} puts coil 1 at 0.9322 (0.8640 to 1.0057) with probability of a loss 0.9360, and Chapter 5's drawing at 0.9146 (0.8398 to 0.9962) with 0.9571. The loss that Chapter 5's method demonstrates, at 0.96, is not demonstrated once the drift is allowed for, at 0.73, and the band reaches above one. Coil 2's change lies between −0.07072 and −0.10434 across the drawings, with standard deviations of 0.046 to 0.055: lower than coil 1's in every drawing, as the record shows it; how it got there is a question for the trend chapters, Chapters 10 to 12.

Figure 7.5 sets the routes to coil 1's index side by side. Each row is a route; its bar runs from $h = 1$ to its median, so its length is the loss the route reports, and the routes with a posterior carry their 5–95 % band as an error bar. From the top: the dashboard's raw-performance index, the readings taken as exact, Chapter 5's drawing, M_{foul}, the reported M_{both}, and the model average, whose median lies on the atom at one. The dotted line is the stated truth. The bars shorten from the top down, and the last three bands all contain one.

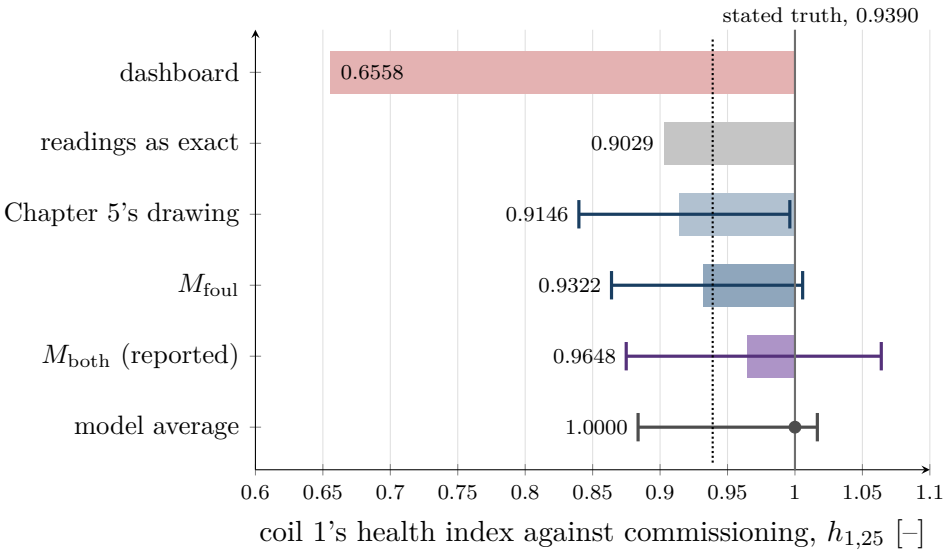


Figure 7.5: Coil 1's health index at window 25 by route. Bars run from $h = 1$ (gray line) to each route's median; error bars are the 5–95 % bands. The dashboard's index is the ratio of approaches; the readings taken as exact solve the air side for each coil; the three drawings are those of Tables 7.5 and 7.6, with Chapter 5's drawing the derate freed with no prior and the sensor taken as sound; the model average is equation (7.10), its median on the atom at one (dot). Dotted: the stated truth, a check marker (Section 7.4.6).

7.3.5 The residual test of the sensor

Asset manager's question. Would the readings themselves have flagged a faulty sensor? The question asks whether the fit the asset manager ran, Chapter 5's drawing with the sensor taken as sound, shows a misfit that would have raised the suspicion before the technician did.

General form. A prediction under a threshold is the probability of an event on a derived quantity, equation (2.12),

$$\Pr(f \geq f^* \mid \mathbf{y}) = \int \mathbf{1}[f \geq f^*]\, p(f \mid \mathbf{y})\, \mathrm{d}f,$$

here on a statistic of the readings rather than on the state.

Specialized form. The question is equation (2.12) with f the misfit $2\mathcal{L}$ of a window's rows at their fit, taken over windows the drawing would produce, and f^* the misfit of this window, under Chapter 5's drawing at window 25: the residual test, or bad-data detection. One choice enters here: the rows are taken as linear over the spread of the fit, which makes the residuals at the fit Gaussian with a covariance the Jacobian gives; it holds to the extent that the fit moves little when the errors are redrawn. For linear rows the residual vector at the MAP state is a

projection of the rows' normalized errors $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$,

$$\mathbf{r}(\hat{\mathbf{z}}) = -(\mathbf{I} - \mathbf{K})\,\boldsymbol{\epsilon}, \qquad \mathbf{K} = \mathbf{J}\boldsymbol{\Sigma}\mathbf{J}^{\mathsf{T}}, \tag{7.18}$$

because the fit absorbs the part of the errors that the columns of $\mathbf{J}$ can explain. The hat matrix $\mathbf{K}$ is symmetric and idempotent, so $\mathbf{I} - \mathbf{K}$ is too, and the residual vector is Gaussian with mean zero and covariance $\mathbf{I} - \mathbf{K}$. Row k's residual has variance $1 - K_{kk}$, its redundancy number, with K_{kk}, the diagonal entry of the hat matrix, its leverage. Summed over the n_{r} rows, by the cyclic property of the trace,

$$\sum_{k=1}^{n_{\mathrm{r}}} (1 - K_{kk}) = n_{\mathrm{r}} - \operatorname{tr}(\boldsymbol{\Sigma}\mathbf{J}^{\mathsf{T}}\mathbf{J}) = n_{\mathrm{r}} - n_{\mathrm{z}}, \tag{7.19}$$

the degrees of freedom of a fit of n_{z} entries, and $2\mathcal{L}(\hat{\mathbf{z}}) = \boldsymbol{\epsilon}^{\mathsf{T}}(\mathbf{I} - \mathbf{K})\boldsymbol{\epsilon}$ is a chi-square on $n_{\mathrm{r}} - n_{\mathrm{z}}$ degrees of freedom, so that the probability asked is $\Pr(\chi^2_{n_{\mathrm{r}} - n_{\mathrm{z}}} > 2\mathcal{L}(\hat{\mathbf{z}}))$. The test for one reading is its normalized residual,

$$r_k^{\mathrm{N}} = \frac{r_k(\hat{\mathbf{z}})}{\sqrt{1 - K_{kk}}} \sim \mathcal{N}(0, 1), \tag{7.20}$$

the reading's departure from what the other rows predict for it, over the spread of that departure; with every residual written predicted less observed, a negative value is a reading above the prediction. A row whose leverage is near one has a redundancy number near zero: its own reading sets its fitted value, and no misfit can show there.

Building blocks. The fit of Chapter 5's drawing at window 25 (Section 7.2.2): its residual vector, $\mathbf{J}$ and $\boldsymbol{\Sigma}$, thirteen rows on seven entries; the chi-square distribution function.

The residual test. Table 7.8 gives the test row by row. Each row of the table is a reading row of the fit; the columns give its residual, its leverage, its redundancy number and its normalized residual (7.20).

Table 7.8: The residual test of Chapter 5's drawing at window 25, the sensor taken as sound: the ten reading rows at the fit.

Reading row	Residual r_k [–]	Leverage K_{kk} [–]	Redundancy $1 - K_{kk}$ [–]	Normalized r_k^{N} [–]
window 0, return air	+0.2153	0.3752	0.6248	+0.2723
window 0, discharge 1	−0.3644	0.7380	0.2620	−0.7120
window 0, discharge 2	−0.4419	0.7362	0.2638	−0.8602
window 0, load meter	−0.1365	0.8446	0.1554	−0.3463
window 0, rack-1 inlet	+0.4814	0.4977	0.5023	+0.6793
window 25, return air	+1.2582	0.3895	0.6105	+1.6104
window 25, discharge 1	−0.3254	0.8126	0.1874	−0.7517
window 25, discharge 2	−0.3254	0.8126	0.1874	−0.7517
window 25, load meter	−0.5586	0.8564	0.1436	−1.4743
window 25, rack-1 inlet	−0.6074	0.6159	0.3841	−0.9801

Twice the objective is 3.4019 on $13 - 7 = 6$ degrees of freedom, and $\Pr(\chi^2_6 > 3.40) = 0.757$. The reading rows' part is 3.1007, against the sum of their redundancy numbers, 3.3213, its expected value by equation (7.19). Unit 1's discharge row of window 25 has residual −0.3254, leverage 0.8126, redundancy number 0.1874 and normalized residual −0.7517, and unit 2's discharge row the same four numbers: with equal capacity rates and equal sensor standard deviations, the energy balance of the two units gives both discharge rows the same residual and leverage, so the residual test cannot even tell which of the two discharge sensors is off.

The largest normalized residual is the return air of window 25, +1.6104, then its load meter, −1.4743. Nothing is flagged. With coil 1's change free, the coil absorbs the sensor's reading: 81 % of that reading's own value sits in its fit, and a drift on a reading whose entry is free is invisible to the residual test. Pin the coil, as M_{neither} does, and the same row's redundancy number rises to 0.4746 and its normalized residual to −1.8068, still under two in size, with $\Pr(\chi^2_7 > 6.33) = 0.502$. The residual test and the evidence agree: the window is consistent with every drawing.

7.4 The checks (step 5)

Each answer of Section 7.3 rests on the Laplace evidence, on the priors of the freed terms and on one record's draws of the sensor errors, and this hall lets each be tested against the calculation it stands for.

7.4.1 The full posterior against the Laplace evidence

The reweighted draws of Section 7.2.4 give each drawing's evidence by equation (7.12), with 200 000 draws per drawing and effective sample sizes of 99.2, 98.9, 98.9 and 97.5 % for M_{neither}, M_{foul}, M_{drift} and M_{both}. The log evidences lie above the Laplace values by 0.0013, 0.0029, 0.0027 and 0.0022, with Monte Carlo standard errors of 0.00018, 0.00027, 0.00025 and 0.00032 by batch means. Each gap is larger than its standard error, so the gaps are resolved and real: the objectives are not quite quadratic. But the gaps nearly cancel in every ratio. No log Bayes factor between two drawings moves by more than 0.0016 and no posterior probability by more than 0.0002: the probabilities by importance sampling are 0.1907, 0.2804, 0.3104 and 0.2185, against 0.1909, 0.2802, 0.3103 and 0.2186, and $\ln B(M_{\text{foul}} : M_{\text{drift}})$ is −0.1019 against −0.1021. The model average's probability of a loss is 0.4200 against 0.4210, with a standard error of 0.00030: resolved, and too small to matter. Within M_{both} the full posterior is slightly skewed, and it does move the answers read from it. Coil 1's index has median 0.9659, against the Laplace 0.9648, so that the maintenance log's 0.96 reads 0.97 on this route; its band is [0.8749, 1.0642], against [0.8749, 1.0640], and its probability of a loss 0.7227 (Monte Carlo standard error 0.00094), against 0.7262. The median drift is 0.4723 K with band [−0.3695, 1.3196] K, against 0.4623 K and [−0.3848, 1.3093] K; the probability of a positive drift is 0.8212 against 0.8153, and the correlation 0.6393 against 0.6414. These moves are resolved by the draws and small beside the bands; none turns a verdict. The Laplace values stay the reported ones, as the one fit of the part reads them, and the full posterior's stand beside them in Table 7.13.

7.4.2 The Savage–Dickey ratio and the closed form

The Savage–Dickey ratio. For a nested pair, in which M_{small} is M_{big} with one entry z_i pinned at zero and the other entries carry the same priors in both, the Bayes factor is the posterior density of the freed entry at the pinned value over its prior density there,

$$B(M_{\text{small}} : M_{\text{big}}) = \frac{p(z_i = 0 \mid \mathbf{y}, M_{\text{big}})}{p(z_i = 0 \mid M_{\text{big}})}, \tag{7.21}$$

which Exercise 7.2 derives. It needs only the larger drawing's fit. On the Laplace marginals it gives $\ln B(M_{\text{foul}} : M_{\text{both}}) = +0.2608$ against the Laplace evidences' +0.2485, $\ln B(M_{\text{drift}} : M_{\text{both}}) = +0.3384$ against +0.3506, $\ln B(M_{\text{neither}} : M_{\text{foul}}) = -0.3851$ against −0.3838, and $\ln B(M_{\text{neither}} : M_{\text{drift}}) = -0.4562$ against −0.4859. The two routes agree for linear rows; the gaps, at most 0.0297, measure the rows' nonlinearity, and none changes a sign.

The closed form of a bias on one reading. For linear rows the Bayes factor (7.8) of a bias freed on reading k against the drawing without it has a closed form in the statistic of the residual test, derived in *Reliability Engineering by Derivation*, Chapter 22, and used in *Operations and Control by Derivation*, Chapter 7. With the leverage K_{kk} and the normalized residual r_k^{N} of equation (7.20), both read in the drawing without the bias,

$$\psi_k = \frac{\sigma_b^2\,(1 - K_{kk})}{\sigma_k^2}, \qquad \ln B = \frac{1}{2}\left[\frac{\psi_k}{1+\psi_k}\left(r_k^{\mathrm{N}}\right)^2 - \ln(1+\psi_k)\right], \qquad \mathrm{sd}(b \mid \mathbf{y}) = \frac{\sigma_b}{\sqrt{1+\psi_k}}, \tag{7.22}$$

with ψ_k the ratio of the bias's prior variance to the variance of the reading's departure from the other rows. The evidence for a drift depends on the window only through the normalized residual, and on the sensor's place in the graph only through ψ_k; a bias loses to the drawing without it when $(r_k^{\mathrm{N}})^2$ falls below $(1+\psi_k)\ln(1+\psi_k)/\psi_k$. For M_{drift} against M_{neither}, unit 1's discharge row has redundancy number 0.4746 in M_{neither}, so $\psi = 5.2731$, and $r^{\mathrm{N}} = -1.8068$, which gives $\ln B = +0.4539$, against the Laplace $+0.4859$; the break-even is $|r^{\mathrm{N}}| = 1.4780$, and the closed form's drift is 0.6614 K with standard deviation 0.3993 K, against the fit's 0.6604 and 0.3978. For M_{both} against M_{foul}, the same row has redundancy number 0.2523 in M_{foul}, larger than the 0.1874 of Chapter 5's drawing because the derate-term prior holds the coil's change in; then $\psi = 2.8034$ and $r^{\mathrm{N}} = -1.0504$, below its break-even of 1.3463, so $\ln B = -0.2613$, against the Laplace -0.2485, with a drift of 0.4624 K and standard deviation 0.5128 K against the fit's 0.4623 and 0.5150. The closed form says why freeing the coil weakens the case for a drift: it cuts the sensor row's redundancy number from 0.47 to 0.25, so the other rows predict that reading less well, and the same departure counts for less.

7.4.3 The limit of exact rows and the curvature term

The fit runs on the reduced state. The check fits M_{both}'s own full state of eighteen entries, the eight of equation (7.1) and each window's $(T_r, T_{d,1}, T_{d,2}, T_c, T_1)$, with the ten physical rows (5.2) at a common finite tolerance beside the readings and priors, and lets the tolerance shrink. Table 7.9 gives the result: each row is a tolerance, the columns the two terms with their standard deviations, and the last row the reduced state. At 0.1 the spreads are wider by the discrepancy the rows allow; at 0.01 the full state differs from the reduced state at most in the last printed digit, and at 0.001 it reproduces it to the printed digits. Each fit took six iterations.

Table 7.9: The limit of exact physical rows on the eighteen-entry full state of M_{both}: coil 1's change and the drift with their standard deviations at three tolerances, against the reduced state.

Tolerance σ_p [kW or K]	$\ln h_{1,25}$ [–]	sd [–]	$b_{d,1}$ [K]	sd [K]
0.1	−0.03382	0.05976	0.4759	0.5166
0.01	−0.03576	0.05950	0.4624	0.5150
0.001	−0.03579	0.05950	0.4623	0.5150
reduced state	−0.03579	0.05950	0.4623	0.5150

The covariance is the Gauss–Newton form, without the curvature term $\mathbf{C} = \mathbf{H} - \mathbf{J}^{\mathsf{T}}\mathbf{J}$ of equation (4.44). The full Hessian of the objective at M_{both}'s MAP state, computed by differences, gives standard deviations of 0.05944 for $\ln h_{1,25}$ and 0.5143 K for $b_{d,1}$, against Gauss–Newton's 0.05950 and 0.5150, ratios of 0.9991 and 0.9988. The largest entry of $\mathbf{C}$ is 5.85×10^{-4} of the largest entry of $\mathbf{J}^{\mathsf{T}}\mathbf{J}$, and the Hessian's determinant in place of $\mathbf{J}^{\mathsf{T}}\mathbf{J}$'s in equation (7.7) moves $\ln p(\mathbf{y} \mid M_{\mathrm{both}})$ by -0.0023. The Gauss–Newton covariance and evidence are the ones reported.

7.4.4 Sensitivity to the priors

The evidence depends on the widths of the freed terms' priors through the Occam factor of equation (7.8), and the ranking on the prior probabilities. Table 7.10 repeats the comparison under other priors. Each row is one pair of widths, or the other prior probabilities of step 2; the columns give $\ln B(M_{\text{foul}} : M_{\text{drift}})$, the four posterior probabilities, the model average's probability of a loss, and the reported answer, M_{both}'s index of coil 1 with its probability of a loss.

Table 7.10: The comparison and the reported index under other priors: the widths of the derate term and the bias, and, in the last row, the prior probabilities 0.5 for M_{neither} and 0.5/3 for each other drawing. Pr(loss) is the model average's; $h_{1,25}$ and P_{both} are M_{both}'s median index and probability of a loss.

$\sigma_{\ln h}$, σ_b [–, K]	$\ln B$ F:D [–]	Pr(M_{neither}) [–]	Pr(M_{foul}) [–]	Pr(M_{drift}) [–]	Pr(M_{both}) [–]	Pr(loss) [–]	$h_{1,25}$ [–]	P_{both} [–]
0.10, 1.0 (the chapter's)	−0.1021	0.1909	0.2802	0.3103	0.2186	0.4210	0.9648	0.7262
0.05, 1.0	−0.1086	0.1804	0.2630	0.2932	0.2634	0.4066	0.9827	0.6631
0.20, 1.0	−0.4877	0.2307	0.2302	0.3750	0.1642	0.3436	0.9524	0.7581
0.10, 0.5	−0.1469	0.1789	0.2626	0.3042	0.2543	0.4561	0.9503	0.8269
0.10, 2.0	+0.3380	0.2367	0.3475	0.2478	0.1680	0.4377	0.9731	0.6695
0.05, 0.5	−0.1534	0.1709	0.2493	0.2906	0.2892	0.4382	0.9732	0.7550
0.20, 2.0	−0.0476	0.2872	0.2866	0.3006	0.1256	0.3605	0.9622	0.6984
prior probabilities 0.5 on M_{neither}	–	0.4145	0.2028	0.2246	0.1582	–	0.9648	0.7262

A wider prior on a term charges every drawing that frees it a larger Occam factor, so its effect runs drawing by drawing. Doubling $\sigma_{\ln h}$ to 0.20 moves probability from the two drawings with a derate term to the two without: M_{neither} rises from 0.1909 to 0.2307 and M_{drift} from 0.3103 to 0.3750, while M_{foul} falls to 0.2302 and M_{both} to 0.1642. Doubling σ_b to 2.0 K moves it from the drift drawings to M_{neither} and M_{foul}, and flips the sign of $\ln B(M_{\text{foul}} : M_{\text{drift}})$, to +0.3380. Halving a width does not simply do the reverse. At $\sigma_b = 0.5$ K, M_{drift} falls slightly, from 0.3103 to 0.3042, although its Occam factor is smaller, because its drift of 0.66 K sits beyond the narrower prior and the fit pays for it; M_{both}, which needs only a drift of 0.46 K, rises to 0.2543. No prior makes the comparison decisive: across the widths $|\ln B|$ is at most 0.4877, and every drawing keeps between 0.1256 and 0.3750. The ranking moves with the widths because the readings say little, Lindley's paradox in the small: where the evidence is weak, the prior widths decide the Occam factor and with it the order. The alternative prior probabilities of step 2, 0.5 on M_{neither}, put M_{neither} first at 0.4145, and the three others keep their order; they leave M_{both}'s own fit, and so the reported index, unchanged.

The reported index moves along the ridge, the other way from the drift. A narrower prior on the drift, $\sigma_b = 0.5$ K, holds the drift to 0.2580 K and gives more of the shift to the coil: the index falls to 0.9503 and the probability of a loss rises to 0.8269. A narrower prior on the coil's change, $\sigma_{\ln h} = 0.05$, does the reverse, with the drift at 0.5640 K, the index at 0.9827 and the probability of a loss at 0.6631. Across the widths M_{both}'s index lies between 0.9503 and 0.9827 and its probability of a loss between 0.6631 and 0.8269: under no prior is a loss demonstrated in the sense of Chapter 5's 0.96. The model average's probability of a loss moves between 0.3436 and 0.4561.

7.4.5 The verdict as one record's draw

The comparison of Section 7.3.1 is one record's draw of the sensor errors. The family is refitted on 300 records with fresh errors at the stated truth, rounded to 0.01 as the record's are, from a

seeded generator; a fraction near 0.5 then has a Monte Carlo standard error of 0.029. Table 7.11 gives the result. Each row is a truth: the record's, with a drift of 0.30 K at window 25, and the same window with the drift set to 1.0 and to 2.0 K. The columns give the median and 5–95 % range of $\ln B(M_{\text{foul}} : M_{\text{drift}})$ over the records, the fraction of records in which the family decides a drift (the two drift drawings together hold 0.95 or more), the fraction in which it decides a change of coil 1 (the two derate drawings hold 0.95 or more), and the fraction in which M_{both} leads.

Table 7.11: The verdict over 300 records with fresh sensor errors at the stated truth of window 25, with the record's drift and with the drift set larger.

Truth used	$\ln B$ F:D, median [5–95 %] [–]	decides a drift [–]	decides coil 1 [–]	M_{both} leads [–]
drift 0.30 K (the record's)	+0.268 [−1.463, +1.772]	0.010	0.010	0.033
drift set to 1.0 K	−0.533 [−3.791, +1.818]	0.150	0.020	0.300
drift set to 2.0 K	−3.609 [−8.989, +0.499]	0.810	0.030	0.673

At the record's truth the family almost never decides: it decides a drift in 0.010 of the records and a change of coil 1 in 0.010, and M_{both}, the drawing of the truth, leads in only 0.033, while M_{neither} leads in 0.320, M_{foul} in 0.397 and M_{drift} in 0.250. This record's −0.1021 sits at the 0.310 quantile of the first row, an ordinary draw. What one window can separate is the part not in question, heat and coil 2, and the shift $s_{d,1}$ of equation (7.15), the sum of the coil's and the sensor's shares. What it cannot separate is their split, for any drift near the record's size: even at 1.0 K the family decides a drift in 0.150 of records, and only at 2.0 K does it decide one in 0.810. A change of coil 1 is decided in at most 0.030 of records in any row; in every row its true change is the record's −0.0629, and only the drift was set larger.

7.4.6 The answers against the stated truth

The last check sets the answers against the stated truth of the record (Section 7.6.5). Table 7.12 collects the routes: each row is a route or a variation, the columns $\ln B(M_{\text{foul}} : M_{\text{drift}})$, M_{both}'s probability, and coil 1's change and the drift under M_{both}, each with its standard deviation; the last row is the truth.

Table 7.12: The routes to the comparison and to M_{both}'s two terms at window 25, and the stated truth. Medians where marked; a dash where a route does not give the quantity. The closed forms' drift is that of the drawing named; their log Bayes factors, and the Savage–Dickey ratios, are compared with the Laplace evidence in Section 7.4.2.

Route	$\ln B$ F:D [–]	$\Pr(M_{\text{both}})$ [–]	$\ln h_{1,25}$ [–]	sd [–]	$b_{d,1}$ [K]	sd [K]
Laplace (reported)	−0.1021	0.2186	−0.03579	0.05950	0.4623	0.5150
Importance sampling, full posterior	−0.1019	0.2185	−0.03471 (med.)	–	0.4723 (med.)	–
Closed form, M_{drift} against M_{neither}	–	–	–	–	0.6614	0.3993
Closed form, M_{both} against M_{foul}	–	–	–	–	0.4624	0.5128
Newton covariance	–	–	–	0.05944	–	0.5143
Full state, $\sigma_p = 0.1$	–	–	−0.03382	0.05976	0.4759	0.5166
Full state, $\sigma_p = 0.001$	–	–	−0.03579	0.05950	0.4623	0.5150
Stated truth (the drawing M_{both})	–	–	−0.0629	–	0.30	–

Coil 1's true change is −0.0629, a true index of 0.9390, and the true drift is 0.30 K: the truth is M_{both}, to which the record gives 0.2186. M_{both}'s values lie +0.46 standard deviations from the true change and +0.32 from the true drift, inside both bands. M_{foul}'s change lies −0.16 standard deviations from the truth and Chapter 5's drawing's −0.51; M_{drift}'s drift lies

+0.91 standard deviations from the true drift. At window 25 the sensor reads 19.81°C against a true discharge of 19.4039°C, an error of +0.4061 K, of which 0.30 is the drift. The truth's split of the rise is heat +0.4740, coil 2 +0.1766, coil 1 +0.5786, sensor +0.3000 and reading errors −0.0281 K, the dotted bar of Figure 7.3. Every drawing's heat term, 0.60 to 0.71 K, exceeds the truth's 0.47 K because the load meter reads 134.28 kW against the true 127.62 kW, +2.22 standard deviations: the luck of the record's draws, not a defect of the method. The true loads are 110.00 and 127.62 kW, and the true β_1 is 0.22. The single-term drawings put the whole of the shared shift on one suspect, and the truth has both.

7.5 The report (step 6)

The record cannot tell whether coil 1 fouled or its sensor drifted: the Bayes factor of fouling against drift is 0.90, and the four drawings hold 0.19 (neither), 0.28 (fouled), 0.31 (drifted) and 0.22 (both). Half of the dashboard's 1.50 K, 49.9 %, is not in question at all: 0.60 K is the load and the season, and 0.14 K is coil 2. The other 0.80 K is shared between coil 1 and the sensor in a proportion the window cannot fix. With the drift allowed for, coil 1 is at 0.96 (0.87 to 1.06) and its loss is not demonstrated, $P = 0.73$; with the sensor taken as sound it would read 0.91, with $P = 0.96$. The residual test sees nothing, because the free coil absorbs the sensor. A side-by-side check of the sensor, ten paired readings, settles the sensor, not the coil: it would decide a drift in 0.680 of its outcomes and a change of coil 1 in 0.007, and the coil's index needs the reference reading of the discharge air taken in the same visit.

7.5.1 The answers and their bands

Table 7.13 gives the answers in the order of Table 7.3. Each row is a question; the second column gives the reported value, from the Laplace evidence and the Laplace normal, the third the full posterior's value beside it, and the fourth the band: the 5–95 % band of a term or an index, the range over the priors of Table 7.10 for the comparison, and the range across the drawings for the split.

7.5.2 The entry in the maintenance log

The condition assessment of window 25 is the next condition entry of the two-unit hall's maintenance log, after the commissioning test of window 0 and the assessment of window 18. It uses nothing later than window 25, and in the values of Table 7.13 it reads:

> *Window 25 (Jan Y3), condition assessment against the commissioning baseline (window 0); unit 1's discharge sensor under suspicion.*
> Unit 1's approach +1.50 K since commissioning: 0.60 K load and season, 0.14 K coil 2; the remaining 0.80 K is coil 1 or the sensor, not separated by the window.
> Fouled against drifted: Bayes factor 0.90. Neither 0.19, fouled 0.28, drifted 0.31, both 0.22.
> Coil 1, drift allowed for: health index 0.96 (5–95 % band 0.87 to 1.06; 0.97 by the full posterior), $P(\text{loss}) = 0.73$ (0.72 by the full posterior). The 0.91 of the fit that takes the sensor as sound is not used.
> Residual test: no flag ($P = 0.76$).
> Action: the cleaning work order for coil 1 is held until the sensor check.
> Next measurement: a side-by-side check of unit 1's discharge sensor against a reference thermometer, ten paired readings, which settles the sensor; in the same visit, a reference reading of unit 1's discharge air, for the coil's index; both before the next condition window.

Whether coil 1 is to be cleaned, once the checks are in, needs the decision level that Chapter 19, in Volume II, calibrates; this entry sets none.

Table 7.13: The report for the two-unit hall at window 25 (Jan Y3) against its commissioning test. Reported values from the Laplace evidence and the Laplace normal of each drawing; bands are 5–95 %, or the ranges stated.

Asset manager's question	Reported value	Full posterior	Band
The approach temperature has risen 1.5 K; is the exchanger fouled or is the sensor drifting?	$B(M_{\text{foul}} : M_{\text{drift}}) = 0.903$, $\ln B = -0.1021$ [–]; M_{neither} 0.1909, M_{foul} 0.2802, M_{drift} 0.3103, M_{both} 0.2186 [–]	$\ln B = -0.1019$; 0.1907, 0.2804, 0.3104, 0.2185 [–]	over the priors: $\ln B$ −0.4877 to +0.3380; $\Pr(M_{\text{both}})$ 0.1256 to 0.2892 [–]
How much of the 1.5 K is in question at all, once the load, the season and the other coil are taken out?	under M_{both}: heat 0.6047, coil 2 0.1436, coil 1 and sensor 0.8015 [K]; not in question 0.7483 [K], 49.9 [%]	–	across the drawings: not in question 0.7262 to 0.9115 [K]
Could it be both, and how much of each?	$\ln h_{1,25} = -0.03579$ (sd 0.05950) [–]; $b_{d,1} = 0.4623$ (sd 0.5150) [K]; correlation +0.6414 [–]	$b_{d,1}$ median 0.4723 [K]; correlation 0.6393 [–]	$b_{d,1}$: −0.3848 to 1.3093 [K]; full posterior −0.3695 to 1.3196 [K]
What is coil 1's health index once the drift is allowed for, and has it lost conductance?	$h_{1,25} = 0.9648$ [–], $\Pr(\text{loss}) = 0.7262$ [–]; model average $\Pr(\text{loss}) = 0.4210$ [–]	$h_{1,25}$ median 0.9659 [–], $\Pr(\text{loss}) = 0.7227$ [–]; model average 0.4200 [–]	$h_{1,25}$: 0.8749 to 1.0640; full posterior 0.8749 to 1.0642 [–]; model average 0.8837 to 1.0166 [–]
Would the readings themselves have flagged a faulty sensor?	no: $2\mathcal{L} = 3.4019$ on 6 degrees of freedom, $\Pr = 0.757$ [–]; largest $\lvert r^{\text{N}}\rvert = 1.6104$, the return air [–]	–	unit 1's discharge row: $r^{\text{N}} = -0.7517$, redundancy 0.1874 [–]

Common mistake

Reading a rise of a coil's approach as fouling when the coil's own discharge sensor reports it. That reading cannot tell the coil from its sensor, and a fit that takes the sensor as sound turns any drift into a loss with a confident band. The comparison of drawings says how far the other readings separate the two, and a reference beside the sensor separates them.

7.6 The next measurement (step 7)

The report leaves the comparison near even and the drift anywhere from −0.38 to 1.31 K. A further measurement in today's window is one more row of the residual vector, and what it removes from each term, and how far it would move the comparison, are known before its value is.

7.6.1 The preposterior of a check

A candidate measurement a is one row or more in today's window, with Jacobian lines $\mathbf{J}_a$ in the columns of M_{both}'s state, each the gradient of the quantity it reads over its standard deviation. Its preposterior covariance is equation (4.65) on M_{both}'s posterior,

$$\boldsymbol{\Sigma}_a = \left(\boldsymbol{\Sigma}^{-1} + \mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a\right)^{-1}, \qquad \text{sd}_a(b_{d,1}) = \sqrt{\mathbf{e}_8^{\mathsf{T}}\boldsymbol{\Sigma}_a\,\mathbf{e}_8}, \quad \text{sd}_a(\ln h_{1,25}) = \sqrt{\mathbf{e}_5^{\mathsf{T}}\boldsymbol{\Sigma}_a\,\mathbf{e}_5}, \tag{7.23}$$

and the width it takes off each term is $1 - \text{sd}_a/\text{sd}$. For the comparison itself, the check is predicted under each of the two drawings the question names, as a reading of a quantity $f_a(\mathbf{z})$

of gradient $\mathbf{g}_a$ with standard deviation σ_a,

$$\tilde{y}_a \mid \mathbf{y}, M_l \approx \mathcal{N}(\hat{y}_{a,l},\ v_{a,l}), \qquad \hat{y}_{a,l} = f_a(\hat{\mathbf{z}}_l), \quad v_{a,l} = \mathbf{g}_a^{\mathsf{T}} \boldsymbol{\Sigma}_l\, \mathbf{g}_a + \sigma_a^2, \qquad l \in \{\text{foul},\ \text{drift}\}. \tag{7.24}$$

By the product rule $p(\mathbf{y}, \tilde{y}_a \mid M) = p(\mathbf{y} \mid M)\, p(\tilde{y}_a \mid \mathbf{y}, M)$, so the check changes $\ln B(M_{\text{foul}} : M_{\text{drift}})$ by the log ratio of the two predictive densities at the value read. If M_l is true, the expected change in its favor is the Kullback–Leibler divergence of the other drawing's predictive from its own; for two normals,

$$U_l(a) = \frac{1}{2}\left[\ln \frac{v_{a,l'}}{v_{a,l}} + \frac{v_{a,l} + (\hat{y}_{a,l} - \hat{y}_{a,l'})^2}{v_{a,l'}} - 1\right], \qquad \nu_a = \frac{|\hat{y}_{a,l} - \hat{y}_{a,l'}|}{\sqrt{\frac{1}{2}\,(v_{a,l} + v_{a,l'})}}, \tag{7.25}$$

with l' the other drawing and ν_a the separation of the two predictions in units of their mean spread. The two directions differ when the spreads differ: the drawing with the wider predictive is the easier one to confirm.

The check over the whole family. The expected change of one Bayes factor does not say whether a check would decide the question. For a check on the drift alone, the four drawings can be carried through it in closed form on their Laplace marginals: a reading $\tilde{y}_a$ of the drift with standard deviation σ_a multiplies each drawing's evidence by its predictive density,

$$\ln p(\mathbf{y}, \tilde{y}_a \mid M_j) = \ln p(\mathbf{y} \mid M_j) + \ln \varphi_{M_j}(\tilde{y}_a), \tag{7.26}$$

with φ_{M_j} the normal density of mean zero and variance σ_a^2 for the drawings that pin the drift, and of mean $\hat{b}_j$ and variance $\mathrm{sd}(b_{d,1} \mid \mathbf{y}, M_j)^2 + \sigma_a^2$ for those that free it, and the probabilities (7.9) follow again. Drawing the outcome from M_{both}'s predictive gives the fraction of outcomes in which the family would decide a drift, the two drift drawings holding 0.95 or more, and the fraction in which it would decide a change of coil 1.

The candidates. The side-by-side check reads the logged sensor and a reference thermometer of 0.1 K in the same air, so that their difference is the drift plus the two errors, a row on $b_{d,1}$ alone,

$$r_{\text{sbs}} = \frac{b_{d,1} - y_b}{\sqrt{0.3^2 + 0.1^2}} = \frac{b_{d,1} - y_b}{0.3162}, \qquad \mathbf{J}_a = \mathbf{e}_8^{\mathsf{T}}/0.3162, \tag{7.27}$$

with y_b the logged reading less the reference, in K, for one pair, and 0.1000 K in place of 0.3162 K for the mean of ten pairs whose errors are independent. It is the one row that sees the sensor and not the coil. The other candidates read the plant: a reference thermometer of 0.1 K at unit 1's discharge in the window, which reads the air, $T_{d,1}$, and not the sensor; a second permanent discharge sensor on unit 1, of 0.3 K, the permanent instrument at the place of the question; a second logged hour of today's five readings, a new load entry with five fresh rows on the same coils and the same drift, formed on the nine-entry state; water-side heat meters on coil 1 or coil 2, of $\sigma_W = 3.0$ kW; a reference return-air sensor and a reference discharge sensor on unit 2, each of 0.1 K; and a clamp-on check of the UPS output, of 1.0 kW, which reads the load. The last line of the comparison is not a measurement: the drift known exactly, the limit of a perfect side-by-side check.

7.6.2 The candidates ranked

Table 7.14 ranks the candidates by the width they take off the drift. Each row is a measurement in today's window; the columns give the standard deviation of the drift after it by equation (7.23) and the share of the present 0.5150 K it removes, the same for $\ln h_{1,25}$ and its present 0.05950, then the separation of equation (7.25) and the expected change of $\ln B$ in favor of the true

Table 7.14: Candidate measurements in today's window, from the posterior of M_{both}; the widths removed are taken from the present $\text{sd}(b_{d,1}) = 0.5150\,\text{K}$ and $\text{sd}(\ln h_{1,25}) = 0.05950$, and the expected changes are those of $\ln B(M_{\text{foul}} : M_{\text{drift}})$ toward the true drawing.

Candidate (today's window)	$\text{sd}_a(b)$ [K]	b [%]	$\text{sd}_a(\ln h_1)$ [–]	$\ln h_1$ [%]	ν_a [–]	U, foul true [–]	U, drift true [–]
Side-by-side check, ten pairs	0.0982	80.9	0.04622	22.3	2.21	2.237	28.306
Side-by-side check, one pair	0.2695	47.7	0.04982	16.3	1.56	1.012	2.497
Reference thermometer at unit 1's discharge, 0.1 K	0.3010	41.5	0.03896	34.5	1.89	1.593	2.058
Second permanent discharge sensor on unit 1, 0.3 K	0.3689	28.4	0.04522	24.0	1.37	0.881	1.007
Second logged hour of today's five readings	0.4242	17.6	0.05299	10.9	–	–	–
Water-side heat meter on coil 1, 3.0 kW, permanent	0.4251	17.4	0.04879	18.0	1.02	0.523	0.525
Reference return-air sensor, 0.1 K	0.4465	13.3	0.05516	7.3	0.75	0.270	0.308
Water-side heat meter on coil 2, 3.0 kW	0.4563	11.4	0.05353	10.0	0.75	0.280	0.290
Reference discharge sensor on unit 2, 0.1 K	0.4948	3.9	0.05599	5.9	0.46	0.110	0.106
Clamp-on check of the UPS output, 1.0 kW	0.5045	2.0	0.05756	3.3	0.34	0.058	0.057
The drift known exactly (the floor)	0	100	0.04564	–	–	–	–

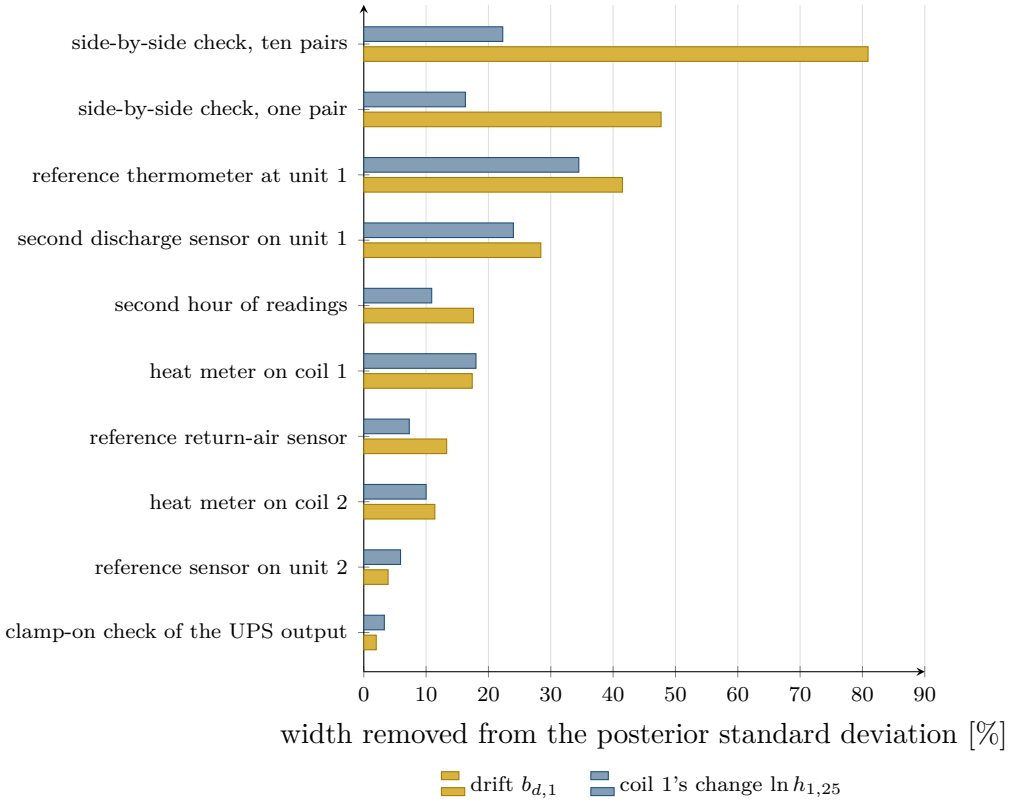


Figure 7.6: The candidate measurements of Table 7.14, ranked by the width they remove from the drift's standard deviation, with the width removed from coil 1's change beside each, by the preposterior (7.23) on M_{both}.

drawing if M_{foul} is true and if M_{drift} is. The widths removed are compared in Figure 7.6, two bars per candidate, ranked by the drift's.

The ranking follows each candidate's row, its line in the columns of $\ln h_{1,25}$ and $b_{d,1}$. The side-by-side check is a line on $b_{d,1}$ alone, so ten pairs take the drift's standard deviation to 0.0982 K and, through the ridge, take 22.3 % off coil 1's change as well. It also separates the drawings most. Under M_{foul} the ten pairs are predicted to average 0.000 with spread 0.100 K, and under M_{drift} 0.660 with spread 0.410 K; the expected change of $\ln B$ is 2.237 if the coil fouled and 28.306 if the sensor drifted, the narrow prediction of M_{foul} meeting a wide one. Carried over the whole family by equation (7.26), with the outcome drawn from M_{both}'s predictive, the ten pairs decide a drift in 0.680 of the outcomes and a change of coil 1 in only 0.007, with a median probability of 0.390 that coil 1 changed; at the true drift of 0.30 K they decide a drift in 0.487 of the outcomes and a change of coil 1 in none, and the median probability that coil 1 changed is 0.445. The check settles the sensor, not the coil: it pins the drift, and what is left of the shift falls to a coil's change that the window still reads only to 0.04622, so the coil's question stays near even. The reference thermometer at unit 1's discharge reads the air, with line $-9.4866/0.1$ on $\ln h_{1,25}$ and zero on $b_{d,1}$, as the logged sensor's line is with the drift removed; it is the best single row for the coil's index, taking 34.5 % off, and the drift follows through the discharge reading the two now share. A second permanent discharge sensor is the same row at 0.3 K and takes 28.4 and 24.0 % off. The heat meter on coil 1 sees the split of the heat between the coils: its gradient has the entries $+36.0104$ and -36.0104 kW per unit on $\ln h_{1,25}$ and $\ln h_{2,25}$, the same on $g_{1,0}$ and $g_{2,0}$, and 0.5189 on Q_{25}, as in Chapter 5, and it reads that split weakly at 3 kW. The second logged hour repeats the drifting sensor, so it narrows the reading errors and not the ridge. A reference return-air sensor separates the two drawings by 0.75 of a predictive standard deviation, the gap of 26.9009 against 26.7356°C seen through its spread. With the drift known exactly, coil 1's change has standard deviation 0.04564 and no lower: the commissioning floor of Chapter 5, Section 5.6.3, lies under it.

7.6.3 The recommendation

The side-by-side check of unit 1's discharge sensor, ten pairs, first: it needs a portable reference thermometer and an hour's work, it takes 80.9 % off the drift's band, and it moves the comparison furthest toward whichever of M_{foul} and M_{drift} is true, by 2.237 in $\ln B$ on average if the coil fouled and by 28.306 if the sensor drifted. It settles the sensor and not the coil. For the coil's index, the reference reading of the discharge air in the same visit takes 34.5 % off, and both go in the log as the next measurement. Of the permanent instruments, a second discharge sensor on unit 1 is worth more for this question than the heat meter on coil 1, 28.4 against 17.4 % off the drift; what either is worth across later windows is Chapter 23's question, in Volume II. The ten pairs' errors are taken as independent of each other and of the logged reading's. Errors correlated from one pair to the next would make the ten pairs worth fewer independent readings, as Chapter 24 (in Volume II) shows for a logged channel; the pairs' own correlation, in which the air's departure cancels and only the two sensors' errors remain, is not modeled in this book.

7.6.4 What the analysis showed

The dashboard said coil 1 had lost 34 %, and Chapter 5's method, with the sensor taken as sound, said 0.91 with a loss demonstrated at 0.96. With the sensor allowed to drift, the record cannot decide between a fouled coil and a drifted sensor, and coil 1's index of 0.96 does not demonstrate a loss. The physics takes the load, the season and coil 2 out of the rise, half of it, and leaves a shift of unit 1's discharge reading that the record knows well. The evidence says that the shift has two explanations of nearly equal weight, because the readings that could tell them apart see the difference by less than their own error. Fresh records at the same truth show that one window would decide only for a drift of about two kelvin, so this is a property of

the window, not of this record's draws. A reference instrument beside the sensor, not a longer look at the same sensors, separates a drift from fouling, which Chapter 3 said of a shared offset.

7.6.5 Limits of the model

A constructed record. The condition record is generated from a stated truth, with the stated sensor errors, and step 5 compares the answers with that truth (Section 7.4.6). It is not field data, and the agreement is a check of the method on its own model, not of the model against a real hall.

One suspect sensor and one suspect coil. The family frees only unit 1's discharge sensor and coil 1. A drift of the return-air sensor or of unit 2's discharge sensor lies outside it, and a diagnosis over every sensor, as in *Operations and Control by Derivation*, Chapter 7, would be the larger family.

A constant bias, calibrated at commissioning. The bias is one value in the window, and a drift that grows from window to window would need trend rows on the bias. The sensors are taken as calibrated at the commissioning test; a sensor replaced or recalibrated between the windows makes $b_{d,1}$ the change across that event. A standing offset is left out of the rows; it is not wholly harmless, because the rating prior holds the commissioning entries and not today's: under Chapter 5's drawing an offset of $+1\,\mathrm{K}$ in both windows moves $\ln h_{1,25}$ by -0.0118 at equal inputs and by $+0.0103$ at the recorded windows, against -0.0995 for a drift of the same size.

Proper priors and the Laplace evidence. The evidence moves with the widths of the freed terms' priors, as Section 7.4.4 shows, and it is the Laplace evidence, with importance sampling beside it.

Chapter 5's limits. Steady state within the window, pinned airflow with no discharge control, known inputs and a shared β_1 unchanged since handover are that chapter's, and they hold here unchanged.

The next chapter. Chapter 8 meets two entries that one record carries in a fixed combination again, in a battery's capacity and resistance, whose correlation is read before either is reported.

7.7 What to retain

Definition 7.1 (Degradation against sensor drift)**.** A rise that an asset's own sensor reports is read as a structure question: drawings of the plant that free a derate term on the asset, a bias term on the sensor, both or neither, nested in the drawing that frees both. Each drawing's evidence is its Laplace evidence, equation (7.7); between nested drawings the log Bayes factor is the gain in fit less the Occam factor, equation (7.8); the drawings' posterior probabilities follow from the evidences and their prior probabilities, and a quantity asked over the family is their mixture. The two terms are told apart only by the rows that see one of them and not the other; the shared reading fixes their combination, and their joint posterior is a ridge along it.

Three conclusions carry beyond the hall. A drift of a sensor and a loss of the asset it reads are confounded in the reading they share, and a fit that takes the sensor as sound reports the drift as a loss with a confident band. The residual test cannot catch a drift on a reading whose entry is free, because that entry absorbs the reading; the comparison of drawings asks the right question and may still answer that the record cannot tell. And what separates the two is a row

that sees one of them alone, a reference beside the sensor, which the preposterior values before it is taken.

Table 7.15 lists the terms of the chapter with their expressions, units and the subsection that defines each.

Table 7.15: The terms of the comparison of degradation and sensor drift, and where each is defined. Units for the two-unit hall.

Term	Expression	Units	Meaning	Defined in
Sensor drift	$b_{d,1}$, the offset's change since commissioning, eq. (7.2)	K	a bias entry in the later window only	7.1.1
Candidate family	Table 7.2	–	nested drawings at the place of the question	7.1.2
Laplace evidence	eq. (7.7)	–	height of the fit times the posterior's volume	7.2.3
Gain in fit, Occam factor	eq. (7.8)	–	the two parts of a log Bayes factor	7.2.3
Bayes factor, drawing probabilities	eq. (7.13), eq. (7.9)	–	the comparison and the ranking	7.3.1
Ridge of a bias and a derate	eq. (7.16); $s_{d,1}$, eq. (7.15)	K, –	the shared reading fixes their combination	7.3.3
Model averaging	eq. (7.10), eq. (7.11)	–	mixture over the drawings; within and between	7.2.4
Residual test, redundancy number	eq. (7.19), r_k^{N} in eq. (7.20)	–	a free entry absorbs its reading	7.3.5
Savage–Dickey ratio	eq. (7.21)	–	the Bayes factor of a nested pin	7.4.2
Closed form of a bias	eq. (7.22), ψ_k	–	the evidence from the normalized residual	7.4.2
Side-by-side calibration check	eq. (7.27)	K	a row on the bias alone	7.6.1
Expected change of the comparison	eq. (7.25)	–	the value of a check before it is made	7.6.1

7.8 Exercises

Exercise 7.1. *Formulation.* Write the question "has the sensor drifted by more than 0.5 K?" in the form of Table 7.3, under M_{both} and over the family, and evaluate it. (A prediction under a threshold, equation (2.12) with $f = b_{d,1}$ and $f^* = 0.5\,\mathrm{K}$: under M_{both}, $\Phi\big((0.4623-0.5)/0.5150\big) = 0.4708$; over the family, by the mixture (7.10), the drawings that pin the drift add nothing, and $0.3103 \times 0.6566 + 0.2186 \times 0.4708 = 0.3067$.)

Exercise 7.2. *Derivation.* Derive the Savage–Dickey ratio (7.21) for a nested entry with an independent prior, and evaluate $\ln B(M_{\mathrm{foul}} : M_{\mathrm{both}})$ from the drift's marginal under M_{both}. (By Bayes' rule in M_{big} at $z_i = 0$, $p(z_i = 0 \mid \mathbf{y}) = p(\mathbf{y} \mid z_i = 0)\, p(z_i = 0)/p(\mathbf{y} \mid M_{\mathrm{big}})$, and $p(\mathbf{y} \mid z_i = 0, M_{\mathrm{big}}) = p(\mathbf{y} \mid M_{\mathrm{small}})$ when the other entries' priors are the same in both, so the ratio of the posterior to the prior density at zero is $p(\mathbf{y} \mid M_{\mathrm{small}})/p(\mathbf{y} \mid M_{\mathrm{big}})$; with $b_{d,1} \sim \mathcal{N}(0.4623,\ 0.5150^2)$ against $\mathcal{N}(0,\ 1.0^2)$, $\ln B = +0.2608$, against the Laplace evidences' $+0.2485$.)

Exercise 7.3. *Derivation.* From the closed form (7.22), show that freeing the coil lowers the sensor row's redundancy number and with it the evidence for a drift, and evaluate $\ln B(M_{\mathrm{drift}} : M_{\mathrm{neither}})$. (Freeing an entry the row touches adds a column to $\mathbf{J}$, which raises the row's leverage K_{kk}, lowers $1 - K_{kk}$ and with it ψ_k, and shrinks the weight $\psi_k/(1+\psi_k)$ on $(r_k^{\mathrm{N}})^2$; in M_{neither} the row has redundancy 0.4746, so $\psi = 5.2731$, and with $r^{\mathrm{N}} = -1.8068$, $\ln B = +0.4539$.)

Exercise 7.4. *Application.* Split the 1.50 K under M_{foul} by equation (7.14), and say why its coil-1 term exceeds M_{both}'s. (M_{foul} pins the drift at zero, so the whole shift of the discharge reading goes to the coil, while M_{both} shares 0.8015 K between the coil's 0.3392 and the sensor's 0.4623; heat +0.6062, coil 2 +0.1200, coil 1 +0.6577 K.)

Exercise 7.5. *Application.* Ten side-by-side pairs average $y_b = 0.25$ K (standard deviation 0.1000 K). Refit the four drawings with the row $(b_{d,1} - y_b)/0.1$, with $b_{d,1} = 0$ in the drawings that pin it, and give the new probabilities. (The row adds the same term to every drawing's evidence where the drift is pinned and a new fit where it is free; M_{neither} 0.0807, M_{foul} 0.1185, M_{drift} 0.4330, M_{both} 0.3678, and $\ln B(M_{\text{foul}} : M_{\text{drift}}) = -1.2957$; under M_{both} the drift is 0.2578 K with standard deviation 0.0982 K.)

Chapter 8

Battery State of Health

8.1 The state-of-health tile and the string's circuit

The site's UPS string is forty valve-regulated lead-acid blocks of 12 V, rated 100 Ah, commissioned with the hall three years ago. Each capacity test discharges it at 40 A to the industry cutoff of 1.75 V per cell, 10.5 V per block or 420 V on the string, and then recharges it at 20 A. The year-3 test, window 36 (Dec Y3), ran 2.0535 h to the cutoff, against 2.4684 h at commissioning. Two readings sit on the dashboard. The state-of-health tile divides the run times, $2.0535/2.4684 = 0.8319$, and reports "capacity 83.2 Ah, lost 16.8 %"; it shows no resistance. The monitor's fit draws a line through each discharge log and reads the circuit from it with the maker's nominal curve rise of 1.55 V per block: it gives 82.34 Ah and 25.31 mΩ today, against 97.52 Ah and 8.08 mΩ at commissioning, "capacity 0.844 of commissioning, resistance up 213 %", with standard deviations of 0.34 Ah and 3.8 mΩ. The conclusions the two readings invite are that the string has lost a sixth of its capacity, that its resistance has tripled, and that it is near the end of its life.

Neither reading can say how much of its number is the maker's curve. The curve is known to ±0.08 V per block, 3.20 V on the string, twice the whole ohmic drop of 1.60 V that a 40 mΩ string shows at 40 A, and that drop is what the monitor's fit reads the resistance from. The fit never uses the recharge, whose ohmic term carries the opposite sign, and it cannot say whether its 82 Ah and 25 mΩ are two findings or one. The question of this chapter is

how a battery's state of health is read as the capacity and the resistance of its equivalent circuit from the discharge and recharge records, with their correlation read first, against the commissioning test, and what each record contributes.

It is answered with the method of Chapter 4, on a joint state over the two tests, as Chapter 5 read the coils.

Jargon: UPS string, capacity test, run time, cutoff voltage, equivalent circuit, state of charge, loaded curve, charge offset, coulomb counting, ohmic resistance, replacement criterion, recharge record, string monitor, confounded parameters

A *UPS string* is the chain of battery blocks in series behind an uninterruptible power supply. A *capacity test* discharges the string at a constant current from full charge to the *cutoff voltage*, below which the maker forbids discharge, and its *run time* is the time that takes. The *equivalent circuit* of the string is a voltage source that depends on the *state of charge*, the share of the usable charge still in the string, in series with a resistance. The *loaded curve* is the maker's source voltage of one block at the test rate as a function of the state of charge, and the *charge offset* the height of the charging curve above it. *Coulomb counting* follows the state of charge by integrating the current over time. The *ohmic resistance* lowers the terminal voltage below the curve while the string discharges and raises it above the curve while it charges. The *replacement criterion* is the share of capacity at which a string is replaced, 80 % by convention. The *recharge record* is the log of the recharge that follows a capacity test. A *string monitor* is a permanent instrument that logs the string's voltage and current. Two parameters are *confounded* when the readings fix a combination of them and not each one.

Key idea

A constant-current log is a straight line, two numbers. A discharge gives two numbers per test, and the circuit has three: the capacity $C_{\mathrm{bat},t}$, the resistance $R_{\mathrm{bat},t}$ and the curve rise Δv. On the discharge alone the maker's tolerance on the curve slides capacity and resistance together along a ridge, with a correlation of $+0.9993$. The recharge adds a second line, whose level carries the resistance with the opposite sign and no curve rise, and that breaks the ridge. The correlation that remains, $+0.84$, is the ordinary trade between a line's level and its slope.

8.1.1 State of health as two parameters of a circuit

Chapter 5 read one parameter per coil against commissioning, the coil's conductance. A battery has two, its usable capacity C_{bat} in Ah and its internal resistance R_{bat} in mΩ, and both are read from the same few numbers of a test record. A log taken at a constant current is close to a straight line in time, so a discharge log holds two numbers, a level and a slope, and those two numbers are all the discharge says about the circuit. They cannot be read one parameter at a time.

The pattern is the ridge of Chapter 1, equation (1.19). When the rows fix a combination of two entries and not the entries themselves, the posterior lies along the direction in which the combination stays fixed, the correlation of the two entries tends to ± 1, and each entry's band is the band of whatever holds the ridge in place, a prior or a pinned quantity. A battery's capacity and resistance meet the same pattern whenever the record measures fewer numbers than the circuit has parameters.

The state of health of the string in window t against its commissioning test, window 0, is the pair

$$h_C = \frac{C_{\mathrm{bat},t}}{C_{\mathrm{bat},0}} = \mathrm{e}^{\xi_t - \xi_0}, \qquad \ln h_C = \xi_t - \xi_0, \qquad \Delta R = R_{\mathrm{bat},t} - R_{\mathrm{bat},0}, \tag{8.1}$$

with $\xi_t = \ln(C_{\mathrm{bat},t}/\bar{C}_{\mathrm{bat}})$ the log-capacity against the rated $\bar{C}_{\mathrm{bat}} = 100\,\mathrm{Ah}$, the entry that carries the capacity. As in equation (5.1), the rating cancels from h_C, and both $\ln h_C$ and ΔR are changes since commissioning in the sense of Chapter 3: differences of two entries of one state, each an estimate. The ratio $R_{\mathrm{bat},t}/R_{\mathrm{bat},0}$ is the *ohmic state of health*. The resistance is carried linearly, in mΩ, because every row is linear in it.

Principle 8.1. *A battery's state of health is the posterior of its capacity and its resistance read together. Their correlation says which combinations the record measures, and it is read before either is reported. A record measures a parameter only if some row moves it in a direction the other parameters cannot copy.*

8.1.2 The UPS string in two tests (step 0)

The plant is the string alone; none of the hall's quantities enters it. Each test is one equivalent circuit of $N = 40$ blocks. The source is the maker's loaded curve, linear in the state of charge, $N(v_0 + \Delta v\,\mathrm{SOC})$, with v_0 a block's loaded voltage at the usable-empty point and Δv its rise to full; on charge the curve sits higher by the charge offset h_{ch} per block. The source is in series with the string's ohmic resistance $R_{\mathrm{bat},t}$, in mΩ, whose drop has the sign of the current. Coulomb counting carries the state of charge: the discharge at I_d starts full and the recharge at I_c starts where the discharge stopped, at the cutoff. With ϑ the time in hours into each log,

$$\begin{aligned}
\text{discharge:}\quad & V = N(v_0 + \Delta v\,\mathrm{SOC}) - \frac{I_d R_{\mathrm{bat},t}}{1000}, \qquad \mathrm{SOC} = 1 - \frac{I_d\,\vartheta}{C_{\mathrm{bat},t}},\\
\text{recharge:}\quad & V = N(v_0 + h_{\mathrm{ch}} + \Delta v\,\mathrm{SOC}) + \frac{I_c R_{\mathrm{bat},t}}{1000}, \qquad \mathrm{SOC} = \mathrm{SOC}_{\mathrm{cut},t} + \frac{I_c\,\vartheta}{C_{\mathrm{bat},t}},\\
& \mathrm{SOC}_{\mathrm{cut},t} = \frac{V_{\mathrm{cut}} + I_d R_{\mathrm{bat},t}/1000 - N v_0}{N\,\Delta v}, \qquad T_{\mathrm{run},t} = \frac{(1 - \mathrm{SOC}_{\mathrm{cut},t})\,C_{\mathrm{bat},t}}{I_d}.
\end{aligned} \tag{8.2}$$

The third line follows from the first two: the discharge reaches the cutoff V_{cut} at the state of charge where its voltage equals V_{cut}, and the state of charge falls from one to that value at the rate $I_d/C_{\mathrm{bat},t}$ per hour.

Pinned and free. Pinned from the maker's sheet are $v_0 = 10.52\,\mathrm{V}$ and $h_{\mathrm{ch}} = 1.40\,\mathrm{V}$ per block; pinned by the test procedure are $I_d = 40\,\mathrm{A}$, $I_c = 20\,\mathrm{A}$ and $V_{\mathrm{cut}} = 420\,\mathrm{V}$. The ratings $\bar{C}_{\mathrm{bat}} = 100\,\mathrm{Ah}$ and $\bar{R}_{\mathrm{bat}} = 40\,\mathrm{m\Omega}$ are the origin of the capacity entry and the starting point of the fit, not rows. Free in each test are $C_{\mathrm{bat},t}$ and $R_{\mathrm{bat},t}$. Free and shared by both tests is the curve rise Δv, which the maker states as $1.55\,\mathrm{V}$ per block to a tolerance of $0.08\,\mathrm{V}$.

The four lines of a test. In elapsed time each log of equation (8.2) is a straight line. The discharge falls from its level $V_{d,t}$ at the slope $\dot{V}_{d,t}$, and the recharge rises from its level $V_{c,t}$ at the slope $\dot{V}_{c,t}$:

$$\begin{aligned}
V &= V_{d,t} - \dot{V}_{d,t}\,\vartheta, \qquad V_{d,t} = N(v_0 + \Delta v) - \frac{I_d R_{\mathrm{bat},t}}{1000}, \qquad \dot{V}_{d,t} = \frac{N\,\Delta v\,I_d}{C_{\mathrm{bat},t}},\\
V &= V_{c,t} + \dot{V}_{c,t}\,\vartheta, \qquad V_{c,t} = N h_{\mathrm{ch}} + V_{\mathrm{cut}} + \frac{(I_d + I_c)\,R_{\mathrm{bat},t}}{1000}, \qquad \dot{V}_{c,t} = \frac{N\,\Delta v\,I_c}{C_{\mathrm{bat},t}}.
\end{aligned} \tag{8.3}$$

The recharge level is free of Δv. At $\vartheta = 0$ the recharge voltage is $N(v_0 + h_{\mathrm{ch}}) + N\,\Delta v\,\mathrm{SOC}_{\mathrm{cut},t} + I_c R_{\mathrm{bat},t}/1000$, and by the third line of equation (8.2) the middle term is $V_{\mathrm{cut}} + I_d R_{\mathrm{bat},t}/1000 - N v_0$: the recharge starts from the state of charge at which the discharge stopped, and that state already carries the discharge's ohmic drop. The curve rise cancels and the two currents' drops add. Each test's record is therefore four numbers. The discharge level carries $R_{\mathrm{bat},t}$ with the sign of $-I_d$ and Δv with the factor N; the recharge level carries $R_{\mathrm{bat},t}$ with the sign of $+(I_d + I_c)$ and no Δv; the two slopes carry the ratio $\Delta v/C_{\mathrm{bat},t}$, in the proportion I_c/I_d of the currents. The least-squares slopes of the logs stand in the ratio 0.4969 at commissioning and 0.4995 today, against $I_c/I_d = 0.5000$, which supports one curve rise shared by the discharge and the recharge.

Modeling choices at the graph. Five choices that no question owns make this graph. Each test is taken as one steady test at constant current and temperature: this makes every parameter a constant of the test, and it fails if the string warms during the discharge, which the record cannot show because temperature is not logged. The loaded curve is taken as linear in the state of charge, with the polarization at the test rates lumped into it: this makes each log a straight line of four numbers, and it fails near the knee at the end of a deep discharge and at another current, which the 80 A candidate of Section 8.6.1 inherits. One ohmic resistance serves charge and discharge: this is what lets the recharge read the resistance the discharge shows, and it fails if the two differ. The charge efficiency is taken as one, with the recharge logged only on its constant-current stage: this makes the recharge's coulomb count the discharge's run backward, and it fails on the constant-voltage tail of a charge, which the record leaves out. And v_0 and h_{ch} are pinned from the maker's sheet: this is what the recharge's answer leans on, and Section 8.4.3 gives the move of every answer per unit error of each.

The record. Table 8.1 gives the two tests the chapter reads. Each row is a quantity of the record, for the commissioning test and for today's. Every voltage is logged every 10 minutes by a voltmeter of standard deviation 0.25 V on the string, and the run time is the instant the discharge reaches 420 V. The string keeps its own maintenance log, which starts at its own commissioning test, window 0.

Table 8.1: The two capacity tests of the UPS string read in this chapter: the commissioning test and today's.

Quantity [unit]	Window 0 (commissioning)	Window 36 (Dec Y3)
Discharge at $I_d = 40$ A: logged voltages [–]	14	12
logged from 10 minutes to [min]	140	120
first and last voltage [V]	477.99, 423.07	476.62, 421.55
Run time to the cutoff of 420 V [h]	2.4684	2.0535
the same [min]	148.1	123.2
Charge delivered, $I_d T_{\mathrm{run},t}$ [Ah]	98.73	82.14
Recharge at $I_c = 20$ A: logged voltages [–]	29	24
logged from 10 minutes to [min]	290	240
first and last voltage [V]	480.42, 539.55	481.26, 539.98
Voltmeter standard deviation on the string [V]	0.25	0.25

The graph. The circuits of the two tests are drawn side by side in Figure 8.1, window 0 on the left and window 36 on the right, the second mirrored so that the two loaded curves face each other. In each copy the currents and the free capacity (green) feed the state of charge by coulomb counting; the state of charge sets the loaded curve, with the pinned v_0 and h_{ch} (gray); the curve and the free resistance (green) set the terminal voltage; and the dashed arrows leave the terminals for the test's three readings. One green node between the copies, the curve rise Δv, is linked to the loaded curve of both tests, and the maker's tolerance is attached to it.

The joint state. With the state of charge a function of the free entries through coulomb counting, the fit runs on the reduced state of Section 4.2.8, over both tests,

$$\mathbf{z} = \left(\xi_0,\ R_{\mathrm{bat},0},\ \xi_{36},\ R_{\mathrm{bat},36},\ \Delta v\right)^{\mathsf{T}}, \tag{8.4}$$

the log-capacities dimensionless, the resistances in mΩ and the curve rise in V per block; z_1 to z_5 in that order. The full state adds the state of charge at each of the 79 logged times, 84 entries in all, under 79 coulomb-counting rows, the second column of equation (8.2) for each logged time; Section 8.4.2 fits it and lets the rows become exact.

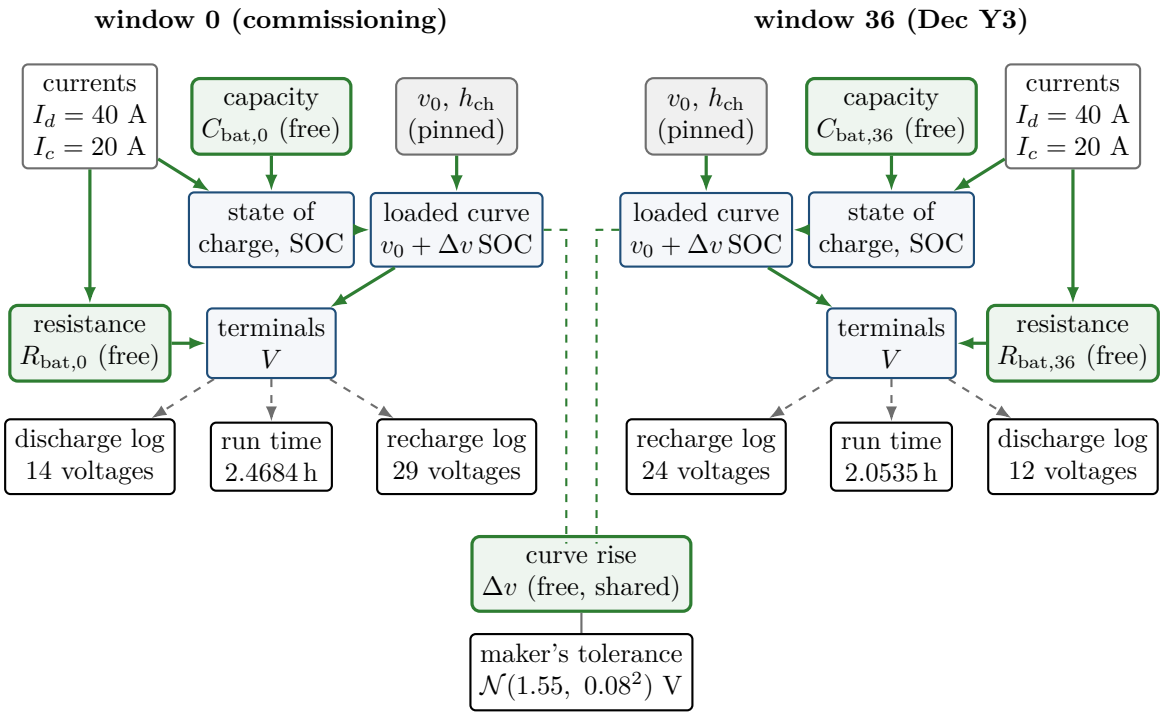


Figure 8.1: The equivalent circuits of the UPS string in the commissioning test (left) and today's test (right, mirrored). Free (green): each test's capacity $C_{\mathrm{bat},t}$ and resistance $R_{\mathrm{bat},t}$, and the curve rise Δv, one entry shared by both tests (dashed green links). Blue: the state of charge, the loaded curve of one block and the terminal voltage of the N blocks, fixed by the rows of equation (8.2). Pinned (gray): v_0 and h_{ch} from the maker's sheet; also pinned, not drawn, $N = 40$ and the cutoff $V_{\mathrm{cut}} = 420\,\mathrm{V}$. Sources (white): the test currents. Observed (dashed): each test's discharge log, run time and recharge log. Prior (solid gray): the maker's tolerance on Δv.

8.1.3 The asset manager's questions (step 1)

The asset manager asks four questions of the two tests, and step 1 writes each as a probability over the joint state (8.4), with the unknown on the left of the bar and the two records and the maker's tolerance on the right. Table 8.2 lists them. The first column is the question in the asset manager's words, the first of them the drill question; the second the general form of its category from Chapter 2, for a quantity $f(\mathbf{z})$ derived from the state and a level f^* of it; the third its mathematical form for the string; the fourth its category; the fifth the computation of Section 2.3.6 that answers it; and the last the subsection where it is answered. The drill question asks for both parameters from the charge and the discharge records, and both records are what step 2 reads; the same answer from the discharge alone stands beside it as the comparison. Which reading would narrow the answers further asks nothing new of today's evidence; it is the next measurement, Section 8.6. A rate of capacity fade needs a third test and is not asked.

Table 8.2: The questions asked of the UPS string at window 36 against its commissioning test. $\mathbf{y}$ is the two tests' discharge and recharge logs, their run times and the maker's tolerance on Δv; h_C and ΔR are the state of health (8.1).

Asset manager's question	General form	Mathematical form	Category	Computation	Answered in
What is the usable capacity now, and the internal resistance?	$\Pr(f(\mathbf{z}) \le f^* \mid \mathbf{y})$; moments; credible interval; the covariance of two entries, (2.9)–(2.11)	$p(C_{\text{bat},36}, R_{\text{bat},36} \mid \mathbf{y})$: the correlation and the eigen-directions first, then each median and 5–95 % band; the discharge alone beside it	estimation	posterior	8.3.1
How much capacity has the string lost since commissioning, and how far has its resistance risen?	moments; credible interval, (2.10), (2.11); $\Pr(f(\mathbf{z}) > f^* \mid \mathbf{y})$, (2.12)	$h_C = \mathrm{e}^{\xi_{36} - \xi_0}$ and $\Delta R = R_{\text{bat},36} - R_{\text{bat},0}$: median and band; $\Pr(\Delta R > 0 \mid \mathbf{y})$	estimation; prediction under a threshold for $\Pr(\Delta R > 0)$	posterior	8.3.2
Is the dashboard's 83.2 % the capacity?	moments and credible interval of a derived quantity, (2.10), (2.11)	$p(T_{\text{run},36}/T_{\text{run},0} \mid \mathbf{y})$ and its split into $\ln h_C$ and a resistance term	estimation	posterior	8.3.3
Is the string still above the replacement criterion of 80 %?	$\Pr(f(\mathbf{z}) < f^* \mid \mathbf{y})$, (2.12)	$\Pr(h_C < 0.80 \mid \mathbf{y})$ against commissioning; $\Pr(C_{\text{bat},36} < 80\,\text{Ah} \mid \mathbf{y})$ against rating	prediction under a threshold	posterior	8.3.4

8.1.4 The evidence as rows (step 2)

Step 2 attaches the evidence of Table 8.1 to the graph as rows on the joint state, 82 in all. Each logged voltage is one row, $14 + 12$ of the discharges and $29 + 24$ of the recharges, all with the voltmeter's standard deviation $\sigma_V = 0.25\,\text{V}$. Each test adds one run-time row, with the standard deviation $\sigma_{\text{run},t} = \sigma_V / \dot{V}_{d,t}$ taken at the slope of the logged discharge line: the cutoff fires on the same voltmeter, so a reading error of σ_V at the cutoff moves the instant by σ_V over the rate at which the voltage falls. That is 0.00983 h (0.59 minutes) at commissioning and 0.00830 h (0.50 minutes) today. The maker's tolerance is a prior row, $\Delta v \sim \mathcal{N}(1.55,\ 0.08^2)$. The capacities and the resistances carry no prior row.

Which row acts on which entry. By equation (8.2) every row of test t depends on $C_{\mathrm{bat},t}$, $R_{\mathrm{bat},t}$ and Δv, and on no entry of the other test. With $*$ for an entry that can be nonzero, the rows grouped by kind and the columns in the order of equation (8.4), the 82×5 Jacobian has the pattern

$$\mathbf{J} = \begin{bmatrix} * & * & 0 & 0 & * \\ * & * & 0 & 0 & * \\ * & * & 0 & 0 & * \\ 0 & 0 & * & * & * \\ 0 & 0 & * & * & * \\ 0 & 0 & * & * & * \\ 0 & 0 & 0 & 0 & * \end{bmatrix} \begin{array}{c} \text{window 0: discharge voltages (14 rows)} \\ \text{run time (1 row)} \\ \text{recharge voltages (29 rows)} \\ \text{window 36: discharge voltages (12 rows),} \\ \text{run time (1 row)} \\ \text{recharge voltages (24 rows)} \\ \text{maker's tolerance on } \Delta v \text{ (1 row)} \end{array} \tag{8.5}$$

$$\mathbf{g}_{\ln h_C}^{\mathsf{T}} = (-1,\ 0,\ 1,\ 0,\ 0), \qquad \mathbf{g}_{\Delta R}^{\mathsf{T}} = (0,\ -1,\ 0,\ 1,\ 0).$$

The second line is not a row: it gives the gradients of the second question's quantities, from equation (8.1). Three facts of the graph are read here, before any number.

The two tests meet only in the column of Δv, as the two windows of Chapter 5 met only in the column of β_1, and each gradient of the second line reaches across that block structure, one entry in each test's block.

On the discharge rows alone the five columns are not independent. Along the direction

$$\mathbf{n} = \Big(\frac{1}{\Delta v},\ \frac{1000N}{I_d},\ \frac{1}{\Delta v},\ \frac{1000N}{I_d},\ 1\Big)^{\mathsf{T}} \tag{8.6}$$

per volt of curve rise, the discharge level of equation (8.3) changes by $N-(I_d/1000)(1000N/I_d) = 0$ and the logarithm of the discharge slope by $1/\Delta v - 1/\Delta v = 0$, in both tests at once; the run time, $(V_{d,t} - V_{\mathrm{cut}})/\dot{V}_{d,t}$ by equation (8.2), is unchanged with them. Every discharge row and every run-time row is a function of the two discharge lines, so all of them are blind to $\mathbf{n}$, and only the maker's tolerance sees it: the discharge rows have a null direction, a ridge along which a larger curve rise goes with a larger capacity and a larger resistance in both tests. At the maker's nominal $\Delta v = 1.55\,\mathrm{V}$ it is $(0.64516,\ 1000.0,\ 0.64516,\ 1000.0,\ 1)$, a capacity $1/\Delta v$ larger in its logarithm and a resistance $1000N/I_d = 1000.0\,\mathrm{m\Omega}$ larger per volt.

The recharge rows see $\mathbf{n}$. The recharge slope is the discharge slope times I_c/I_d and is blind to it too, but the recharge level changes along $\mathbf{n}$ by $(I_d+I_c)/1000 \times 1000N/I_d = N(I_d+I_c)/I_d = 60.0\,\mathrm{V}$ per volt, with the sign opposite to the discharge level's resistance term. In the column of $R_{\mathrm{bat},t}$ a normalized discharge row has the entry $-I_d/(1000\sigma_V) = -0.1600$ per $\mathrm{m\Omega}$ and a normalized recharge row the entry $+(I_d + I_c)/(1000\sigma_V) = +0.2400$ per $\mathrm{m\Omega}$; the signs differ, and that is the whole of what the recharge adds.

Modeling choices at the rows. Six choices make these rows. Each reading is taken as Gaussian, unbiased and independent of the others given the state, the reading rows of Chapter 4: this makes each logged voltage one normalized row, and it fails if the voltmeter drifted between the two tests, the class of fault Chapter 7 takes up for a sensor of the hall. The run time is taken as a reading of the same voltmeter, with the standard deviation $\sigma_V/\dot{V}_{d,t}$: this adds the cutoff's instant as one more row on the discharge line, and it fails if the cutoff fires on a different instrument or after a delay. One Δv is shared by both tests and by the charge and the discharge, so the curve does not age: this lets the record measure the rise once from both tests, and it fails if the blocks' chemistry changes the curve's shape with age. The recharge is taken to start at the cutoff state of charge, with no rest and no charge lost between the two logs: this is what makes $V_{c,t}$ free of Δv in equation (8.3), and it fails if the string rests on open circuit or self-discharges before the recharge. No prior is placed on the capacity or the resistance, because the ratings carry no stated tolerance: the readings alone fix both. And the rows carry no trend

rows: Chapter 4's trend rows (4.68) would tie the two tests through a rate, and two tests fix no rate. Section 8.4.3 gives the effect of each choice that can be varied on every answer.

8.2 The posterior of the state (step 3)

All four questions of Table 8.2 are read from one posterior, that of the joint state (8.4) under the 82 rows of step 2, with the pinned quantities of Section 8.1.2.

8.2.1 The rows on the joint state

Write $V_{\mathrm{d}}(\vartheta;\mathbf{z})$ and $V_{\mathrm{c}}(\vartheta;\mathbf{z})$ for the discharge and recharge voltages of equation (8.2) at the entries of test t, with $C_{\mathrm{bat},t}=\bar{C}_{\mathrm{bat}}\,\mathrm{e}^{\xi_t}$, and $y_{\mathrm{d},t,j}$, $y_{\mathrm{c},t,j}$ and $y_{\mathrm{run},t}$ for the logged voltages at the times ϑ_j and the logged run time. Each test gives one row per reading, each divided by its standard deviation,

$$\mathbf{r}_t(\mathbf{z})=\begin{bmatrix}(V_{\mathrm{d}}(\vartheta_j;\mathbf{z})-y_{\mathrm{d},t,j})/\sigma_V, & j=1,\dots,n_{\mathrm{d},t}\\ (T_{\mathrm{run},t}(\mathbf{z})-y_{\mathrm{run},t})/\sigma_{\mathrm{run},t} & \\ (V_{\mathrm{c}}(\vartheta_j;\mathbf{z})-y_{\mathrm{c},t,j})/\sigma_V, & j=1,\dots,n_{\mathrm{c},t}\end{bmatrix},\qquad t=0,36, \tag{8.7}$$

with $n_{\mathrm{d},t}$ and $n_{\mathrm{c},t}$ the numbers of logged voltages of Table 8.1, and the residual vector stacks the two tests and the prior in the order of the pattern (8.5),

$$\mathbf{r}(\mathbf{z})=\begin{bmatrix}\mathbf{r}_0(\mathbf{z})\\ \mathbf{r}_{36}(\mathbf{z})\\ (\Delta v-1.55)/0.08\end{bmatrix},\qquad \mathcal{L}(\mathbf{z})=\tfrac{1}{2}\,\mathbf{r}(\mathbf{z})^{\mathsf{T}}\mathbf{r}(\mathbf{z}), \tag{8.8}$$

82 dimensionless rows on five entries. The rows are nonlinear in the state through e^{ξ_t} and the product $\Delta v\,\mathrm{SOC}$, and their Jacobian lines follow from equations (8.2) and (8.3). In the columns of test t, $(\xi_t, R_{\mathrm{bat},t}, \Delta v)$, a discharge row, the run-time row and a recharge row have the lines

$$\begin{aligned}\text{discharge:}\quad & \frac{1}{\sigma_V}\Big(\frac{N\,\Delta v\,I_d\,\vartheta_j}{C_{\mathrm{bat},t}},\ -\frac{I_d}{1000},\ N\,\mathrm{SOC}_j\Big),\\ \text{run time:}\quad & \frac{1}{\sigma_{\mathrm{run},t}}\Big(T_{\mathrm{run},t},\ -\frac{C_{\mathrm{bat},t}}{1000\,N\,\Delta v},\ \frac{C_{\mathrm{bat},t}\,\mathrm{SOC}_{\mathrm{cut},t}}{I_d\,\Delta v}\Big),\\ \text{recharge:}\quad & \frac{1}{\sigma_V}\Big(-\frac{N\,\Delta v\,I_c\,\vartheta_j}{C_{\mathrm{bat},t}},\ \frac{I_d+I_c}{1000},\ N\big(\mathrm{SOC}_j-\mathrm{SOC}_{\mathrm{cut},t}\big)\Big),\end{aligned} \tag{8.9}$$

with SOC_j the state of charge at the logged time ϑ_j, and zeros in the other test's columns; the prior row is $1/0.08=12.5$ in the column of Δv. The capacity column grows with the charge moved since the start of the log, the resistance column is constant within each log and changes sign between the two, and the curve-rise column is the charge in the string on the discharge and the charge put back on the recharge.

8.2.2 The normal equations and the covariance

The MAP state $\hat{\mathbf{z}}$ minimizes the objective (8.8). Each Gauss–Newton iteration solves the normal equations of equation (4.29),

$$(\mathbf{J}^{\mathsf{T}}\mathbf{J})\,\Delta\mathbf{z}=-\mathbf{J}^{\mathsf{T}}\mathbf{r}(\mathbf{z}),\qquad \mathbf{z}^{(l+1)}=\mathbf{z}^{(l)}+\Delta\mathbf{z}, \tag{8.10}$$

with $\mathbf{J}$ and $\mathbf{r}$ at $\mathbf{z}^{(l)}$. The iteration starts from the ratings and the maker's curve, $\mathbf{z}^{(0)}=(0,\ 40,\ 0,\ 40,\ 1.55)^{\mathsf{T}}$, and it stops when the gradient $\mathbf{J}^{\mathsf{T}}\mathbf{r}$ is below 10^{-4}, not when the step is small. The same normal equations fit the comparison route, the joint state on the discharge and

run-time rows alone, without the recharge and with the same start; along its ridge (8.6) a step criterion never settles. With every row normalized the information matrix is $\mathbf{\Lambda} = \mathbf{J}^{\mathsf{T}}\mathbf{J}$; with unnormalized rows and their weights it is $\mathbf{J}^{\mathsf{T}}\mathbf{W}\mathbf{J}$. At the minimum the Laplace covariance is

$$\mathbf{\Sigma} = \left(\mathbf{J}^{\mathsf{T}}\mathbf{J}\right)^{-1}\big|_{\hat{\mathbf{z}}}, \tag{8.11}$$

the Gauss–Newton form of equation (4.38), which drops the curvature term $\mathbf{C}$ of equation (4.44); Section 8.4.2 sets the full Hessian beside it. By equation (4.48) this $\mathbf{\Sigma}$ is the free-entry block of the 84-entry full state's covariance in the limit of exact coulomb counting.

The block structure. The pattern (8.5) fixes the shape of $\mathbf{\Lambda}$ before any number. A product of two columns of $\mathbf{J}$ is zero unless some row touches both, and only the column of Δv is touched by the rows of both tests, so

$$\mathbf{\Lambda} = \begin{bmatrix} \mathbf{\Lambda}_{0,0} & \mathbf{0} & \mathbf{\Lambda}_{0,v} \\ \mathbf{0} & \mathbf{\Lambda}_{36,36} & \mathbf{\Lambda}_{36,v} \\ \mathbf{\Lambda}_{0,v}^{\mathsf{T}} & \mathbf{\Lambda}_{36,v}^{\mathsf{T}} & \Lambda_{vv} \end{bmatrix}, \tag{8.12}$$

with $\mathbf{\Lambda}_{t,t}$ the 2×2 block of test t's entries $(\xi_t, R_{\mathrm{bat},t})$, $\mathbf{\Lambda}_{t,v}$ their coupling to Δv, and Λ_{vv} the information on Δv from both tests and the maker's tolerance. Given Δv the two tests are independent; with Δv integrated out each test's entries move with it, and the uncertainty of the curve rise both share leaves a covariance between the tests, the smaller the better Δv is known.

One state over two tests. Chapter 2's commissioning row carries an earlier fit's marginal into a later one, and it is the right tool when the earlier record is no longer at hand. Here both records are at hand, and the joint state makes the commissioning entries entries of the one posterior every answer is read from. A row carrying the commissioning fit forward reproduces that posterior only if it carries the full covariance of the entries the tests share: here ξ_0 and $R_{\mathrm{bat},0}$ with Δv. Marginal rows on ξ_0 and $R_{\mathrm{bat},0}$ alone, with the maker's tolerance stated again for the second test, drop what window 0 said about the curve rise, and so act as a second, unshared Δv. Section 8.4.4 sets both forms against the joint state.

8.2.3 The fit of the two tests

The iterations. Table 8.3 runs the normal equations (8.10) from $\mathbf{z}^{(0)}$. Each row is an iteration, with the objective and the norm of its gradient. The objective falls from 14447.403527 to 124.913929 in the first step, to 42.638350 in the second and to 42.638067 in the third, and the gradient is below 10^{-4} at the fourth.

Table 8.3: Gauss–Newton iterations of the joint state of the two tests, equation (8.10), from the rated capacity and resistance and the maker's curve rise.

Iteration l [–]	Objective $\mathcal{L}(\mathbf{z}^{(l)})$ [–]	$\lVert\mathbf{J}^{\mathsf{T}}\mathbf{r}\rVert$ [–]
0	14447.403527	1.375×10^{5}
1	124.913929	1.187×10^{4}
2	42.638350	2.245×10^{1}
3	42.638067	1.792×10^{-3}
4	42.638067	$< 10^{-4}$

The Jacobian at the MAP state. The lines of equation (8.9) at $\hat{\mathbf{z}}$, for the first and the last row of each group of window 0 and for the prior row, in the columns $(\xi_0, R_{\text{bat},0}, \Delta v)$, are

$$\begin{bmatrix} 16.8776 & -0.1600 & 149.3040 \\ 236.2860 & -0.1600 & 10.2554 \\ 250.5192 & -0.1607 & 1.9607 \\ -8.4388 & 0.2400 & 5.3480 \\ -244.7248 & 0.2400 & 155.0927 \\ 0 & 0 & 12.5000 \end{bmatrix} \begin{matrix} \text{discharge at 10 min} \\ \text{discharge at 140 min} \\ \text{run time} \\ \text{recharge at 10 min} \\ \text{recharge at 290 min} \\ \text{maker's tolerance} \end{matrix},$$

the capacity column dimensionless, the resistance column per mΩ and the curve-rise column per V; their entries in the columns of window 36 are zero. Window 36's rows have the same pattern in its own columns. The capacity column grows along each log with the charge moved; the resistance column is −0.1600 through the discharge and +0.2400 through the recharge; the curve-rise column is the charge left in the string on the discharge, $N\,\mathrm{SOC}_j/\sigma_V$, largest at its start, and the charge put back on the recharge, $N(\mathrm{SOC}_j - \mathrm{SOC}_{\text{cut},t})/\sigma_V$, largest at its end. The reading rows, all but the prior, have rank 5 of 5, with singular values 1084.2, 929.85, 624.09, 0.73716 and 0.67668. The discharge and run-time rows alone have rank 4 of 5: their singular values are 646.86, 577.88, 359.80 and 0.28044, and the fifth is below 10^{-6}, the null direction (8.6). Along that direction, normalized to unit length, the full record's readings carry the information 1.5082, against the maker's tolerance's 0.0001: the recharge level's slope of 60.0 V per volt on the ridge gives the readings a view the discharge rows did not have.

The information. The information matrix has the block form (8.12),

$$\mathbf{J}^{\mathsf{T}}\mathbf{J} = \begin{bmatrix} 961113.315 & -1204.817 & 0 & 0 & -285290.93 \\ -1204.817 & 2.055 & 0 & 0 & 379.312 \\ 0 & 0 & 815768.925 & -1012.187 & -227024.076 \\ 0 & 0 & -1012.187 & 1.715 & 308.622 \\ -285290.93 & 379.312 & -227024.076 & 308.622 & 652935.376 \end{bmatrix},$$

with the two test blocks on the diagonal, zeros between them, and the coupling to Δv in the last row and column. The resistance diagonals, 2.055 and 1.715 per $\mathrm{m}\Omega^2$, are the sums of 0.16^2 over the discharge rows and 0.24^2 over the recharge rows, with the run-time row's small share; the recharge rows carry most of them.

The covariance. Its inverse, the Laplace covariance (8.11), is, in the order of equation (8.4), with the log-capacities dimensionless, the resistances in mΩ and the curve rise in V,

$$\boldsymbol{\Sigma} = \begin{bmatrix} 4.0487 \times 10^{-6} & 2.2834 \times 10^{-3} & 1.0096 \times 10^{-7} & -2.8806 \times 10^{-5} & 4.9122 \times 10^{-7} \\ 2.2834 \times 10^{-3} & 1.8403 & -1.6288 \times 10^{-5} & 4.6471 \times 10^{-3} & -7.9248 \times 10^{-5} \\ 1.0096 \times 10^{-7} & -1.6288 \times 10^{-5} & 4.6615 \times 10^{-6} & 2.6772 \times 10^{-3} & 4.0892 \times 10^{-7} \\ -2.8806 \times 10^{-5} & 4.6471 \times 10^{-3} & 2.6772 \times 10^{-3} & 2.1839 & -1.1667 \times 10^{-4} \\ 4.9122 \times 10^{-7} & -7.9248 \times 10^{-5} & 4.0892 \times 10^{-7} & -1.1667 \times 10^{-4} & 1.9895 \times 10^{-6} \end{bmatrix}.$$

The correlations read by sign, and the rows fix each sign. A test's log-capacity and resistance are correlated by +0.8365 at commissioning and +0.8391 today. The resistance comes from the recharge level, and the recharge slope is one of the readings of the capacity. The level and the slope of a line fitted to one log are estimated with a negative correlation, because the log's readings lie after its start: a line through them with a higher level must run flatter to pass through their middle. A higher recharge level is a higher resistance, and a flatter recharge slope a larger capacity, so the two entries move together; the discharge pushes the same way, since a lower discharge level, a higher resistance, goes with a gentler fall, a larger capacity. The curve

rise is correlated with the log-capacities by $+0.1731$ and $+0.1343$ and with the resistances by -0.0414 and -0.0560: small, because Δv is fixed by the two levels together, the recharge level giving the resistance and the discharge level then the rise. Across the two tests every correlation is below 0.03 in size, 0.0232 between the log-capacities and 0.0023 between the resistances: with the curve rise measured, the two tests are nearly independent.

The MAP state. Table 8.4 gives the MAP state with the posterior spreads $\sqrt{\Sigma_{ii}}$ and the prior spread of each entry; the last column names the rows that inform it. The capacities are $C_{\mathrm{bat},0} = 99.725$ and $C_{\mathrm{bat},36} = 83.897\,\mathrm{Ah}$. The record measures the curve rise itself, to $0.00141\,\mathrm{V}$ against the maker's $0.08\,\mathrm{V}$, at $1.57793\,\mathrm{V}$ per block.

Table 8.4: The MAP state of the joint state of the two tests, with the posterior and prior spreads of each entry.

Entry [unit]	$\hat{z}_i$ [entry's]	$\sqrt{\Sigma_{ii}}$ [entry's]	Prior sd [entry's]	Informed by
ξ_0 [–]	-0.00275	0.00201	none	window 0's discharge, run time and recharge
$R_{\mathrm{bat},0}$ [mΩ]	39.24966	1.35659	none	window 0's recharge level, mostly
ξ_{36} [–]	-0.17558	0.00216	none	window 36's discharge, run time and recharge
$R_{\mathrm{bat},36}$ [mΩ]	53.97711	1.47779	none	window 36's recharge level, mostly
Δv [V]	1.57793	0.00141	0.08	both tests' discharge levels; the maker's tolerance

The residuals. At the MAP state twice the objective is $2\mathcal{L}(\hat{\mathbf{z}}) = 85.2761$ on $82 - 5 = 77$ degrees of freedom, and $\Pr(\chi^2_{77} > 85.28) = 0.243$: the readings sit within their errors. The largest residual of a normalized row is 3.0679, window 36's recharge reading at 200 minutes, one row in 82. The two run-time rows have the residuals -0.5733 and -0.1469, and the prior row 0.3491: the record puts Δv 0.35 of the maker's tolerance above the nominal.

The lines from the state. The four lines of each test, equation (8.3), are derived quantities of the state, and their bands follow by the delta method (4.53). At window 0 the discharge line has the level $482.3471\,\mathrm{V}$ (standard deviation 0.0799) and the slope $25.3164\,\mathrm{V/h}$ (0.0520), and the recharge line $478.3550\,\mathrm{V}$ (0.0814) and $12.6582\,\mathrm{V/h}$ (0.0260). At window 36 they are $481.7580\,\mathrm{V}$ (0.0840) and $30.0926\,\mathrm{V/h}$ (0.0669), and $479.2386\,\mathrm{V}$ (0.0887) and $15.0463\,\mathrm{V/h}$ (0.0334).

The logs and these lines are plotted in Figure 8.2, one panel per test. The horizontal axis is the time from switch-on in minutes, the vertical axis the string voltage; the dots are the logged voltages, the discharge on the left of each panel and the recharge after the cutoff on the right, the solid lines the two lines from the state, the dashed horizontal line the cutoff of $420\,\mathrm{V}$ and the dotted vertical line the logged run time. Each log lies on its line within the voltmeter's scatter. Today's discharge falls faster and ends sooner, and today's recharge starts higher above the cutoff than at commissioning: the step at the reversal, from $420\,\mathrm{V}$ to the recharge level, is $(I_d + I_c)R_{\mathrm{bat},t}/1000$ above $Nh_{\mathrm{ch}} = 56.00\,\mathrm{V}$, and it is that step the resistance is read from.

8.2.4 The full posterior by reweighted draws

The Laplace normal $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ is the reported posterior. Beside it stands the full posterior of the same rows, computed by the reweighted draws of Section 5.2.4: draws from a proposal, each weighted by the ratio of the posterior to the proposal, equation (5.13), with the effective sample size $N_{\mathrm{eff}} = 1/\sum_j \varpi_j^2$ and Monte Carlo standard errors by batch means over 20 batches. The

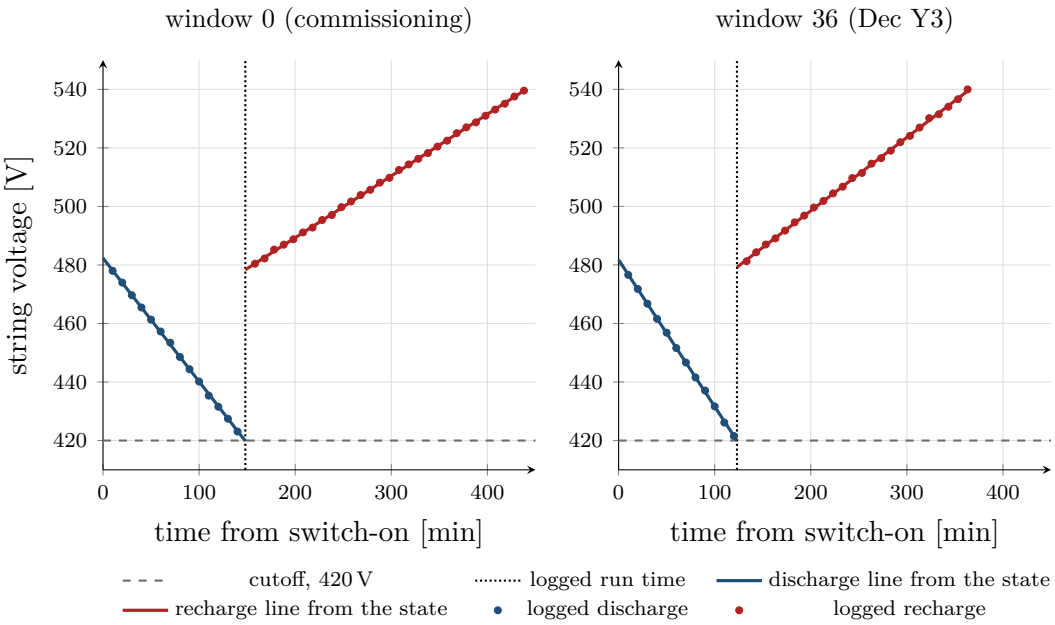


Figure 8.2: The two capacity tests of the string: the logged voltages (dots) and the four lines from the state, equation (8.3) at $\hat{\mathbf{z}}$ (solid): at window 0 the discharge $482.3471 - 25.3164\,\vartheta$ and the recharge $478.3550 + 12.6582\,\vartheta$, at window 36 $481.7580 - 30.0926\,\vartheta$ and $479.2386 + 15.0463\,\vartheta$, in V with ϑ in hours from the start of each log. The recharge starts at the logged run time (dotted), 148.1 and 123.2 minutes. Dashed: the cutoff of 420 V.

draws are good only where the proposal covers the posterior, and the two routes of this chapter need different proposals.

The full record. The Laplace normal in the coordinates of the state is the proposal, with $\mathcal{L}$ the objective (8.8) through the forward equations (8.2). For this record 200 000 draws from a seeded generator give $N_{\text{eff}} = 199\,992$, 100.0 % of the draws, and a largest weight of 5.79×10^{-6}: the Laplace normal is close to the posterior, and no draw dominates.

The discharge alone. The comparison route of Section 8.2.2 fits the discharge and run-time rows without the recharge. Its Laplace normal in the state's coordinates is a poor proposal, with $N_{\text{eff}} = 8422$, 4.21 % of the draws: along the ridge (8.6) the log-capacity changes by $1/\Delta v$ per volt, so the ridge is curved in $(\xi_t, R_{\text{bat},t}, \Delta v)$, ξ_t logarithmic in Δv, while the Laplace normal is straight. That route is drawn in the coordinates of its lines instead. Let $\mathbf{s} = (V_{d,0},\ \dot{V}_{d,0},\ V_{d,36},\ \dot{V}_{d,36})^{\mathsf{T}}$ be the two discharge lines of equation (8.3). Every discharge row and run-time row is a function of $\mathbf{s}$ alone, and the map from $(\mathbf{s}, \Delta v)$ to the state,

$$\xi_t = \ln \frac{N\,\Delta v\,I_d}{\bar{C}_{\text{bat}}\,\dot{V}_{d,t}}, \qquad R_{\text{bat},t} = \frac{1000\,\big(N(v_0 + \Delta v) - V_{d,t}\big)}{I_d}, \tag{8.13}$$

is one-to-one, the inverse of equation (8.3). Its change-of-variable factor (Section 4.3.3) is $|\partial \mathbf{z}/\partial(\mathbf{s}, \Delta v)| = (1000/I_d)^2/(\dot{V}_{d,0}\,\dot{V}_{d,36})$, which does not involve Δv. The posterior of the discharge route, flat in the state apart from the maker's tolerance, is therefore in the new coordinates

$$p(\mathbf{s}, \Delta v \mid \mathbf{y}_{\text{d}}) \propto \frac{\exp\big[-\mathcal{L}_{\text{d}}(\mathbf{s})\big]}{\dot{V}_{d,0}\,\dot{V}_{d,36}}\ \exp\Big[-\tfrac{1}{2}\Big(\frac{\Delta v - 1.55}{0.08}\Big)^2\Big], \tag{8.14}$$

with $\mathbf{y}_{\text{d}}$ the discharge and run-time readings and $\mathcal{L}_{\text{d}}$ their objective: a product of a posterior of the lines and the maker's normal, so that the maker's normal is the exact marginal of Δv on

this route. The proposal is the Laplace normal of $\mathbf{s}$ from the fit of the lines times the maker's normal, the weights correct the lines alone, and 200 000 draws give $N_{\text{eff}} = 199\,996$, 100.0 % of the draws. Each draw is mapped to the state by equation (8.13). Both routes give the full posterior of the stated model up to its Monte Carlo error, named so here and in what follows.

8.3 The answers (step 4)

Every answer is read from the pair $(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ of Section 8.2.3, the full record's, and that value is the one reported. A derived quantity f is carried by its gradient $\mathbf{g} = \partial f/\partial \mathbf{z}$ at $\hat{\mathbf{z}}$, with variance $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$. The full posterior of Section 8.2.4 stands beside each value, a median beside a median.

8.3.1 The capacity and the resistance now

Asset manager's question. What is the usable capacity now, and the internal resistance? The two are asked together, and the drill asks that the correlation between them be read before either is reported.

General form. An estimation question about entries of the state asks for their posterior distribution, equation (2.9),

$$\Pr(f(\mathbf{z}) \le f^* \mid \mathbf{y}) = \int \mathbf{1}\big[f(\mathbf{z}) \le f^*\big]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z},$$

for each entry, summarized by its moments, equation (2.10), and its credible interval, equation (2.11), at the levels 0.05 and 0.95; for two entries asked together the moments include their covariance,

$$\mathrm{Cov}(z_i, z_j \mid \mathbf{y}) = \int \big(z_i - \mathbb{E}[z_i \mid \mathbf{y}]\big)\big(z_j - \mathbb{E}[z_j \mid \mathbf{y}]\big)\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z},$$

and its correlation $\mathrm{Cov}(z_i, z_j \mid \mathbf{y})/\sqrt{\mathrm{Var}(z_i \mid \mathbf{y})\,\mathrm{Var}(z_j \mid \mathbf{y})}$.

Specialized form. The question is equations (2.9), (2.10) and (2.11) with $\mathbf{z}$ the joint state (8.4), the two entries $z_3 = \xi_{36}$ and $z_4 = R_{\text{bat},36}$, and $p(\mathbf{z} \mid \mathbf{y})$ the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ of Section 8.2.2. Their joint posterior is the 2×2 block of $\boldsymbol{\Sigma}$,

$$\boldsymbol{\Sigma}_{36} = \begin{bmatrix} \Sigma_{33} & \Sigma_{34} \\ \Sigma_{34} & \Sigma_{44} \end{bmatrix}, \qquad \varrho = \frac{\Sigma_{34}}{\sqrt{\Sigma_{33}\Sigma_{44}}}, \qquad \tilde{\boldsymbol{\Sigma}}_{36} = \begin{bmatrix} 1 & \varrho \\ \varrho & 1 \end{bmatrix} = (1+\varrho)\,\mathbf{e}_+\mathbf{e}_+^{\mathsf{T}} + (1-\varrho)\,\mathbf{e}_-\mathbf{e}_-^{\mathsf{T}}, \tag{8.15}$$

with $\tilde{\boldsymbol{\Sigma}}_{36}$ the block of the two entries each divided by its own standard deviation and $\mathbf{e}_\pm = (1,\ \pm 1)^{\mathsf{T}}/\sqrt{2}$ its eigenvectors, as in *Reliability Engineering by Derivation*, Chapter 7. The combination along $\mathbf{e}_-$, the standardized capacity less the standardized resistance, has the standard deviation $\sqrt{1-\varrho}$, and the one along $\mathbf{e}_+$, their sum, $\sqrt{1+\varrho}$: as $\varrho \to 1$ the record fixes the difference and leaves the sum to whatever holds the ridge. The marginals then follow. The capacity is the monotone image $C_{\text{bat},36} = \bar{C}_{\text{bat}}\, \mathrm{e}^{\xi_{36}}$, so its median and band are the images of the normal's; the resistance is the entry itself,

$$C_{\text{bat},36}:\ \bar{C}_{\text{bat}}\, \mathrm{e}^{\hat{\xi}_{36}},\ \big[\bar{C}_{\text{bat}}\, \mathrm{e}^{\hat{\xi}_{36} \mp 1.6449\sqrt{\Sigma_{33}}}\big], \qquad R_{\text{bat},36}:\ \hat{R}_{\text{bat},36},\ \big[\hat{R}_{\text{bat},36} \mp 1.6449\sqrt{\Sigma_{44}}\big]. \tag{8.16}$$

Which rows set each answer is read from the lines. Inverting equation (8.3) for today's test,

$$\begin{aligned} R_{\text{bat},36} &= \frac{1000\,(V_{c,36} - N h_{\text{ch}} - V_{\text{cut}})}{I_d + I_c}, \\ \Delta v &= \frac{V_{d,36} + I_d R_{\text{bat},36}/1000}{N} - v_0, \qquad C_{\text{bat},36} = \frac{N\,\Delta v\, I_d}{\dot{V}_{d,36}}: \end{aligned} \tag{8.17}$$

the resistance from the recharge level alone, the curve rise from the discharge level once the resistance is known, and the capacity from the curve rise and the discharge slope. Beside the reported route stands the same block from the discharge and run-time rows alone, the fit of the same state without the recharge, whose rank deficiency step 2 found; its posterior along the ridge is the maker's tolerance.

Building blocks. The pair $(\hat{\mathbf{z}}, \mathbf{\Sigma})$ of Section 8.2.3 and its block (8.15); the bands (8.16); the lines from the state and the inversion (8.17); the fit of the discharge alone; the full posterior of Section 8.2.4 for both routes.

The correlation, read first. On the full record $\Sigma_{33} = 4.6615 \times 10^{-6}$, $\Sigma_{44} = 2.1839\,\mathrm{m\Omega}^2$ and $\Sigma_{34} = 2.6772 \times 10^{-3}\,\mathrm{m\Omega}$, so $\varrho = +0.8391$. The standardized eigenvalues are 0.1609 and 1.8391, with standard deviations 0.4011 along $\mathbf{e}_-$ and 1.3561 along $\mathbf{e}_+$, a ratio of 3.4: the record knows the two entries' difference better than their sum, and it knows both. On the discharge alone the correlation is +0.99926, the eigenvalues 0.00074 and 1.99926, the standard deviations 0.0273 and 1.4140, a ratio of 51.9. There capacity and resistance lie on one line, the ridge (8.6), with $C_{\mathrm{bat},36}$ growing as Δv and $R_{\mathrm{bat},36} = 25.517 + 1000\,(\Delta v - 1.55)\,\mathrm{m\Omega}$, held in place only by the maker's tolerance.

The two routes. Table 8.5 sets the routes side by side: each row is a route, the columns the capacity with its band, the resistance with its standard deviation and band, the correlation and the curve rise. On the discharge alone the Laplace normal puts 0.4551 of $R_{\mathrm{bat},0}$ and 0.3750 of $R_{\mathrm{bat},36}$ below zero, a resistance no string has. With the recharge the capacity's spread narrows 24.0 times, the standard deviation of ξ_{36} from 0.05175 to 0.00216, and the resistance's 54.2 times, from 80.086 to 1.478 mΩ. The full posterior agrees with the Laplace normal on both routes.

Table 8.5: The capacity and the resistance at window 36 by route. Bands are 5–95 %; the reported route is the full record's Laplace normal.

Route	$C_{\mathrm{bat},36}$ [Ah]	band [Ah]	$R_{\mathrm{bat},36}$ [mΩ]	sd [mΩ]	band [mΩ]	corr [–]	Δv [V]
Discharge alone, Laplace	82.372	75.65–89.69	25.517	80.086	−106.21 to 157.25	0.99926	1.55000
Discharge alone, full posterior	82.369	75.327–89.400	25.463	–	−106.849 to 157.428	–	–
Discharge and recharge, Laplace (reported)	83.897	83.600–84.195	53.977	1.478	51.546–56.408	0.8391	1.57793
Discharge and recharge, full posterior	83.897	83.601–84.195	53.978	–	51.551–56.405	0.8388	1.57792
The monitor (Δv taken as 1.55)	82.341	sd 0.343	25.311	3.847	–	–	1.55

The source of each answer. By equation (8.17) the resistance is the recharge level less $56.00 + 420.00\,\mathrm{V}$, over $(I_d + I_c)/1000 = 0.06\,\mathrm{V}$ per mΩ: the recharge level's standard deviation of 0.0887 V is 1.478 mΩ, the whole of the resistance's. The curve rise then comes from the discharge level, and the capacity from the curve rise and the discharge slope; the curve rise's own variance, $(\sqrt{\Sigma_{55}}/\widehat{\Delta v})^2$ on the logarithm, is 17.1 % of the variance of ξ_{36}. The recharge alone, with the run times, gives 83.971 Ah and 54.064 mΩ with a correlation of 0.5009: the recharge carries most of the resistance and the discharge most of the capacity.

The ridge through the answer. The discharge alone's Laplace normal, conditioned on the full record's curve rise of 1.57793 V, gives 83.869 Ah and 53.443 mΩ, close to the full record's answer, on its Laplace conditional, which is linear in ξ_{36} along the ridge; conditioned on the maker's nominal 1.55 V, it gives 82.372 Ah and 25.517 mΩ, the monitor's reading of 82.341 Ah and 25.311 mΩ to within its own fit of the line. Against the full record's answer the monitor's capacity lies −8.67 of the posterior's standard deviations and its resistance −19.40. In closed form, by equation (8.13), a discharge read with the curve rise taken as 1.55 gives the resistance $\hat{R}_{\mathrm{bat},36} - 1000N(\widehat{\Delta v} - 1.55)/I_d = 26.051\,\mathrm{m\Omega}$: the measured rise is 0.02793 V per block above the nominal, and it is worth 27.926 mΩ of resistance that the monitor does not see.

The joint posterior of the two quantities is shown in Figure 8.3, in two panels. The horizontal axis is the capacity $C_{\mathrm{bat},36}$ and the vertical axis the resistance $R_{\mathrm{bat},36}$. The left panel shows the discharge alone: its contours at one and two standard deviations, a needle along the ridge, the shaded region $R_{\mathrm{bat},36} < 0$, the monitor's point and the stated truth, a check marker for Section 8.4.5. The right panel enlarges the neighborhood of the full record's answer: the full record's contours at one and two standard deviations, the Laplace conditional of the discharge alone along its ridge, and the truth. The needle crosses the zero line, its 5–95 % band of resistance running from −106.21 to 157.25 mΩ; the full record's contours are a small tilted ellipse that sits next to that line.

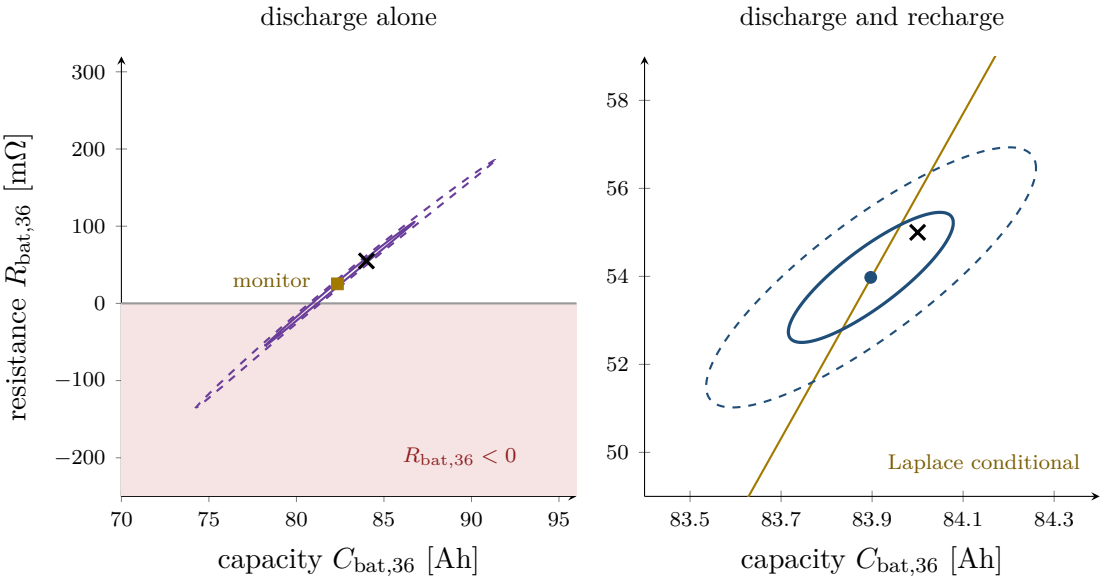


Figure 8.3: The joint posterior of today's capacity and resistance under the Laplace normal. Left, the discharge alone: $\ln(C/100) \sim$ normal with mean $\ln 0.82372$ and standard deviation 0.05175, R with mean 25.517 and standard deviation 80.086 mΩ, correlation 0.99926; solid and dashed, the contours at one and two standard deviations, drawn as $\hat{\mathbf{m}} + n_\sigma \mathbf{L}(\cos\phi, \sin\phi)^{\mathsf{T}}$ with $\mathbf{L}$ the Cholesky factor of the 2×2 block and $n_\sigma = 1, 2$, the capacity mapped through $C = 100\,\mathrm{e}^{\xi}$; shaded, $R < 0$; square, the monitor. Right, the full record: mean (83.897, 53.977), standard deviations 0.00216 in ξ and 1.478 mΩ, correlation 0.8391; gold, the Laplace conditional of the discharge alone along its ridge, $C = 82.372\,\mathrm{e}^{0.64516(\Delta v - 1.55)}$ and $R = 25.517 + 1000(\Delta v - 1.55)$; dot, the MAP pair. Cross: the stated truth, a check marker (Section 8.4.5).

The answer. The capacity now is 83.897 Ah, band [83.600, 84.195], and the resistance 53.977 mΩ, standard deviation 1.478, band [51.546, 56.408], correlated +0.8391 through each log's level–slope trade. The full posterior gives 83.897 Ah with [83.601, 84.195] and 53.978 mΩ with [51.551, 56.405], correlation 0.8388. The discharge record alone could have said neither: its

capacity band runs from 75.65 to 89.69 Ah and its resistance band from −106.21 to 157.25 mΩ.

8.3.2 The state of health against commissioning

Asset manager's question. How much capacity has the string lost since commissioning, and how far has its resistance risen? The first is the capacity index h_C, the second the rise ΔR, with the probability that the resistance has risen at all.

General form. An estimation question about a derived quantity is summarized by its moments, equation (2.10), and its credible interval, equation (2.11); whether it exceeds a level is a tail under the posterior, equation (2.12),

$$\Pr(f(\mathbf{z}) \geq f^* \mid \mathbf{y}) = \int \mathbf{1}\big[f(\mathbf{z}) \geq f^*\big]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}.$$

Specialized form. The question is equations (2.10), (2.11) and (2.12) with $f(\mathbf{z}) = \ln h_C = z_3 - z_1$ and $f(\mathbf{z}) = \Delta R = z_4 - z_2$ of equation (8.1), and $p(\mathbf{z} \mid \mathbf{y})$ the Laplace normal. One choice enters here: the capacity index is carried as its logarithm, as Chapter 5 carried h, and the resistance rise linearly. This makes both answers closed forms in $\hat{\mathbf{z}}$ and $\boldsymbol{\Sigma}$, and it is accurate only if each is close to normal in the posterior; the full posterior beside it tests that. Both quantities are linear in the state, with the gradients of the second line of the pattern (8.5), so by the delta method (4.53) their variances are the change-since-commissioning variances (3.38),

$$\mathrm{Var}(\ln h_C \mid \mathbf{y}) = \Sigma_{11} + \Sigma_{33} - 2\Sigma_{13}, \qquad \mathrm{Var}(\Delta R \mid \mathbf{y}) = \Sigma_{22} + \Sigma_{44} - 2\Sigma_{24}, \tag{8.18}$$

each the budget of a commissioning term, today's term and a cross term whose sign is that of minus the covariance. The capacity index has the median $\mathrm{e}^{\widehat{\ln h_C}}$ and the band $\mathrm{e}^{\widehat{\ln h_C} \mp 1.6449\,\mathrm{sd}(\ln h_C \mid \mathbf{y})}$, and the rise the normal band; whether the resistance has risen is

$$\Pr(\Delta R > \Delta R^* \mid \mathbf{y}) = \Phi\left(\frac{\widehat{\Delta R} - \Delta R^*}{\mathrm{sd}(\Delta R \mid \mathbf{y})}\right), \tag{8.19}$$

with $\Delta R^* = 0$ for any rise. The ohmic state of health, $R_{\mathrm{bat},36}/R_{\mathrm{bat},0}$, is carried as its logarithm, with the gradient $(0,\ -1/\hat{R}_{\mathrm{bat},0},\ 0,\ 1/\hat{R}_{\mathrm{bat},36},\ 0)$. The covariance of the two answers is $\mathbf{g}_{\ln h_C}^{\mathsf{T}} \boldsymbol{\Sigma}\, \mathbf{g}_{\Delta R} = \Sigma_{12} + \Sigma_{34} - \Sigma_{14} - \Sigma_{23}$.

The curve rise drops out of both quantities. The ridge direction (8.6) has equal entries in the two log-capacities and equal entries in the two resistances, so it is orthogonal to both gradients: $\mathbf{g}_{\ln h_C}^{\mathsf{T}} \mathbf{n} = -1/\Delta v + 1/\Delta v = 0$, and the same for ΔR. Against commissioning the maker's tolerance cancels, on the discharge alone as on the full record.

Building blocks. The pair $(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ of Section 8.2.3; the gradients of the pattern (8.5); the variances (8.18); the tail (8.19); the normal quantile 1.6449.

The capacity index. By equation (8.18), $\widehat{\ln h_C} = -0.175580 - (-0.002751) = -0.172830$ and

$$\mathrm{Var}(\ln h_C \mid \mathbf{y}) = 4.0487 \times 10^{-6} + 4.6615 \times 10^{-6} - 2\,(1.0096 \times 10^{-7}) = 8.5082 \times 10^{-6} = 0.002917^2.$$

The index has median 0.84128 and band [0.83725, 0.84533]: the string has lost 15.87 % of its commissioned capacity, between 15.47 and 16.27 %. The full posterior gives 0.84128 and [0.83725, 0.84532].

The resistance rise. $\widehat{\Delta R} = 53.9771 - 39.2497 = 14.7274\,\mathrm{m\Omega}$ and

$$\mathrm{Var}(\Delta R \mid \mathbf{y}) = 1.8403 + 2.1839 - 2\,(4.6471 \times 10^{-3}) = 4.0149\,\mathrm{m\Omega}^2 = (2.0037\,\mathrm{m\Omega})^2,$$

so the rise has the band $[11.432,\ 18.023]\,\mathrm{m\Omega}$ and, by equation (8.19), $\Pr(\Delta R > 0 \mid \mathbf{y}) = \Phi(7.350)$, 1.000000 to six places; $\Pr(\Delta R > 10\,\mathrm{m\Omega} \mid \mathbf{y}) = 0.9908$. The full posterior gives $14.732\,\mathrm{m\Omega}$ with $[11.427,\ 18.017]$, a probability of any rise of 1.000000 and of more than $10\,\mathrm{m\Omega}$ of 0.9905. The ohmic state of health is 1.3752, band $[1.2791,\ 1.4785]$, a resistance up 37.5 % (full posterior 1.3752 and $[1.2797,\ 1.4793]$), against the monitor's "up 213 %".

The budget of the variance. Table 8.6 splits each variance into the three terms of equation (8.18); each row is a term, with its value and its share for each quantity. For the capacity index the commissioning test carries 47.6 %, today's test 54.8 % and the cross term −2.4 %; for the resistance rise 45.8, 54.4 and −0.2 %. Both cross terms are small and negative. The two tests share only the curve rise, which both records now measure, and the shared part cancels in the difference, as in Chapter 5's budget. The two answers are correlated with each other by +0.8565, through the same level–slope trade that correlates each test's two entries.

Table 8.6: The variance budget of the state of health, equation (8.18), at the MAP state.

Term	$\ln h_C$ [–]	Share [%]	ΔR [$\mathrm{m\Omega}^2$]	Share [%]
Commissioning entry	4.0487×10^{-6}	47.6	1.8403	45.8
Today's entry	4.6615×10^{-6}	54.8	2.1839	54.4
Cross term	$-2\,(1.0096 \times 10^{-7})$	−2.4	$-2\,(4.6471 \times 10^{-3})$	−0.2
Total	8.5082×10^{-6}	100	4.0149	100

The discharge alone against commissioning. On the discharge alone the state of health is $h_C = 0.84323$, with standard deviation of the logarithm 0.005183, and $\Delta R = 16.482\,\mathrm{m\Omega}$, standard deviation 5.0428: the ridge cancels, as the specialized form showed, and the state of health is sharp on that route too. The recharge takes a further 43.7 % off the first standard deviation and 60.3 % off the second. The state of health was never the ridge's casualty; the absolute values were.

The densities of the two answers are plotted in Figure 8.4. The left panel has the capacity index on its horizontal axis and the right the resistance rise; the vertical axes are the posterior densities. The solid curves are the full record's, the dash-dotted curves the discharge alone's; the dashed lines are the dashboard's tile on the left and the monitor's readings on both, and the dotted lines the stated truth. On both panels the full record's density is the narrower and the two routes overlap. The tile lies below the full record's capacity density, in the left tail of the discharge alone's; the monitor's capacity index and rise lie inside the discharge alone's densities, because against commissioning its error cancels: what the monitor gets wrong are the absolute values and the ratio built from them.

8.3.3 The dashboard's run-time index

Asset manager's question. Is the dashboard's 83.2 % the capacity? The tile divides today's run time by the commissioning run time and reports the ratio as the capacity against commissioning.

General form. An estimation question about a derived quantity is summarized by its moments, equation (2.10),

$$\mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}] = \int f(\mathbf{z})\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z},$$

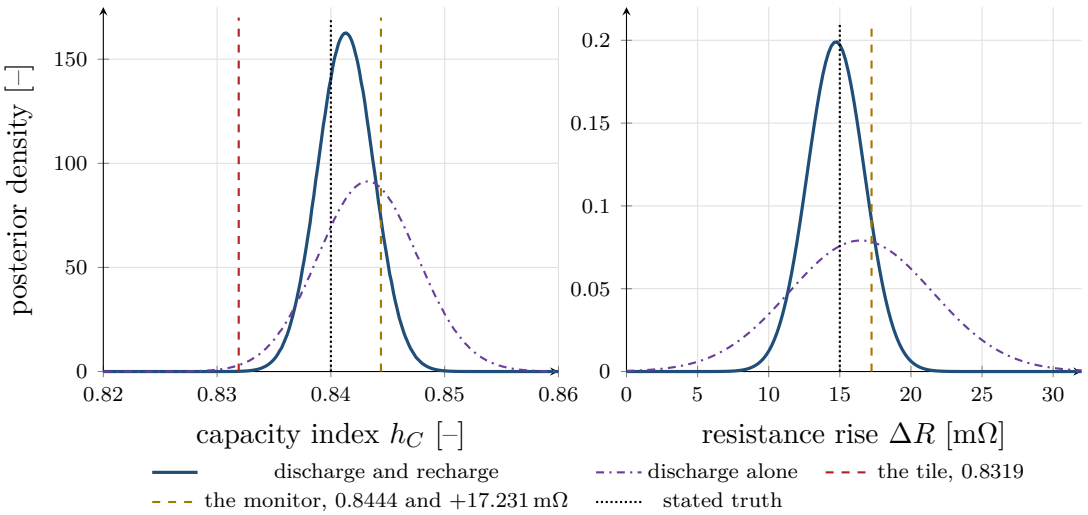


Figure 8.4: The posterior densities of the state of health at window 36 against commissioning. Left: the capacity index, lognormal with $\ln h_C \sim \mathcal{N}(-0.172830,\ 0.002917^2)$ on the full record (solid) and median 0.84323 with standard deviation of the logarithm 0.005183 on the discharge alone (dash-dotted). Right: the resistance rise, normal with mean 14.727 and standard deviation 2.0037 mΩ (solid) and 16.482 with 5.0428 mΩ (dash-dotted). Dashed: the run-time tile (0.8319) and the monitor (0.8444 and +17.231 mΩ). Dotted: the stated truth, a check marker (Section 8.4.5).

and by its credible interval, equation (2.11), at the levels 0.05 and 0.95.

Specialized form. The question is equations (2.10) and (2.11) with $f(\mathbf{z}) = \ln(T_{\text{run},36}/T_{\text{run},0})$ and $p(\mathbf{z} \mid \mathbf{y})$ the Laplace normal. By the third line of equation (8.2), $1 - \text{SOC}_{\text{cut},t} = V_{\text{h},t}/(N\,\Delta v)$ with $V_{\text{h},t} = N(v_0 + \Delta v) - V_{\text{cut}} - I_d R_{\text{bat},t}/1000$ the loaded head, the voltage the discharge starts above the cutoff, so that $T_{\text{run},t} = C_{\text{bat},t}\, V_{\text{h},t}/(N\,\Delta v\, I_d)$, and the logarithm of the ratio splits into two terms,

$$\ln \frac{T_{\text{run},36}}{T_{\text{run},0}} = \ln h_C + \ln \frac{N(v_0 + \Delta v) - V_{\text{cut}} - I_d R_{\text{bat},36}/1000}{N(v_0 + \Delta v) - V_{\text{cut}} - I_d R_{\text{bat},0}/1000}, \tag{8.20}$$

the capacity index and the ratio of the heads. A resistance that has risen lowers the head, so the run time falls even at an unchanged capacity: the run-time ratio is the capacity index only when the resistance has not changed. The gradient of f has -1 and 1 in the log-capacity columns, $\pm I_d/(1000 V_{\text{h},t})$ in the resistance columns, + for window 0 and − for window 36, and $N/V_{\text{h},36} - N/V_{\text{h},0}$ in the column of Δv; the delta method (4.53) carries it.

Building blocks. The pair $(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$; the split (8.20) and its gradient; the capacity index of Section 8.3.2.

The run-time index from the state. At the MAP state the logarithm of the ratio is −0.182323, with the standard deviation 0.001929. The index therefore has the median 0.83333 and the band [0.83069, 0.83598], beside the logged run times' 0.83192; the full posterior gives 0.83334 and [0.83068, 0.83598]. By equation (8.20) the logarithm splits as

$$-0.182323 = -0.172830 + (-0.009494)\ :$$

the head falls from 62.3471 V at commissioning to 61.7580 V today, and the rise of 14.727 mΩ takes 0.5891 V of it at 40 A. The run-time index from the state charges the resistance rise to

the capacity: 5.2 % of its logarithm is the resistance, and it lies 0.00795 below the capacity index, 3.25 standard deviations of $\ln h_C$. The tile's own 0.83192, from the logged run times, lies 0.00936 below it, 3.83 standard deviations. The cutoff leaves 0.0122 of the charge in the string at commissioning and 0.0215 today, a share the higher resistance strands below the cutoff.

8.3.4 The replacement criterion

Asset manager's question. Is the string still above the replacement criterion of 80 %? The criterion is read against the commissioned capacity, as the record states it, and against the rating, as the maker's warranty would.

General form. A prediction under a threshold is a tail of a derived quantity under the posterior, equation (2.12),

$$\Pr(f(\mathbf{z}) < f^* \mid \mathbf{y}) = \int \mathbf{1}\big[f(\mathbf{z}) < f^*\big]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}.$$

Specialized form. The question is equation (2.12) with $f(\mathbf{z}) = \ln h_C$ and $f^* = \ln 0.80$ against commissioning, and $f(\mathbf{z}) = \xi_{36}$ and $f^* = \ln 0.80$ against the rating, and $p(\mathbf{z} \mid \mathbf{y})$ the Laplace normal, so that

$$\Pr(h_C < 0.80 \mid \mathbf{y}) = \Phi\left(\frac{\ln 0.80 - \widehat{\ln h_C}}{\mathrm{sd}(\ln h_C \mid \mathbf{y})}\right), \qquad \Pr(C_{\mathrm{bat},36} < 80\,\mathrm{Ah} \mid \mathbf{y}) = \Phi\left(\frac{\ln 0.80 - \hat{\xi}_{36}}{\sqrt{\Sigma_{33}}}\right). \tag{8.21}$$

One choice enters here: the criterion of 80 % is a stated input of the record, the convention the site's maintenance plan adopts, and not a quantity of this posterior. This keeps the answer to today's margin on today's evidence; when the margin is used up is a question of a rate, which two tests do not fix.

Building blocks. The capacity index of Section 8.3.2; the entry $\hat{\xi}_{36}$ and $\sqrt{\Sigma_{33}}$; Φ.

The margin today. Against commissioning, equation (8.21) gives $\Pr(h_C < 0.80 \mid \mathbf{y}) = \Phi(-17.25) = 5.7 \times 10^{-67}$; the margin is 0.0413 in the index, 17.2 standard deviations of $\ln h_C$. Against the rating, $\Pr(C_{\mathrm{bat},36} < 80\,\mathrm{Ah} \mid \mathbf{y}) = \Phi(-22.03) = 7.5 \times 10^{-108}$. No draw of the full posterior falls below either level. On the discharge alone the same probability against the rating is $\Phi(-0.5646) = 0.2862$: the rating question belonged to the maker's tolerance until the recharge measured Δv, while against commissioning the tolerance cancels. The monitor's capacity, read with the nominal curve, gives 4.3×10^{-12}, and the tile's 0.8319 leaves a margin of 0.0319.

8.4 The checks (step 5)

Each answer of Section 8.3 rests on approximations and pinned inputs that this string lets be tested against the calculation they stand for.

8.4.1 The full posterior against the Laplace answers

The reweighted draws of Section 8.2.4 have been set beside every answer of step 4. On the full record, with 200 000 draws in 20 batches, the largest gap between the full posterior and the Laplace normal, over the median and the two band ends of each quantity, and the largest Monte Carlo standard error of the three are: for $C_{\mathrm{bat},36}$, 0.00109 Ah against a standard error of 0.00095; for $R_{\mathrm{bat},36}$, 0.00510 mΩ against 0.00606; for Δv, 0.00001 V against 0.00001; for h_C, 0.00000 against 0.00001; for ΔR, 0.00608 mΩ against 0.00959; and for the run-time ratio,

0.00001 against 0.00001. The correlation of today's two entries is 0.8388 with a standard error of 0.0005, against 0.8391, and the probability of a resistance rise is 1.000000 on both routes. Every gap is within about one standard error, so the draws resolve no difference: on the full record the Laplace normal is the posterior to the precision the answers are reported in.

No draw of the full record has a resistance at or below zero; the weight there is 0, and under the Laplace normal the probabilities are 2.3×10^{-184} for $R_{\mathrm{bat},0}$ and 2.2×10^{-292} for $R_{\mathrm{bat},36}$. Restricting the posterior to positive resistances would change nothing, and the chapter does not impose it. On the discharge alone the same restriction would remove 0.4554 of the weight and move the median resistance from 25.5 to 74.1 mΩ: there the support, not a reading, would set the answer, one more sign that that route measures no resistance.

8.4.2 The limit of exact rows and the curvature term

The fit runs on the reduced state, which equation (4.48) makes the limit $\sigma_p \to 0$ of the full state. The check fits the chapter's own full state of 84 entries, the five of equation (8.4) and the state of charge at each of the 79 logged times, with the 79 coulomb-counting rows of equation (8.2), from full for the discharge and from the cutoff for the recharge, at a common tolerance σ_p beside the reading rows and the prior, and lets the tolerance shrink. Table 8.7 gives the result: each row is a tolerance, the columns the two answers of the second question and today's resistance and the curve rise, each with its standard deviation, and the last row the reduced state. At 0.1 and 0.01 the spreads are wider by the slack the rows allow; from 10^{-4} the full state reproduces the reduced state to the precision of the table.

Table 8.7: The limit of exact coulomb counting on the 84-entry full state of the two tests, against the reduced state.

σ_p [–]	$\ln h_C$ [–]	sd [–]	ΔR [mΩ]	sd [mΩ]	$R_{\mathrm{bat},36}$ [mΩ]	sd [mΩ]	Δv [V]	sd [V]
10^{-1}	−0.175156	0.026725	13.6578	39.5777	55.1036	30.0765	1.57458	0.03053
10^{-2}	−0.173804	0.006272	14.2796	4.9697	54.0955	3.6872	1.57831	0.00371
10^{-3}	−0.172846	0.002999	14.7200	2.0636	53.9794	1.5220	1.57794	0.00145
10^{-4}	−0.172830	0.002918	14.7274	2.0043	53.9771	1.4782	1.57793	0.00141
10^{-5}	−0.172830	0.002917	14.7274	2.0037	53.9771	1.4778	1.57793	0.00141
reduced state	−0.172830	0.002917	14.7274	2.0037	53.9771	1.4778	1.57793	0.00141

The limit is approached more slowly than in Chapter 5, and the reason is the rows' scale: a slack σ_p in the state of charge moves a logged voltage by $N\,\Delta v\,\sigma_p$, which is 0.631 V at $\sigma_p = 0.01$ and 0.063 V at 0.001, against the voltmeter's 0.25 V. Until the slack is well below the voltmeter's error, the readings rather than the coulomb count set each state of charge, and the spreads stay wider.

The covariance (8.11) is the Gauss–Newton form, without the curvature term $\mathbf{C} = \mathbf{H} - \mathbf{J}^{\mathsf{T}}\mathbf{J}$ of equation (4.44). The full Hessian of the objective at $\hat{\mathbf{z}}$, computed by differences, gives the Newton covariance, in which the standard deviations are 0.002918 for $\ln h_C$, 2.003966 for ΔR, 0.002159 for ξ_{36} and 1.477871 for $R_{\mathrm{bat},36}$, against Gauss–Newton's 0.002917, 2.003717, 0.002159 and 1.477786, ratios of 1.0001 to 1.0002. The largest entry of $\mathbf{C}$ is 3.0×10^{-4} of the largest entry of $\mathbf{J}^{\mathsf{T}}\mathbf{J}$. The Gauss–Newton covariance is the one reported.

8.4.3 Sensitivity to the priors and the pinned inputs

Table 8.8 repeats the fit under other priors and other choices of step 2, and under errors in the pinned charge offset. Each row is a variant; the columns give today's capacity and resistance with their spreads and correlation, the two answers of the second question with theirs, and the fitted curve rise.

Table 8.8: The answers under other priors, other modeling choices and other values of the pinned charge offset; the full record unless named.

Variant	$C_{\mathrm{bat},36}$ [Ah]	sd ξ_{36} [–]	$R_{\mathrm{bat},36}$ [mΩ]	sd [mΩ]	corr [–]	h_C [–]	sd $\ln h_C$ [–]	ΔR [mΩ]	sd [mΩ]	Δv [V]
tolerance sd 0.04 V	83.897	0.00216	53.979	1.478	0.8391	0.84128	0.002917	14.728	2.004	1.57790
tolerance sd 0.08 V (the chapter's)	83.897	0.00216	53.977	1.478	0.8391	0.84128	0.002917	14.727	2.004	1.57793
tolerance sd 0.16 V	83.897	0.00216	53.977	1.478	0.8391	0.84128	0.002917	14.727	2.004	1.57793
tolerance centered 1.63 V	83.897	0.00216	53.976	1.478	0.8391	0.84128	0.002917	14.727	2.004	1.57795
tolerance centered 1.47 V	83.897	0.00216	53.979	1.478	0.8391	0.84128	0.002917	14.728	2.004	1.57790
no run-time rows	83.885	0.00229	53.928	1.515	0.8446	0.84152	0.003079	14.857	2.054	1.57781
each test its own Δv	83.899	0.00218	53.969	1.481	0.8197	0.84133	0.002985	14.714	2.009	1.57780
h_{ch} pinned 0.02 V high	83.188	0.00216	40.644	1.478	0.8385	0.84128	0.002917	14.727	2.004	1.56460
h_{ch} pinned 0.02 V low	84.606	0.00216	67.311	1.478	0.8397	0.84128	0.002917	14.728	2.004	1.59126
h_{ch} freed, sd 0.02 V	83.857	0.00861	53.222	13.235	0.9855	0.84128	0.002917	14.727	2.004	1.57717
h_{ch} freed, sd 0.005 V	83.894	0.00302	53.929	3.644	0.8823	0.84128	0.002917	14.727	2.004	1.57788
discharge alone	82.372	0.05175	25.517	80.086	0.9993	0.84323	0.005183	16.482	5.043	1.55000
discharge alone, each test its own Δv	82.372	0.05175	25.517	80.086	0.9993	0.84323	0.073175	16.482	113.249	1.55000

The maker's tolerance after the recharge. Halving or doubling the tolerance, or moving its center by one tolerance either way, moves no answer beyond the third decimal of a resistance and the fifth of the curve rise. The record measures Δv itself, and the prior row holds nothing the readings do not.

The choices of step 2. Without the run-time rows the capacity index moves to 0.84152 and the rise to 14.857 mΩ, with the standard deviations 0.003079 and 2.054 mΩ in place of 0.002917 and 2.004: the cutoff's instant is a reading worth keeping, not one that sets the answer. Giving each test its own curve rise moves the index to 0.84133 and the rise to 14.714 mΩ, and widens $\ln h_C$ from 0.002917 to 0.002985: on the full record each test measures its own rise, so an aging curve would barely touch the state of health. On the discharge alone the same choice widens $\ln h_C$ 14.1 times, to 0.073175, because there only the shared tolerance made the two tests' ridges move together; the recharge also protects the state of health against a curve that ages.

The charge offset behind the absolute values. The recharge moves the confound from Δv to h_{ch}, which the maker's sheet pins. By equation (8.17) a change of the charge offset moves the resistance by $1000N/(I_d + I_c) = 666.7$ mΩ per volt per block, 13.33 mΩ for 0.02 V, with the opposite sign, and through the discharge level it moves Δv and the capacity with it: pinned 0.02 V high, the offset gives 40.644 mΩ and 83.188 Ah; pinned 0.02 V low, 67.311 mΩ and 84.606 Ah. Freed with a tolerance of 0.02 V it rebuilds a ridge, with a correlation of 0.9855 and a resistance spread of 13.235 mΩ; with 0.005 V, the spread is 3.644 mΩ.

Effects test by test. An error in the charge offset moves both tests' resistances alike, because the same pinned value enters both recharge levels, and it leaves h_C and ΔR where they were: 0.84128 and 14.727 or 14.728 mΩ in every row of the offset. The capacity of both tests moves with the curve rise, by the same factor, and cancels from the index. The state of health is again the robust answer, and the absolute resistance is as good as the charge offset.

The two answers are plotted against the pinned offset in Figure 8.5. The horizontal axis is the value at which h_{ch} is pinned, the vertical axis a resistance; the descending line is today's resistance $R_{\mathrm{bat},36}$, with the three fitted values as dots, and the flat line the rise ΔR; the shaded band is the resistance's 5–95 % band at the chapter's offset, and the dashed lines mark ±0.02 V. The descending line crosses the band in a narrow strip of offsets: an error of 0.00222 V per block moves today's resistance by its own standard deviation, 1.478 mΩ. The flat line does not move.

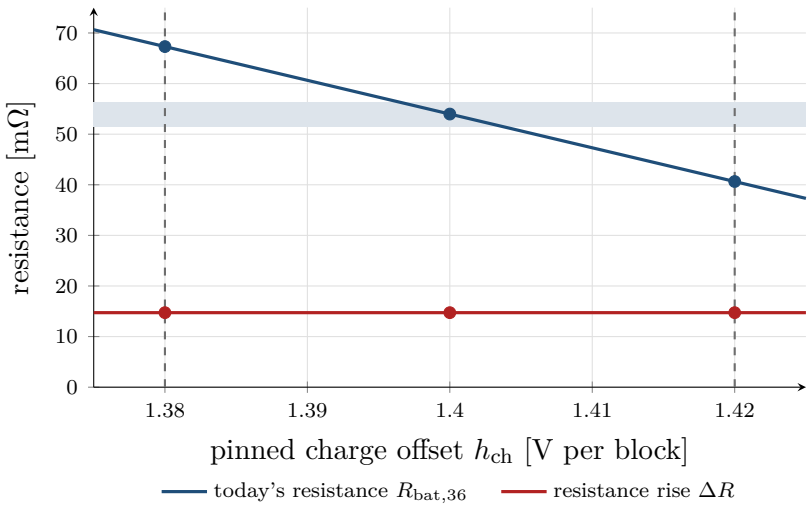


Figure 8.5: Today's resistance and the resistance rise against the pinned charge offset. Blue line: $R_{\mathrm{bat},36} = 53.977 - 666.7\,(h_{\mathrm{ch}} - 1.40)\,\mathrm{m\Omega}$, the slope $1000N/(I_d + I_c)$ of equation (8.17); dots, the refits at 1.38, 1.40 and 1.42 V (Table 8.8). Shaded: the 5–95 % band of $R_{\mathrm{bat},36}$ at the pinned 1.40 V. Red: ΔR, 14.727 or 14.728 mΩ at every offset. Dashed: the offsets 0.02 V either side.

This is the surprise of the checks. The recharge did not remove the confound of a discharge; it moved it from the curve rise, which the record now measures, to the charge offset, which the sheet pins and no row of the record sees. The state of health against commissioning does not depend on it, and the absolute resistance does, one for one.

8.4.4 Other forms of the state

Section 8.2.2 chose one state over two tests. The commissioning test fitted alone, on the state $(\xi_0, R_{\mathrm{bat},0}, \Delta v)$, gives $\xi_0 = -0.00278$ with standard deviation 0.00204, $R_{\mathrm{bat},0} = 39.255\,\mathrm{m\Omega}$ with 1.358, and $\Delta v = 1.57780$ with 0.00191, with the correlations 0.8198 between ξ_0 and $R_{\mathrm{bat},0}$, 0.2314 between ξ_0 and Δv, and -0.0560 between $R_{\mathrm{bat},0}$ and Δv. Carried into a fit of window 36 as Chapter 2's marginal rows on ξ_0 and $R_{\mathrm{bat},0}$, with the maker's tolerance stated again, it gives $h_C = 0.84133$ with $\mathrm{sd}(\ln h_C) = 0.002985$ and $\Delta R = 14.714\,\mathrm{m\Omega}$ with 2.0087, against the joint state's 0.002917 and 2.0037: ratios of 1.02 and 1.00, the same 0.002985 that giving each test its own curve rise produced. Carried as one row on $(\xi_0, R_{\mathrm{bat},0}, \Delta v)$ with the full covariance of the commissioning fit, it gives 0.002917 and 2.0037, the joint state. The settlement of one state over the tests is kept. Here, with the curve rise pinned down by both records, the marginal rows cost 2 %; on the discharge alone the same loss of the shared rise widens $\ln h_C$ 14.1 times, and there it decides everything.

8.4.5 The answers against the stated truth

The last check sets the answers against the stated truth of the record (Section 8.6.6). Table 8.9 collects the routes: each row is a route or a variation, the columns today's capacity and resistance with their correlation, the capacity index with its band or the standard deviation of its logarithm, and the resistance rise with its band or standard deviation. Every entry of the state lies inside its band. The truth puts the capacity index at 0.8400, today's capacity at 84.0 Ah, today's resistance at 55.0 mΩ, the rise at 15.0 mΩ and the curve rise at 1.58 V. The posterior's $\ln h_C$ lies +0.52 standard deviations from the truth and ΔR −0.14; ξ_{36} lies −0.57 and $R_{\mathrm{bat},36}$ −0.69; and Δv lies −1.47 standard deviations from the true 1.58 V, after the record moved it from the maker's 1.55 V.

The monitor lies −4.8 of its own standard deviations from the true capacity and −7.7 from the true resistance, and it reads 8.08 mΩ for a string commissioned at 40 mΩ. The discharge alone lies −0.37 of its standard deviations from the true resistance, with a standard deviation of 80.086 mΩ and a band from −106.21 to 157.25 mΩ: uninformative, not wrong. The tile's logged 0.83192 lies 0.00808 below the true capacity index, 3.31 standard deviations of $\ln h_C$, which is the resistance rise it charges to the capacity.

Table 8.9: The routes to the state of health at window 36 and the stated truth. Bands are 5–95 %; where marked, the standard deviation of $\ln h_C$ or of ΔR stands in place of a band.

Route	$C_{\mathrm{bat},36}$ [Ah]	$R_{\mathrm{bat},36}$ [mΩ]	corr [–]	h_C [–]	band [–]	ΔR [mΩ]	band [mΩ]
Laplace, full record (reported)	83.897	53.977	0.8391	0.84128	0.83725–0.84533	14.727	11.432–18.023
Full posterior, full record	83.897	53.978	0.8388	0.84128	0.83725–0.84532	14.732	11.427–18.017
Newton covariance	–	sd 1.4779	–	–	sd 0.002918	–	sd 2.0040
Full state, $\sigma_p = 10^{-5}$	–	53.9771	–	–	sd 0.002917	14.7274	sd 2.0037
Discharge alone, Laplace	82.372	25.517	0.99926	0.84323	sd 0.005183	16.482	sd 5.0428
Discharge alone, full posterior	82.369	25.463	–	–	–	–	–
Recharge alone, with the run times	83.971	54.064	0.5009	0.83994	sd 0.003591	13.943	sd 2.2541
Commissioning as marginal rows	–	–	–	0.84133	sd 0.002985	14.714	sd 2.0087
The monitor	82.341	25.311	–	0.8444	–	17.231	–
Stated truth	84.0	55.0	–	0.8400	–	15.0	–

8.5 The report (step 6)

The monitor's "resistance up 213 %" is the maker's curve rise read as exact. The record puts this string's curve 0.02793 V per block above the nominal, worth 27.926 mΩ of the resistance the monitor reports, and a discharge record cannot see that: on its own it puts capacity and resistance on one line, correlated +0.9993. The recharge, whose resistance term has the opposite sign, measures both, and the curve rise as well. The capacity is 83.9 Ah (83.6 to 84.2), 0.841 of commissioning (0.837 to 0.845); the resistance is 54.0 mΩ (51.5 to 56.4), up 14.7 mΩ (11.4 to 18.0) or 37.5 % since commissioning; the two are correlated +0.84. The string is 17 standard deviations above the 80 % criterion. The run-time index charges the resistance rise to the capacity: read from the state it lies 3.25 standard deviations of $\ln h_C$ below the capacity index, and the tile's logged 0.8319 lies 3.83 below it.

8.5.1 The answers and their bands

Table 8.10 gives the answers in the order of Table 8.2. Each row is a question; the second column gives the reported value from the Laplace normal of the full record, the third the full posterior's value beside it, and the fourth the 5–95 % band, or for a probability the margin it rests on.

8.5.2 The entry in the maintenance log

The capacity test of window 36 is the first condition entry of the UPS string's own maintenance log, after its window-0 commissioning test. It uses nothing later than window 36, and in the values of Table 8.10 it reads:

Table 8.10: The report for the UPS string at window 36 (Dec Y3) against its commissioning test. Reported values from the Laplace normal of the joint state on the discharge and recharge records; bands are 5–95 %.

Asset manager's question	Reported value	Full posterior	Band
What is the usable capacity now, and the internal resistance?	$C_{\mathrm{bat},36} = 83.897$ [Ah]; $R_{\mathrm{bat},36} = 53.977$ [mΩ], sd 1.478 [mΩ]; correlation 0.8391 [–]	83.897 [Ah]; 53.978 [mΩ]; 0.8388 [–]	83.600 to 84.195 [Ah]; 51.546 to 56.408 [mΩ]
How much capacity has the string lost since commissioning, and how far has its resistance risen?	$h_C = 0.84128$ [–], lost 15.87 %; $\Delta R = 14.727$ [mΩ], $P(\Delta R > 0) = 1.000000$ [–]	0.84128 [–]; 14.732 [mΩ]; 1.000000 [–]	0.83725 to 0.84533 [–]; 11.432 to 18.023 [mΩ]
Is the dashboard's 83.2 % the capacity?	no: run-time index 0.83333 [–], of which 5.2 % of the logarithm is the resistance; capacity index 0.84128 [–]	0.83334 [–]	0.83069 to 0.83598 [–]
Is the string still above the replacement criterion of 80 %?	yes: $P(h_C < 0.80) = 5.7 \times 10^{-67}$ [–]; $P(C_{\mathrm{bat},36} < 80\ \mathrm{Ah}) = 7.5 \times 10^{-108}$ [–]	no draw below either level	margin 0.0413 [–], 17.2 sd of $\ln h_C$

Window 36 (Dec Y3), UPS string, capacity test (discharge at 40 A, recharge at 20 A) against the commissioning test (window 0).
Capacity: 83.9 Ah (5–95 % band 83.6 to 84.2), 0.841 of commissioning (0.837 to 0.845), lost 15.9 %.
Internal resistance: 54.0 mΩ (51.5 to 56.4), risen 14.7 mΩ (11.4 to 18.0) since commissioning, 37.5 %. Capacity and resistance correlated +0.84.
Curve rise measured at 1.578 V per block, 0.35 of the maker's tolerance above the nominal; the monitor's discharge-only resistance (25 mΩ, "up 213 %") is not a reading.
Replacement criterion, 80 % of commissioning: met, $P(\text{below}) = 6 \times 10^{-67}$.
Absolute resistance as good as the maker's charge offset (13 mΩ per 0.02 V per block); the rise does not depend on it.
Action: no replacement. Correct the monitor's curve rise to the measured value.
Next measurement: a string monitor logging the next test every minute; the charge offset of a sample block.

Common mistake

Reading a resistance from a discharge record alone. A discharge fixes a level and a slope; the circuit has a capacity, a resistance and a curve rise, and the third slides the other two together along a ridge. A discharge-only resistance is the maker's tolerance seen through the level, not a reading of the string.

8.6 The next measurement (step 7)

The report puts today's capacity between 83.600 and 84.195 Ah and today's resistance between 51.546 and 56.408 mΩ. A further measurement in today's test is one more set of rows of the residual vector, and what it removes from each band is known before its value is.

8.6.1 The preposterior of a candidate row

A candidate measurement a is one row or more in today's test, with Jacobian lines $\mathbf{J}_a$ in the columns of the joint state, each the gradient of the quantity it reads over its standard deviation.

Its preposterior covariance is equation (4.65),

$$\boldsymbol{\Sigma}_a = \left(\boldsymbol{\Sigma}^{-1} + \mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a\right)^{-1}, \qquad \mathrm{sd}_a(f) = \sqrt{\mathbf{g}_f^{\mathsf{T}}\boldsymbol{\Sigma}_a\,\mathbf{g}_f}, \tag{8.22}$$

for each answer f with its gradient $\mathbf{g}_f$, and the width it takes off is $1 - \mathrm{sd}_a/\mathrm{sd}$. The candidates open to the asset manager are of five kinds. A string monitor, a new logger installed on the string and not the existing monitor's fit, logging today's test every minute, adds the 333 readings between the 10-minute ones, rows of equation (8.7) at the new times; it is the permanent instrument at the place of the question. The reversal step, the recharge's first reading taken at the instant the current reverses, reads the recharge level $V_{c,36}$ itself, and the switch-on reading, taken at the instant the discharge starts, reads the discharge level $V_{d,36}$; both by the same voltmeter. A 30-minute discharge at 80 A with three readings is a second test current. One choice enters with it: the loaded curve of equation (8.2) is taken to hold at 80 A as at 40, with the polarization at that rate lumped into the same v_0 and Δv, and the usable capacity is taken not to depend on the discharge rate, which for lead-acid it does; this is what lets three readings at another current act on the same five entries, and it fails if the maker's curve at 80 A differs or the capacity falls at the higher rate, which the record cannot test. And the maker's tolerance can be narrowed to 0.01 V by measuring a sample block's curve. Two limits stand beside them, which are not measurements: Δv known exactly, and today's four lines known exactly, the most any reading of today's test can give.

8.6.2 The candidates

Table 8.11 ranks the candidates. Each row is a measurement in today's test; the columns give the share of the present standard deviation it removes from each of four answers, today's two entries and the two answers of the state of health, by equation (8.22), and the correlation of today's entries after it. The widths removed are compared as bars in Figure 8.6, the absolute values in the left panel and the state of health in the right, ranked by the share taken off today's resistance, with the limit of today's four lines known exactly as dashed lines in the right panel.

Table 8.11: Candidate measurements in today's test, from the posterior of the joint state; the width removed is taken from the present standard deviations 0.00216 of ξ_{36}, 1.478 mΩ of $R_{\mathrm{bat},36}$, 0.002917 of $\ln h_C$ and 2.0037 mΩ of ΔR.

Candidate (today's test)	ξ_{36} [%]	$R_{\mathrm{bat},36}$ [%]	$\ln h_C$ [%]	ΔR [%]	corr after [–]
A string monitor logging the test every minute, 333 more readings (permanent)	66.7	68.9	28.0	28.6	0.8224
The reversal step, read at the instant of reversal	4.0	5.8	2.2	3.1	0.8239
The switch-on reading at full charge	1.3	2.8	1.1	1.4	0.8355
A 30-minute discharge at 80 A, three readings (the loaded curve taken to hold)	4.8	12.7	3.5	6.4	0.8322
The maker's tolerance narrowed to 0.01 V (a sample block's curve)	0.0	0.0	0.0	0.0	0.8394
Δv known exactly	0.9	0.2	0.0	0.0	0.8557
Today's four lines known exactly (the commissioning floor)	100.0	100.0	32.1	32.4	0.2992

A curve rise already measured. A sample block's curve, narrowing the maker's tolerance to 0.01 V, takes nothing off any answer, and even Δv known exactly takes 0.9 % off the log-capacity and 0.2 % off the resistance: the record already measures the rise, to 0.00141 V. The candidate a discharge-only reading of the string would rank first is worth nothing once the recharge is read.

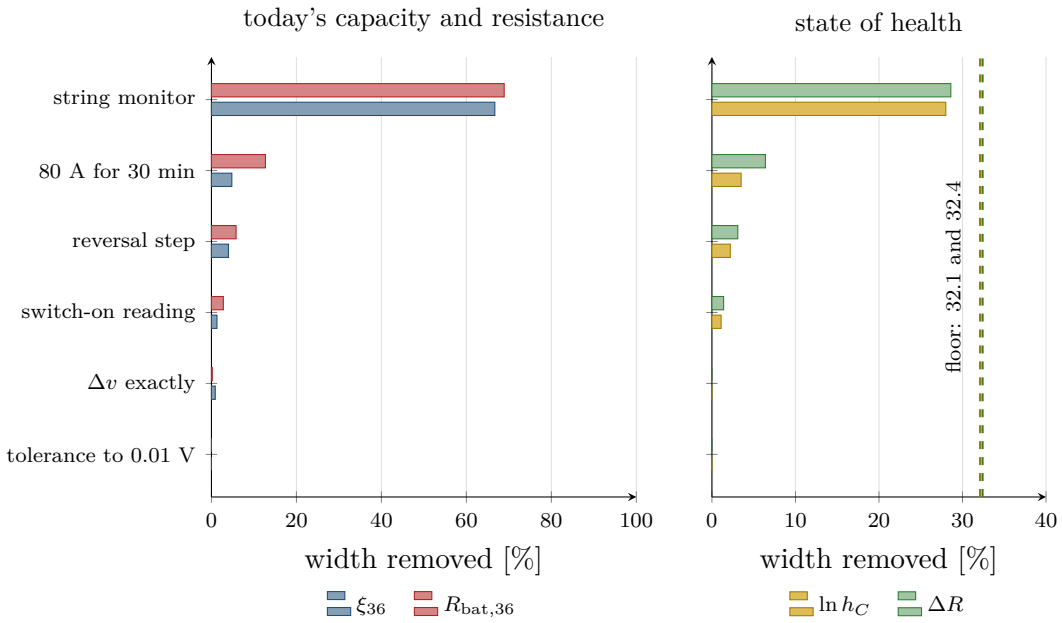


Figure 8.6: The candidate measurements of Table 8.11, ranked by the width they remove from today's resistance, by the preposterior (8.22). Left: today's log-capacity and resistance, whose limit with today's four lines known exactly is 100 %. Right: the capacity index and the resistance rise; dashed, the same limit, 32.1 and 32.4 %, the most any measurement in today's test can remove.

The same logs, more often, as the best candidate. The ranking follows each candidate's line in the columns of today's entries. The string monitor adds many rows of the kind already present and takes 66.7 and 68.9 % off the absolute values and 28.0 and 28.6 % off the state of health, near the limit of 32.1 and 32.4 %. The 80 A discharge is the best single test after it on the resistance, 12.7 %, because a second current moves the resistance term of the discharge level without moving the curve; it rests on the loaded curve and the capacity holding at that rate. The reversal step reads the recharge level directly and takes 5.8 % off the resistance; the switch-on reading takes less, because the discharge level carries the curve rise as well as the resistance.

8.6.3 The commissioning floor and the charge offset

Today's four lines known exactly fix today's capacity and resistance completely and take only 32.1 and 32.4 % off the state of health's standard deviations, because the commissioning test's own entries are untouched: they carry 46.2 % of the variance of $\ln h_C$ and 45.8 % of that of ΔR, which no row of today's test removes. The first is below the budget's 47.6 % because the four lines also fix Δv, which takes $\Sigma_{15}^2/\Sigma_{55}$ off the variance of ξ_0, 1.4 points of the variance of $\ln h_C$. That is the commissioning floor of Chapter 5, the gold bars of its variance budget, met again on a battery: the commissioning test sets the precision of every later state of health, and after handover it can only be carried, not improved.

What the preposterior cannot rank is the charge offset. It is pinned, so no candidate row narrows it, and yet it moves today's resistance by 13.33 mΩ for each 0.02 V per block of error, Section 8.4.3. Measuring a sample block's charge offset is a check on a pinned input, not a row of the preposterior: it has no width removed to report, and it goes into the recommendation and the limits for what it protects, the absolute resistance.

8.6.4 The recommendation

Install the string monitor and log every capacity test every minute: it takes 66.7 % off the capacity's band and 68.9 % off the resistance's, and 28.0 and 28.6 % off the state of health, near the floor. Have the maker measure a sample block's charge offset, on which the absolute resistance rests. And correct the monitor's curve rise to the measured 1.578 V per block, so that its discharge-only readings stop reporting the maker's tolerance as a resistance. A sample block's loaded curve is not worth requesting for this question.

8.6.5 What the analysis showed

The monitor said the string had 82.3 Ah and 25.3 mΩ, its resistance "up 213 %", and the tile said 0.832 of its capacity; read from both records as the parameters of one circuit, the string has 83.9 Ah and 54.0 mΩ, up 37.5 %, at 0.841 of its commissioned capacity. The difference is the curve. A discharge log is a line of two numbers, the circuit has three, and the correlation read first, +0.9993 on the discharge alone, says the record measures one combination of capacity and resistance and leaves the rest to the maker's tolerance; the monitor's resistance is that tolerance read as exact. The recharge, whose ohmic term has the opposite sign and whose level holds no curve rise, measures both parameters and the curve rise as well. The state of health against commissioning was sharp on either record, because the tolerance cancels between two tests of one curve, and the recharge only made it sharper; the absolute values exist because of the recharge, and the confound did not vanish but moved, to the charge offset the sheet pins. The resistance read today is as good as that pinned value, and the rise since commissioning does not depend on it.

8.6.6 Limits of the model

A constructed record. The string's two tests are generated from a stated truth, with the stated voltmeter error, and step 5 compares the answers with that truth (Section 8.4.5). They are not field data, and the agreement is a check of the method on its own model, not of the model against a real string. The constructed run time carries no reading error, so the run-time rows, which give it the voltmeter's error over the slope, are conservative.

Linear curves. The loaded curve is linear in the state of charge, with the polarization at the test rates lumped into it. Near the knee of a deep discharge, or at another current, the logs bend, and the four lines no longer carry the whole record.

One resistance, full efficiency, no rest. The same ohmic resistance is taken in charge and discharge, the charge efficiency as one, and the recharge as starting at the cutoff with no rest. A difference between the charge and discharge resistances, a charge lost to gassing, or a rest on open circuit before the recharge would each go into the resistance read from the recharge level.

Pinned v_0 and h_{ch}. The maker's sheet pins the loaded curve's empty point and the charge offset. The absolute resistance moves by 13.33 mΩ for each 0.02 V per block of error in the offset; the state of health against commissioning does not move.

No temperature. Temperature is not logged, and both the capacity and the resistance of a lead-acid block depend on it; a test run warmer than the commissioning test would read as a change of condition.

A curve that does not age, and two tests. One curve rise is shared by both tests. On this record that matters little (Section 8.4.3), and on a discharge-only record it would decide the state of health. Two tests say nothing about how fast the capacity fades or whether it accelerates; the rate is Chapter 10's question and a remaining life that of Chapter 14, in Volume II.

The next chapter. Chapter 9 reads one unit's parameter with the help of fifty siblings of its design: a shared prior where this chapter had a shared curve.

8.7 What to retain

Definition 8.1 (State of health from an equivalent circuit)**.** The state of health of a battery is the posterior of the capacity and the resistance of its equivalent circuit in two tests, read from one joint state that carries both tests and the curve they share: the capacity index $h_C = \mathrm{e}^{\xi_t - \xi_0}$ and the resistance rise $\Delta R = R_{\mathrm{bat},t} - R_{\mathrm{bat},0}$, each with the variance of a difference. Today's capacity and resistance are read together, their correlation first. A constant-current discharge fixes a level and a slope, and with an uncertain curve rise its rows have a null direction, along which capacity and resistance slide together; a recharge, whose level carries the resistance with the opposite sign and no curve rise, removes it. Whatever a record cannot see is held by a prior or a pinned input, and the answer is as good as that input.

Three conclusions carry beyond the string. The correlation of two parameters read from one record says which combination the record measures, and a correlation near one means that an entry's band is the band of whatever holds the ridge, not a reading. A second record whose rows move the parameters in a direction the first cannot copy breaks the ridge, and the direction is read from a column of the Jacobian, here the sign of the resistance column. And a change since commissioning can be sharp when the absolute values are not, because what the two tests share cancels from it: the state of health is the robust answer of a battery, and the absolute values are as good as their pinned inputs.

Table 8.12 lists the terms of the chapter with their expressions, units and the subsection that defines each.

Table 8.12: The terms of the battery's state of health and where each is defined. Units for the UPS string.

Term	Expression	Units	Meaning	Defined in
Equivalent circuit	eq. (8.2)	V, Ah, mΩ	loaded curve, ohmic resistance, coulomb counting	8.1.2
The four lines of a test	eq. (8.3)	V, V/h	two levels and two slopes per test	8.1.2
State of health	h_C, ΔR, eq. (8.1)	–, mΩ	capacity index and resistance rise	8.1.1
Ohmic state of health	$R_{\mathrm{bat},t}/R_{\mathrm{bat},0}$	–	the resistance ratio	8.1.1
Rows and posterior				
Confounded parameters	the ridge $\mathbf{n}$, eq. (8.6)	mixed	the null direction of the discharge rows	8.1.4
The recharge's opposite sign	$-I_d/1000$ against $+(I_d+I_c)/1000$ in the R column	V/mΩ	what breaks the ridge	8.1.4
Block structure	eq. (8.12)	mixed	two test blocks meeting in Δv	8.2.2
Reweighted draws in the line's coordinates	eqs. (8.13), (8.14)	–	the full posterior of a curved ridge	8.2.4
Questions and answers				
Correlation read first	eq. (8.15)	–	which combination the record measures	8.3.1
Answers from the lines	eq. (8.17)	mΩ, V, Ah	which row sets each answer	8.3.1
Variance of the state of health	eq. (8.18)	–, $\mathrm{m\Omega}^2$	a difference's variance	8.3.2
Run-time index	eq. (8.20)	–	capacity index times the head ratio	8.3.3
Replacement criterion	eq. (8.21)	–	a stated input, 80 %	8.3.4
Pinned charge offset	$1000N/(I_d+I_c)$ per V	mΩ/V	the move of R the posterior cannot show	8.4.3
Preposterior of a candidate	eq. (8.22)	–	the band after a measurement	8.6.1

8.8 Exercises

Exercise 8.1. *Formulation.* Write the question "has the resistance risen by more than 20 mΩ since commissioning?" in the form of Table 8.2 and evaluate it from the fit of this chapter. (A prediction under a threshold, equation (2.12) with $f=\Delta R$ and $f^*=20\,\mathrm{m\Omega}$; by equation (8.19), $\Pr(\Delta R>20\mid\mathbf{y})=\Phi\big((14.727-20)/2.0037\big)=0.0043$.)

Exercise 8.2. *Derivation.* From the discharge lines of equation (8.3), show that the discharge rows are blind to the direction $\mathrm{d}\xi_t/\mathrm{d}\Delta v=1/\Delta v$, $\mathrm{d}R_{\mathrm{bat},t}/\mathrm{d}\Delta v=1000N/I_d$, and evaluate how far one tolerance of the curve rise moves today's capacity and resistance along it. (Along the direction $\mathrm{d}V_{d,t}=N\,\mathrm{d}\Delta v-(I_d/1000)\,\mathrm{d}R_{\mathrm{bat},t}=0$ and $\mathrm{d}\ln\dot V_{d,t}=\mathrm{d}\Delta v/\Delta v-\mathrm{d}\xi_t=0$, and the run time is a function of the two; at $\Delta v=1.55$, one tolerance of 0.08 V moves the capacity by 5.16 % and the resistance by 80.0 mΩ.)

Exercise 8.3. *Derivation.* Show that the recharge level is $V_{c,t}=Nh_{\mathrm{ch}}+V_{\mathrm{cut}}+(I_d+I_c)R_{\mathrm{bat},t}/1000$, free of the curve rise, and evaluate it for today's test at the MAP state. (At $\vartheta=0$ the recharge of equation (8.2) reads $N(v_0+h_{\mathrm{ch}})+N\,\Delta v\,\mathrm{SOC}_{\mathrm{cut},t}+I_cR_{\mathrm{bat},t}/1000$, and $N\,\Delta v\,\mathrm{SOC}_{\mathrm{cut},t}=V_{\mathrm{cut}}+I_dR_{\mathrm{bat},t}/1000-Nv_0$; at $\hat R_{\mathrm{bat},36}$, $56.00+420.00+3.2386=479.2386$ V.)

Exercise 8.4. *Derivation.* Split the logarithm of the run-time ratio into the capacity index and a resistance term, and evaluate the split. (By equation (8.2), $T_{\mathrm{run},t} = C_{\mathrm{bat},t} V_{\mathrm{h},t} / (N\,\Delta v\, I_d)$ with $V_{\mathrm{h},t} = N(v_0 + \Delta v) - V_{\mathrm{cut}} - I_d R_{\mathrm{bat},t}/1000$, so the ratio is h_C times $V_{\mathrm{h},36}/V_{\mathrm{h},0}$, equation (8.20): $-0.182323 = -0.172830 + (-0.009494)$.)

Exercise 8.5. *Application.* The reversal step of today's test is read at 479.50 V (standard deviation 0.25 V), against the fitted recharge level of 479.2386 V, whose posterior standard deviation is 0.0887 V. Update today's resistance, capacity and resistance rise by the correction form and by a refit. (The row's gradient is that of $V_{c,36}$, which by equation (8.3) lies in the column of $R_{\mathrm{bat},36}$; by the rank-one update (4.66) with the innovation $479.50 - 479.2386$, $R_{\mathrm{bat},36}$ moves from 53.977 to 54.464 mΩ and its standard deviation from 1.478 to 1.393, the capacity from 83.897 to 83.947 Ah and the rise from 14.727 to 15.213 mΩ. The refit with the row gives 54.464 mΩ with standard deviation 1.393, 83.947 Ah and 15.213 mΩ.)

Chapter 9

Fleet Pooling and Hierarchical Models

9.1 The fleet chart and the coil's place in it

Coil 1 of the two-unit hall is also coil 0 of a fleet: fifty coils of the same model, each rated 7.0 kW/K, at the operator's sites, commissioned in different years, so that at the end of the record they are one to twelve years old. Each is tested once a year, the commissioning test repeated, and the fleet's condition-monitoring chart draws each coil's own straight line through its tests, the tests taken as exact, and ranks the coils by the slope. Coil 0 is three years old and has three tests. Its line falls by 0.17 %/yr, 0.075 of the chart's fleet mean of −2.27 %/yr, and that ranks coil 0 39th of the 47 coils that have a slope, with 1 the fastest-fouling: only eight coils foul more slowly by their lines. Three coils have one test each, no slope, and no place on the chart. The conclusion the chart invites is that coil 0 is one of the slowest-fouling coils of its design, fouling at a tenth of the fleet's rate, and can go to the back of the cleaning queue.

The chart cannot say how much three tests over two years, each read to a standard deviation of 0.04, know of a slope: the standard error of coil 0's own slope is $0.0283\,\mathrm{yr}^{-1}$, larger than the slope itself. Nor can it say what the other forty-nine coils of the same design say about coil 0, or what coil 0's tests say about the design. The question of this chapter is

where coil 0 sits among its fifty siblings, read from its own record pooled with the fleet's through a shared prior whose mean and spread are learned, with its band, against the coil read alone and against one line for the whole fleet.

It is answered with the method of Chapter 4, on a state of 104 entries.

Jargon: partial pooling, shrinkage, exchangeable units, the population's spread, fleet position, separate and fully pooled drawings, funnel, centered and non-centered parametrizations, marginal of the spreads, composition draws, information cap

Partial pooling (a hierarchical model) estimates each unit of a fleet with a shared prior whose mean and spread are themselves estimated from the fleet; *shrinkage* is the move of a unit's estimate from its own record's value toward the fleet's. *Exchangeable units* are units that nothing distinguishes before their records are read. *The population's spread* is the standard deviation of a parameter across the units of a design, the unit-to-unit variability. A unit's *fleet position* is its *rank* among the units of the fleet by a parameter, 1 the fastest-fouling, and its *percentile* in the design's population of that parameter. The *separate drawing* reads every unit alone; the *fully pooled drawing* gives every unit one shared value. A *funnel* is the shape of the joint density of a population's spread and its units' parameters: wide at a large spread and narrowing to a neck as the spread shrinks with the units pulled to the mean. In the *centered parametrization* the units' parameters are entries of the state; in the *non-centered parametrization* the entries are standardized deviations, multiplied by the spread inside the rows. The *marginal of the spreads* is their posterior with every other entry integrated out. *Composition draws* are exact draws of a mixture: a component by its weight, then a draw from that component. The *information cap* of a unit is the most it can tell the fleet's mean, whatever its record.

Key idea

A coil's rate is read from its own record and the fleet's at once. The tie rows that hold each coil to the fleet have a tolerance, the population's spread, and the record learns it. Where a coil's record is short the coil borrows the fleet's information, and where it is long the coil keeps its own. The fleet's mean and spread are entries of the same state, so each coil's tests also inform them, up to a cap that the spread sets.

9.1.1 Partial pooling as tie rows with a free tolerance

Chapter 2 carried an earlier fit into a later one as a row, the commissioning row $(g-\hat{g}_0)/\sigma_{\hat{g}_0}$, and Chapter 4, fitting the trend of one air path from two condition tests, left the trend entry without a prior and named "a prior on it from the coils of this design" as this chapter's (Section 4.6.5). A prior from the coils of a design is a row of the same kind, with one difference: its mean and its width are not known, and the coils of the design are what inform them.

Reliability Engineering by Derivation, Chapter 6, tied the rates of three pumps together by tie rows: each pump's departure from a shared rate entered as a Gaussian row of a chosen tolerance. At a tolerance of zero the rows hold exactly and the three pumps share one rate; at an infinite tolerance the rows vanish and each pump keeps its own. Between the two limits the tolerance was the engineer's choice. Here the tolerance is an entry of the state. Each coil's commissioning value and fouling rate are tied to the design's mean by a row whose standard deviation is the population's spread, and the spread is estimated with everything else.

> **Principle 9.1.** *The population's spread is a quantity of the state like any other. Fixed at zero it gives the fleet's one line, fixed at infinity it gives each coil alone, and between them the record decides. A coil's position in its fleet is a derived quantity of the whole state, its own entries and the fleet's together.*

9.1.2 The fleet as a graph (step 0)

The fleet has fifty coil nodes and one fleet node. Coil i, $i = 0, \ldots, 49$, carries two free entries against its rating: its commissioning value $g_{i,0}$, the log-derate of Chapter 5 at age zero, and its fouling rate γ_i, the change of the log-derate per year of age. At age a its log-derate is

$$g_i(a) = g_{i,0} + \gamma_i\, a, \tag{9.1}$$

its derate factor $d_i(a) = \mathrm{e}^{g_i(a)}$ and its health index against commissioning $h_i(a) = \mathrm{e}^{g_i(a) - g_{i,0}} = \mathrm{e}^{\gamma_i a}$. Each coil's annual tests hang on it as readings. The fleet node carries four free entries: the means μ_c and μ_γ of the design's commissioning values and rates, and their spreads τ_c and τ_γ, carried as $\ell_\mathrm{c} = \ln \tau_\mathrm{c}$ and $\ell_\gamma = \ln \tau_\gamma$ so that they stay positive. Each coil is tied to the fleet node by two tie rows, one on each of its entries.

The fleet has no physical rows. Each annual test is a condition test, the commissioning test repeated at the same inputs, already reduced to one reading of g: Chapter 5's fit of the two-unit hall through its physical rows is what one such test is, and it is not redrawn here. The rows of the fleet act on g directly.

The record. Table 9.1 gives the record the chapter reads. Its first rows are coil 0's three tests, at ages one to three; the rest summarize the fleet, the number of coils with each number of tests and with each age at the end of the record. Each coil has one test for each of its last $\min(\text{age}, 5)$ years, at whole years of age, 192 tests in all, each read with a standard deviation $\sigma_\mathrm{test} = 0.04$. No coil has a test at age 0: the fleet record holds no commissioning test.

Table 9.1: The fleet record: coil 0's annual condition tests of g, and the fifty coils by number of tests and by age at the end of the record.

Coil 0 (the hall's coil 1), age 3						
Age [yr]	1	2	3			
Condition test of g [–]	−0.0605	−0.1008	−0.0639			
The fleet, 50 coils, 192 tests						
Tests per coil [–]	1	2	3	4	5	
Coils [–]	3	9	7	5	26	
Age at the end of the record [yr]	1	2	3	4	5	6
Coils [–]	3	9	7	5	4	0
Age at the end of the record [yr]	7	8	9	10	11	12
Coils [–]	2	6	5	2	4	3

All fifty coils' tests are plotted against age in Figure 9.1. The horizontal axis is the coil's age at the test, the vertical axis the test's reading of g; each gray dot is a test of one of coils 1 to 49 and each thin gray segment that coil's own line on the chart, drawn over the ages it was tested at; the red dots and the red segment are coil 0's. The tests fall with age across the fleet, each coil's tests scatter about its own line, and coil 0's line is nearly flat because its middle test is its lowest.

Modeling choices at the graph. Four choices that no question owns make this graph. Each test is taken as one reading of g with standard deviation 0.04, independent of the others: this gives one row per test, and it fails if a site's test rig has an offset of its own, a shared sensor offset of Chapter 3, which would read as that coil's commissioning value. Each coil follows the straight line (9.1): it is Chapter 4's trend rows (4.68) in the limit $\sigma_q \to 0$, with each coil's unread months integrated out as in equation (4.69), and it buys two entries per coil; it fails for a coil whose rate changes, which Chapter 11 takes up, and the checks free the process noise. The fifty coils are taken as exchangeable: before their records are read, nothing distinguishes

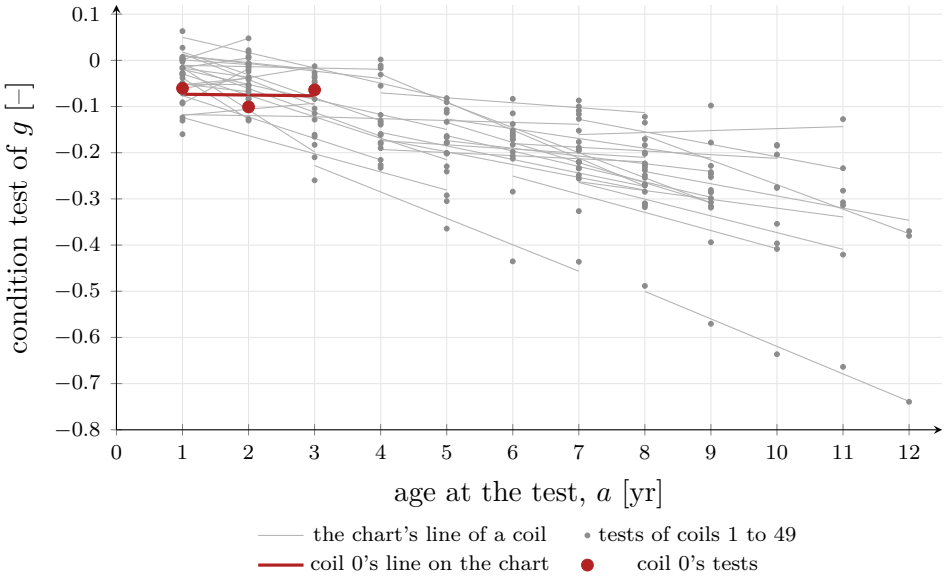


Figure 9.1: The fleet record: the 192 annual condition tests of the fifty coils against age (dots), and the fleet chart's line of each coil with two tests or more, its least-squares line through its own tests taken as exact (segments). Red: coil 0, the hall's coil 1, with tests at ages 1, 2 and 3 and its line of slope $-0.0017\,\mathrm{yr}^{-1}$.

one coil's rate from another's, since they are one model in one duty class. This is what makes one population prior the right prior, and it fails if sites differ in a known way, by water quality or duty, which would enter as a known input of each coil's rate, a covariate this chapter does not use. And the population carries no correlation between a coil's commissioning value and its rate: this keeps the two tie rows of a coil separate, and it fails if poorly built coils also foul faster.

The graph. The rows are drawn in Figure 9.2. The fleet node (green) sits above a row of coil nodes; coil 0 is drawn in full, with its three tests as dashed readings below it, and coils 1 to 49 as one dashed box with their tests. Black squares are the tie rows, two between each coil and the fleet node, each labeled with its residual; the four boxes above the fleet node are its prior rows.

The state. The free entries form the state

$$\mathbf{z} = (g_{0,0},\ \ldots,\ g_{49,0},\ \gamma_0,\ \ldots,\ \gamma_{49},\ \mu_c,\ \mu_\gamma,\ \ell_c,\ \ell_\gamma)^{\mathsf{T}}, \tag{9.2}$$

104 entries, the rates and μ_γ in yr^{-1} and the others dimensionless. Two sub-vectors recur: $\mathbf{x}_i = (g_{i,0},\ \gamma_i)^{\mathsf{T}}$, the two entries of coil i, and $\boldsymbol{\mu} = (\mu_c,\ \mu_\gamma)^{\mathsf{T}}$; the 102 entries other than the spreads are $\mathbf{x}$, and the spreads $\boldsymbol{\ell} = (\ell_c,\ \ell_\gamma)^{\mathsf{T}}$.

9.1.3 The asset manager's questions (step 1)

The asset manager asks five questions of the fleet record, and step 1 writes each as a probability over the state (9.2), with the unknown on the left of the bar and the 192 tests and the priors on the right. Table 9.2 lists them, in the columns of Chapter 5: the question in the asset manager's words, the first the drill question; the general form of its category from Chapter 2; its mathematical form for the fleet; the category; the computation of Section 2.3.6 that answers it; and the subsection where it is answered. "This unit" is coil 0. The first two questions share

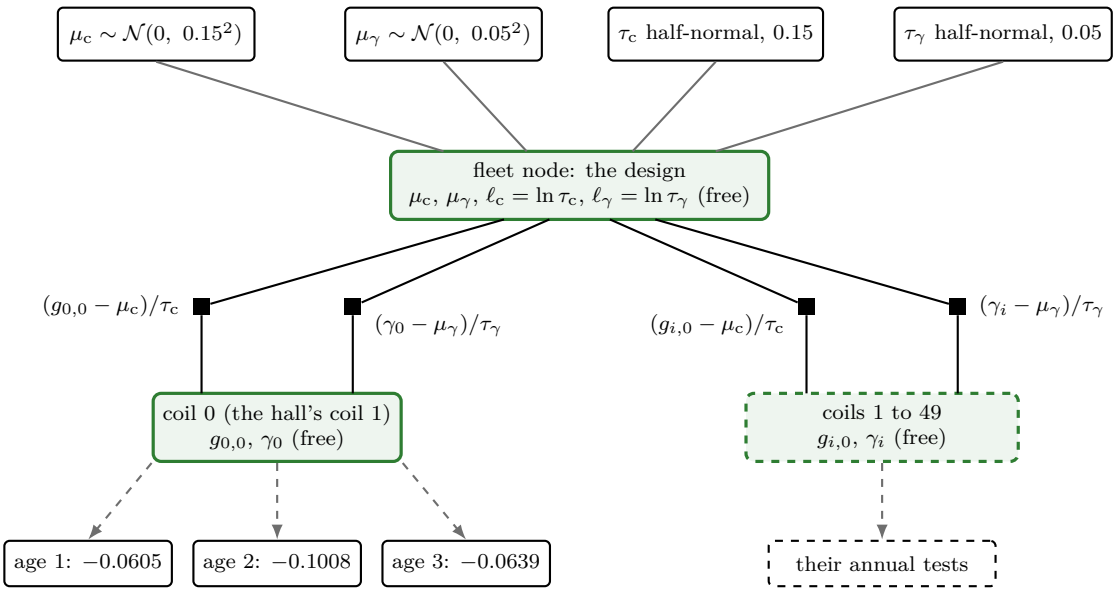


Figure 9.2: The rows of the fleet. Free (green): the fleet node's four entries, the means μ_c and μ_γ of the design's commissioning values and rates and the logarithms ℓ_c and ℓ_γ of their spreads; and each coil's commissioning value $g_{i,0}$ and rate γ_i (coil 0 drawn alone, coils 1 to 49 as one dashed box). Black squares: the tie rows, Gaussian rows whose tolerances τ_c and τ_γ are entries of the state. Observed (dashed): each coil's annual condition tests of g, standard deviation 0.04, coil 0's three with their values. Priors (solid gray): the four prior rows on the fleet node. No quantity is pinned.

one answer block. The rate of the hall's coil from its own monthly record is Chapter 10's; this chapter's rate is the one the fleet's annual tests give. Which coil the crew visits first is Chapter 20's question. Which measurement would place coil 0 better asks nothing new of today's evidence; it is the next measurement for these five questions, Section 9.6.

9.1.4 The evidence as rows (step 2)

Step 2 attaches the record of Table 9.1 to the graph as rows on the state. Each test y_{ij}, the j-th test of coil i, at age a_{ij}, reads the coil's line (9.1). Each coil's two entries are tied to the design's means by the two tie rows, whose standard deviations are the spreads. And four priors carry what was known before the record: $\mu_c \sim \mathcal{N}(0,\ 0.15^2)$, $\mu_\gamma \sim \mathcal{N}(0,\ 0.05^2)$ in yr^{-1}, and half-normal priors of scales 0.15 and $0.05\,\mathrm{yr}^{-1}$ on τ_c and τ_γ, the density $2\varphi(\tau_\bullet/s_{\mathrm{hn}})/s_{\mathrm{hn}}$ on $\tau_\bullet > 0$ for a scale s_{hn}. The rows are

$$\begin{aligned} r_{ij}(\mathbf{z}) &= (g_{i,0} + \gamma_i\, a_{ij} - y_{ij})/\sigma_{\mathrm{test}}, && \text{192 test rows,} \\ r_{c,i}(\mathbf{z}) &= (g_{i,0} - \mu_c)/\tau_c, \quad r_{\gamma,i}(\mathbf{z}) = (\gamma_i - \mu_\gamma)/\tau_\gamma, && \text{100 tie rows,} \\ r_{\mu,c}(\mathbf{z}) &= \mu_c/0.15, \quad r_{\mu,\gamma}(\mathbf{z}) = \mu_\gamma/0.05, && \text{priors on the means,} \end{aligned} \tag{9.3}$$

with $\sigma_{\mathrm{test}} = 0.04$, and the two half-normal priors on the spreads, 296 rows in all. A tie row's tolerance is an entry, so its term in the objective is not its half square alone: by equation (4.39) each row contributes its negative log-probability, and the normal density of $g_{i,0}$ about μ_c with standard deviation τ_c has the term $\frac{1}{2}r_{c,i}^2 + \ell_c$, up to a constant. A half-normal prior carried on $\ell = \ln \tau_\bullet$ has the term $\tau_\bullet^2/(2s_{\mathrm{hn}}^2) - \ell$, the last its change-of-variable factor. With $n_{\mathrm{coil}} = 50$ coils the objective is

$$\mathcal{L}(\mathbf{z}) = \tfrac{1}{2}\,\mathbf{r}(\mathbf{z})^{\mathsf{T}}\mathbf{r}(\mathbf{z}) + n_{\mathrm{coil}}\,(\ell_c + \ell_\gamma) + \frac{\tau_c^2}{2 \times 0.15^2} + \frac{\tau_\gamma^2}{2 \times 0.05^2} - \ell_c - \ell_\gamma + \text{constant}, \tag{9.4}$$

Table 9.2: The questions asked of the fleet record. $\mathbf{y}$ is the 192 tests and the four prior rows; $\mathcal{K}_0$ is coil 0's rank among the fifty by rate (1 the fastest-fouling), π_0 its percentile in the design's population of rates, Φ the standard normal distribution function, and $\hat{\gamma}_0^{\text{chart}} = -0.0017\,\text{yr}^{-1}$ the chart's slope of coil 0.

Asset manager's question	General form	Mathematical form	Category	Computation	Answered in
Where does this unit sit among its fifty siblings?	moments, credible interval, (2.9)–(2.11)	$p(\mathcal{K}_0 \mid \mathbf{y})$, $\mathcal{K}_0 = 1 + \#\{i : \gamma_i < \gamma_0\}$; $p(\pi_0 \mid \mathbf{y})$, $\pi_0 = \Phi\big((\gamma_0 - \mu_\gamma)/\tau_\gamma\big)$: medians and 5–95 % bands	estimation	posterior	9.3.1
Is it fouling faster than the design's typical coil, or as slowly as the chart says?	$\Pr(f(\mathbf{z}) < f^* \mid \mathbf{y})$, (2.12)	$\Pr(\gamma_0 < \mu_\gamma \mid \mathbf{y})$; $\Pr(\pi_0 > 0.9 \mid \mathbf{y})$; $\Pr(\gamma_0 > \hat{\gamma}_0^{\text{chart}} \mid \mathbf{y})$	prediction under a threshold	posterior	9.3.1
How fast is it fouling, and what is its health index today, read with the fleet?	moments, credible interval, (2.9)–(2.11)	$p(\gamma_0 \mid \mathbf{y})$; $d_0(3) = \mathrm{e}^{g_{0,0}+3\gamma_0}$; $h_0(3) = \mathrm{e}^{3\gamma_0}$; beside the separate and fully pooled drawings	estimation	posterior	9.3.2
How fast does this design foul, how alike are its coils, and what does one coil tell the fleet?	moments, credible interval; the predictive, (2.13)	$p(\mu_\gamma, \tau_\gamma, \mu_\text{c}, \tau_\text{c} \mid \mathbf{y})$; $p(\gamma_\text{new} \mid \mathbf{y})$; the fleet with coil 0 left out	estimation, prediction	posterior	9.3.3
Are the coils alike enough to share one line, or different enough to need their own?	$B_{AB} = p(\mathbf{y} \mid M_A)/p(\mathbf{y} \mid M_B)$, (2.14)	$\ln p(\mathbf{y} \mid M)$ of the separate, hierarchical and fully pooled drawings	hypothesis comparison	evidence	9.3.4

with $\mathbf{r}$ the Gaussian rows of equation (9.3), the tests, the tie rows and the priors on the means.

Which row acts on which entry. A test row of coil i touches only $\mathbf{x}_i = (g_{i,0},\ \gamma_i)$; a tie row touches its coil's entry, its mean and its spread; each prior row touches one entry of the fleet node. With the columns grouped by coil, $(\mathbf{x}_0, \ldots, \mathbf{x}_{49}, \boldsymbol{\mu}, \boldsymbol{\ell})$, a reordering of equation (9.2), and each $*$ a block that can be nonzero, the Jacobian has the pattern of an arrow,

$$\mathbf{J} = \begin{bmatrix} * & & & & \mathbf{0} & \mathbf{0} \\ & * & & & \mathbf{0} & \mathbf{0} \\ & & \ddots & & \vdots & \vdots \\ & & & * & \mathbf{0} & \mathbf{0} \\ * & & & & * & * \\ & * & & & * & * \\ & & \ddots & & \vdots & \vdots \\ & & & * & * & * \\ & & & & * & \mathbf{0} \\ & & & & \mathbf{0} & * \end{bmatrix} \begin{array}{l} \text{coil 0's tests, } k_0 \times 2 \\ \text{coil 1's tests} \\ \\ \text{coil 49's tests} \\ \text{coil 0's two tie rows} \\ \text{coil 1's two tie rows} \\ \\ \text{coil 49's two tie rows} \\ \text{priors on } \mu_{\mathrm{c}}, \mu_\gamma \\ \text{priors on } \ell_{\mathrm{c}}, \ell_\gamma \end{array} , \tag{9.5}$$

$$\mathbf{g}_{\pi_0} = \frac{\partial \pi_0}{\partial \mathbf{z}}: \quad \frac{\partial \pi_0}{\partial \gamma_0} = \frac{\varphi(\tilde{\gamma}_0)}{\tau_\gamma}, \quad \frac{\partial \pi_0}{\partial \mu_\gamma} = -\frac{\varphi(\tilde{\gamma}_0)}{\tau_\gamma}, \quad \frac{\partial \pi_0}{\partial \ell_\gamma} = -\tilde{\gamma}_0\, \varphi(\tilde{\gamma}_0), \quad \tilde{\gamma}_0 = \frac{\gamma_0 - \mu_\gamma}{\tau_\gamma},$$

with k_i the number of coil i's tests, and every other entry of $\mathbf{g}_{\pi_0}$ zero. Each test block has the lines $(1,\ a_{ij})/0.04$; coil 0's three are $(25.0,\ 25.0)$, $(25.0,\ 50.0)$ and $(25.0,\ 75.0)$. Each tie block is $\mathrm{diag}(1/\tau_{\mathrm{c}},\ 1/\tau_\gamma)$ in its coil's columns, minus the same in the columns of $\boldsymbol{\mu}$, and $-\mathrm{diag}(r_{\mathrm{c},i},\ r_{\gamma,i})$ in the columns of $\boldsymbol{\ell}$. Of the 296×104 entries of $\mathbf{J}$, 688 can be nonzero, 2.23 %. The second line is not a row: it gives the gradient of the percentile π_0 of the first question, which reaches three columns, those of γ_0, μ_γ and ℓ_γ. The rank $\mathcal{K}_0$ has no gradient: it is a count, constant between the states at which two coils' rates cross, and it is read from draws of the state.

Two facts of the graph are read here, before any number. The coils meet only in the four columns of the fleet node: given those four entries, the fifty coils are independent of one another. And once the spreads are fixed, every row of equation (9.3) is linear in the other 102 entries; the posterior of step 3 is built on that.

Modeling choices at the rows. Four choices make these rows. The maker's rating moves from each coil to the design's mean. Chapter 5 put the rating prior $\mathcal{N}(0,\ 0.15^2)$ on each coil's commissioning value; here it describes the design, so it sits on μ_{c}, and each coil's commissioning value is tied to the design by τ_{c}; the checks vary it. The maker makes no statement on fouling: $\mu_\gamma \sim \mathcal{N}(0,\ 0.05^2)$ puts the design's rate within ± 10 %/yr at two standard deviations, and the checks triple it. The spreads have half-normal priors, of scales 0.15, the rating's spread, and $0.05\,\mathrm{yr}^{-1}$: a flat prior on $\ell = \ln \tau_\bullet$ would leave the posterior improper, because as $\tau_\bullet \to 0$ the evidence of the record levels off and a flat density on ℓ keeps its mass all the way down; the checks narrow and widen each scale about threefold. And the tie rows carry the population, not a chosen tolerance: fixed at zero they give the fully pooled drawing, removed they give the separate drawing, and in between the record sets them, where the pumps of *Reliability Engineering by Derivation*, Chapter 6, had a tolerance chosen by the engineer.

9.2 The posterior of the state (step 3)

All five questions of Table 9.2 are read from one posterior, that of the state (9.2) under the rows of step 2. The second fact of step 2, that the rows are linear in all but two entries once those two are fixed, sets the form it is computed in.

9.2.1 The posterior given the spreads

The quantity owed is $p(\mathbf{z} \mid \mathbf{y})$. By the product rule it splits into the posterior of the 102 entries $\mathbf{x}$ given the spreads and the posterior of the spreads,

$$p(\mathbf{z} \mid \mathbf{y}) = p(\mathbf{x} \mid \boldsymbol{\ell}, \mathbf{y})\; p(\boldsymbol{\ell} \mid \mathbf{y}), \tag{9.6}$$

and each factor is derived in turn.

The rows at fixed spreads. With $\boldsymbol{\ell}$ fixed, the test rows, the tie rows and the priors on the means of equation (9.3) are linear Gaussian rows on $\mathbf{x}$, the rows of Chapter 3 (Section 3.5.3). Their residual vector is $\mathbf{r}(\mathbf{x}) = \mathbf{J}\mathbf{x} - \tilde{\mathbf{y}}$, with $\mathbf{J}$ the 102 columns of the pattern (9.5) that belong to $\mathbf{x}$, evaluated at the fixed spreads, and $\tilde{\mathbf{y}}$ the tests divided by σ_{test}, with zeros in the places of the tie rows and the priors. The normal equations (3.27), $(\mathbf{J}^\mathsf{T}\mathbf{J})\,\hat{\mathbf{x}} = \mathbf{J}^\mathsf{T}\tilde{\mathbf{y}}$, reach the minimum in one step from any start, and the covariance $(\mathbf{J}^\mathsf{T}\mathbf{J})^{-1}$ of equation (3.28) is the exact conditional covariance, not an approximation. With every row normalized the information is $\mathbf{J}^\mathsf{T}\mathbf{J}$; with unnormalized rows and their weights it is $\mathbf{J}^\mathsf{T}\mathbf{W}\mathbf{J}$.

The information as an arrow. The pattern fixes the shape of the information before any number. With the columns grouped by coil, a coil's test block gives $\mathbf{J}_i^\mathsf{T}\mathbf{J}_i$ in its own 2×2 diagonal block, with $\mathbf{J}_i$ the $k_i \times 2$ block of lines $(1,\ a_{ij})/\sigma_{\text{test}}$; its two tie rows add $\boldsymbol{\Omega}^{-1}$ to that block, $-\boldsymbol{\Omega}^{-1}$ to its coupling with $\boldsymbol{\mu}$ and $\boldsymbol{\Omega}^{-1}$ to the block of $\boldsymbol{\mu}$, with $\boldsymbol{\Omega} = \operatorname{diag}(\tau_c^2,\ \tau_\gamma^2)$ the population's covariance. So

$$\boldsymbol{\Lambda}_{\mathbf{x}}(\boldsymbol{\ell}) = \mathbf{J}^\mathsf{T}\mathbf{J} = \begin{bmatrix} \boldsymbol{\Lambda}_0 & & & -\boldsymbol{\Omega}^{-1} \\ & \ddots & & \vdots \\ & & \boldsymbol{\Lambda}_{49} & -\boldsymbol{\Omega}^{-1} \\ -\boldsymbol{\Omega}^{-1} & \cdots & -\boldsymbol{\Omega}^{-1} & \boldsymbol{\Lambda}_{\text{pr}} + n_{\text{coil}}\boldsymbol{\Omega}^{-1} \end{bmatrix}, \qquad \begin{aligned} \boldsymbol{\Lambda}_i &= \boldsymbol{\Omega}^{-1} + \mathbf{J}_i^\mathsf{T}\mathbf{J}_i, \\ \boldsymbol{\Lambda}_{\text{pr}} &= \operatorname{diag}(0.15^{-2},\ 0.05^{-2}), \end{aligned} \tag{9.7}$$

and the right side of the normal equations is $\mathbf{J}^\mathsf{T}\tilde{\mathbf{y}} = (\mathbf{q}_0, \ldots, \mathbf{q}_{49}, \mathbf{0})$, with $\mathbf{q}_i = \mathbf{J}_i^\mathsf{T}\mathbf{y}_i/\sigma_{\text{test}}$ and $\mathbf{y}_i$ the tests of coil i. Every coil block touches $\boldsymbol{\mu}$ and no other coil.

The coils eliminated. The marginal of a joint Gaussian is the Schur complement of the entries integrated out (Section 3.5.3). Integrating the fifty coil blocks out of equation (9.7) takes from the block of $\boldsymbol{\mu}$ the term $\boldsymbol{\Omega}^{-1}\boldsymbol{\Lambda}_i^{-1}\boldsymbol{\Omega}^{-1}$ of each coil, and moves the right side by $\boldsymbol{\Omega}^{-1}\boldsymbol{\Lambda}_i^{-1}\mathbf{q}_i$, so that

$$\boldsymbol{\Lambda}_\mu = \boldsymbol{\Lambda}_{\text{pr}} + \sum_{i=0}^{49}(\boldsymbol{\Omega}^{-1} - \boldsymbol{\Omega}^{-1}\boldsymbol{\Lambda}_i^{-1}\boldsymbol{\Omega}^{-1}), \qquad \boldsymbol{\mu} \mid \boldsymbol{\ell}, \mathbf{y} \sim \mathcal{N}\Big(\boldsymbol{\Lambda}_\mu^{-1}\sum_{i=0}^{49}\boldsymbol{\Omega}^{-1}\boldsymbol{\Lambda}_i^{-1}\mathbf{q}_i,\ \boldsymbol{\Lambda}_\mu^{-1}\Big), \tag{9.8}$$

and each coil, given the means, is the conditional of its own block,

$$\mathbf{x}_i \mid \boldsymbol{\mu}, \boldsymbol{\ell}, \mathbf{y} \sim \mathcal{N}\Big(\boldsymbol{\Lambda}_i^{-1}(\mathbf{q}_i + \boldsymbol{\Omega}^{-1}\boldsymbol{\mu}),\ \boldsymbol{\Lambda}_i^{-1}\Big). \tag{9.9}$$

With the means integrated out as well, coil i has the mean $\boldsymbol{\Lambda}_i^{-1}(\mathbf{q}_i + \boldsymbol{\Omega}^{-1}\hat{\boldsymbol{\mu}})$, the covariance $\boldsymbol{\Sigma}_i = \boldsymbol{\Lambda}_i^{-1} + \boldsymbol{\Lambda}_i^{-1}\boldsymbol{\Omega}^{-1}\boldsymbol{\Lambda}_\mu^{-1}\boldsymbol{\Omega}^{-1}\boldsymbol{\Lambda}_i^{-1}$ and the covariance $\boldsymbol{\Lambda}_i^{-1}\boldsymbol{\Omega}^{-1}\boldsymbol{\Lambda}_\mu^{-1}$ with the means, where $\hat{\boldsymbol{\mu}}$ is the mean of equation (9.8). The two equations are the fleet's two directions. The sum in $\boldsymbol{\Lambda}_\mu$ is what the coils tell the fleet, one term per coil. The conditional mean (9.9) is what the fleet tells the coil: its own tests $\mathbf{q}_i$, and the design's means weighted by $\boldsymbol{\Omega}^{-1}$, together, through the coil's own information $\boldsymbol{\Lambda}_i$.

Coil 0 at the spreads' medians. At $\tau_c = 0.03153$ and $\tau_\gamma = 0.01199\,\mathrm{yr}^{-1}$, the medians of the spreads' marginal derived next, coil 0's tests give

$$\mathbf{J}_0^\mathsf{T}\mathbf{J}_0 = \begin{bmatrix} 1875.0 & 3750.0 \\ 3750.0 & 8750.0 \end{bmatrix}, \qquad \mathbf{q}_0 = \begin{bmatrix} -140.7500 \\ -283.6250 \end{bmatrix}, \qquad \mathbf{\Lambda}_0 = \begin{bmatrix} 2880.7 & 3750.0 \\ 3750.0 & 15704.3 \end{bmatrix},$$

the last the tests' information plus $\mathbf{\Omega}^{-1} = \mathrm{diag}(1005.7,\ 6954.3)$, the rate entries in yr and yr^2. The covariance is given here in the blocks of equation (9.7), the rows and columns of coil 0 in the order $(g_{0,0},\ \gamma_0)$ and those of the means in the order $(\mu_c,\ \mu_\gamma)$:

$$\mathbf{\Lambda}_0^{-1} = \begin{bmatrix} 5.0372 \times 10^{-4} & -1.2028 \times 10^{-4} \\ -1.2028 \times 10^{-4} & 9.2399 \times 10^{-5} \end{bmatrix}, \qquad \mathbf{\Lambda}_\mu^{-1} = \begin{bmatrix} 9.0322 \times 10^{-5} & -1.6459 \times 10^{-5} \\ -1.6459 \times 10^{-5} & 8.0238 \times 10^{-6} \end{bmatrix},$$

$$\mathbf{\Sigma}_{i=0} = \begin{bmatrix} 5.4646 \times 10^{-4} & -1.3715 \times 10^{-4} \\ -1.3715 \times 10^{-4} & 9.9592 \times 10^{-5} \end{bmatrix},$$

$$\mathbf{\Lambda}_0^{-1}\mathbf{\Omega}^{-1}\mathbf{\Lambda}_\mu^{-1} = \begin{bmatrix} 5.9523 \times 10^{-5} & -1.5049 \times 10^{-5} \\ -2.1502 \times 10^{-5} & 7.1468 \times 10^{-6} \end{bmatrix},$$

the last with coil 0's entries as rows and the means as columns. With the means integrated, coil 0's mean is $(g_{0,0},\ \gamma_0) = (-0.01292,\ -0.02836)$, with standard deviations 0.02338 and 0.00998 and correlation -0.5879; the means are $\mu_c = -0.00280$ and $\mu_\gamma = -0.03022$, with standard deviations 0.00950 and 0.00283 and correlation -0.6114. The correlations read by sign. Both diagonal-block correlations are negative for one reason: a line through tests at ages 1 to 3, and for the fleet a population of lines with no test at age 0, trades a higher start for a faster fall, and the positive off-diagonal 3750.0 of $\mathbf{J}_0^\mathsf{T}\mathbf{J}_0$ is that trade. A coil moves with the means its tie rows hold it to, so $g_{0,0}$ covaries positively with μ_c and γ_0 with μ_γ, and the two cross entries are negative through the negative correlations within the coil and within the means. The uncertainty of the means adds $\mathbf{\Lambda}_0^{-1}\mathbf{\Omega}^{-1}\mathbf{\Lambda}_\mu^{-1}\mathbf{\Omega}^{-1}\mathbf{\Lambda}_0^{-1}$ to the coil's covariance and takes the standard deviation of γ_0 from 0.00961, its value given the means, to 0.00998.

9.2.2 The marginal of the spreads

The second factor of equation (9.6) is, by Bayes' rule, $p(\boldsymbol{\ell} \mid \mathbf{y}) \propto p(\mathbf{y} \mid \boldsymbol{\ell})\, p(\boldsymbol{\ell})$, with $p(\mathbf{y} \mid \boldsymbol{\ell})$ the evidence (4.18) of the linear fit at fixed spreads, the integral over $\mathbf{x}$ of the rows' densities. The rows are linear and Gaussian, so the volume-times-height form of equation (4.19) is exact here. The tie rows' normalizing constants $(2\pi\tau_\bullet^2)^{-1/2}$ give $-n_{\mathrm{coil}}(\ell_c + \ell_\gamma)$, the Gaussian integral over $\mathbf{x}$ gives $\det(\mathbf{\Lambda}_\mathbf{x})^{-1/2}$, and the block elimination of equation (9.8) factors that determinant as $\det \mathbf{\Lambda}_\mu \prod_i \det \mathbf{\Lambda}_i$. With $\ln\det\mathbf{\Omega} = 2(\ell_c + \ell_\gamma)$ the terms in ℓ combine, and

$$\ln p(\boldsymbol{\ell} \mid \mathbf{y}) = -\tfrac{1}{2}\,\Psi(\boldsymbol{\ell}) - \tfrac{1}{2}\sum_{i=0}^{49} \ln\det(\mathbf{\Omega}\mathbf{\Lambda}_i) - \tfrac{1}{2}\ln\frac{\det\mathbf{\Lambda}_\mu}{\det\mathbf{\Lambda}_{\mathrm{pr}}} + \ln p(\boldsymbol{\ell}) + \text{constant}, \tag{9.10}$$

with $\Psi(\boldsymbol{\ell}) = \mathbf{r}^\mathsf{T}\mathbf{r}$ at the conditional minimum $\hat{\mathbf{x}}(\boldsymbol{\ell})$ and $\ln p(\boldsymbol{\ell})$ the two half-normal terms with their change-of-variable factors. The determinant terms are the volume of the linear entries' conditional posterior, and the tie rows' $n_{\mathrm{coil}} \ln \tau_\bullet$ terms have joined them as $\det(\mathbf{\Omega}\mathbf{\Lambda}_i) = \det(\mathbf{I} + \mathbf{\Omega}\mathbf{J}_i^\mathsf{T}\mathbf{J}_i)$, which tends to one as the spreads shrink.

The grid. The two spreads are integrated on a 430×200 grid in $(\ell_c,\ \ell_\gamma)$, with τ_c from 10^{-8} to 0.300 and τ_γ from 10^{-4} to $0.100\,\mathrm{yr}^{-1}$, both at nearly one spacing in ℓ. Each node $\boldsymbol{\ell}_m$ has the weight ϖ_m of equation (9.10), normalized over the grid, and the posterior of the other 102 entries is the mixture of the exact conditionals,

$$p(\mathbf{x} \mid \mathbf{y}) = \sum_m \varpi_m\, \mathcal{N}\big(\mathbf{x};\ \hat{\mathbf{x}}(\boldsymbol{\ell}_m),\ \mathbf{\Lambda}_\mathbf{x}(\boldsymbol{\ell}_m)^{-1}\big), \qquad \varpi_m = \frac{p(\boldsymbol{\ell}_m \mid \mathbf{y})}{\sum_{m'} p(\boldsymbol{\ell}_{m'} \mid \mathbf{y})}, \tag{9.11}$$

exact up to the grid. The commissioning spread's axis reaches far down because the marginal falls only in proportion to τ_c toward zero, the hyperprior's change-of-variable factor: the mass below $\tau_c = 10^{-4}$ is 2.22×10^{-4} and below 10^{-6} it is 2.18×10^{-6}, so what lies below the grid's 10^{-8} is negligible.

The normal on the spreads' marginal. A second route keeps the exact conditionals and replaces only the grid. It fits a normal to $\ln p(\boldsymbol{\ell} \mid \mathbf{y})$ at its maximum, with the covariance the inverse of the negative Hessian of the two entries, and mixes the conditionals over the 15×15 nodes of the Gauss–Hermite rule of Chapter 2 on that normal's axes, in place of the grid's nodes in equation (9.11). The normal has its mode at $\tau_c = 0.03356$ and $\tau_\gamma = 0.01165$, standard deviations 0.2561 and 0.1737 on ℓ_c and ℓ_γ, and a correlation of -0.2911. It is the book's usual Laplace order applied where it can work, to the two entries that the linear rows do not fix. The spreads fixed at the marginal's mode, the plug-in route of empirical Bayes, is one node with all the weight.

The numbers. Table 9.3 gives the four fleet entries. Each row is an entry; the columns give its median and 5–95 % band from the mixture, and the same from the normal on the spreads' marginal. The marginal of the spreads has its mode at $\tau_c = 0.03300$ and $\tau_\gamma = 0.01162\,\mathrm{yr}^{-1}$, and the correlation of ℓ_c and ℓ_γ is -0.3349: the scatter of a coil's tests about the fleet's line at the ages it was tested can be carried by a spread of commissioning values or by a spread of rates, and the record knows their sum better than their split. The record separates the rates' spread from zero, $\Pr(\tau_\gamma < 0.005 \mid \mathbf{y}) = 1.54 \times 10^{-5}$. It barely separates the commissioning spread from zero, $\Pr(\tau_c < 0.01 \mid \mathbf{y}) = 0.0295$, because no coil has a test at age 0 and every commissioning value is an extrapolation.

Table 9.3: The four entries of the fleet node: medians and 5–95 % bands from the mixture (9.11) (reported), with the normal on the spreads' marginal beside them.

Entry [unit]	Median [unit]	5 % [unit]	95 % [unit]	Normal on the spreads, median [unit]	Normal, 5 % [unit]	Normal, 95 % [unit]
μ_c, mean commissioning value [–]	−0.00281	−0.01882	0.01289	–	–	–
μ_γ, mean rate [yr^{-1}]	−0.03022	−0.03488	−0.02537	−0.03021	−0.03482	−0.02544
τ_c, commissioning spread [–]	0.03153	0.01361	0.04752	0.03356	0.02202	0.05114
τ_γ, spread of rates [yr^{-1}]	0.01199	0.00887	0.01600	0.01165	0.00875	0.01550

9.2.3 The funnel and the two joint modes

The usual route of this book is the normal at the mode of the whole state, Chapter 4's Laplace approximation. On this state it does not exist in one parametrization and misleads in the other, and the two failures show what equation (9.10) adds. Minimizing the objective (9.4) over the 102 linear entries at fixed spreads gives the *centered profile* $\mathcal{L}_{\mathrm{cen}}$. In the non-centered parametrization, $g_{i,0} = \mu_c + \tau_c\,\varepsilon_{c,i}$ and $\gamma_i = \mu_\gamma + \tau_\gamma\,\varepsilon_{\gamma,i}$ with the standardized deviations $\varepsilon_{c,i}$ and $\varepsilon_{\gamma,i}$ as entries, each tie row becomes the row ε with no $\ln \tau_\bullet$ term, the change of variable having absorbed it, and the same minimization gives the *non-centered profile* $\mathcal{L}_{\mathrm{nc}}$. At fixed

spreads both minimize the same sum of squares, so

$$\begin{aligned} \mathcal{L}_{\mathrm{nc}}(\boldsymbol{\ell}) &= \tfrac{1}{2}\,\Psi(\boldsymbol{\ell}) - \ell_{\mathrm{c}} - \ell_{\gamma} + \frac{\tau_{\mathrm{c}}^2}{2\times 0.15^2} + \frac{\tau_{\gamma}^2}{2\times 0.05^2}, \\ \mathcal{L}_{\mathrm{cen}}(\boldsymbol{\ell}) &= \mathcal{L}_{\mathrm{nc}}(\boldsymbol{\ell}) + n_{\mathrm{coil}}\,(\ell_{\mathrm{c}} + \ell_{\gamma}), \\ -\ln p(\boldsymbol{\ell} \mid \mathbf{y}) &= \mathcal{L}_{\mathrm{nc}}(\boldsymbol{\ell}) + \tfrac{1}{2}\sum_{i=0}^{49} \ln\det(\boldsymbol{\Omega}\boldsymbol{\Lambda}_i) + \tfrac{1}{2}\ln\frac{\det\boldsymbol{\Lambda}_{\mu}}{\det\boldsymbol{\Lambda}_{\mathrm{pr}}} + \text{constant}, \end{aligned} \tag{9.12}$$

the last from equation (9.10).

The centered profile. It has no minimum. On a grid of τ_{c} from 10^{-6} to 2 and τ_{γ} from 10^{-6} to $1\,\mathrm{yr}^{-1}$, the centered profile has no interior local minimum: its least value is at the grid's corner, $\tau_{\mathrm{c}} = \tau_{\gamma} = 10^{-6}$. Along $\tau_{\gamma} = 0.01199$, its median, it takes the values -293.5825, -392.6490, -505.2366, -618.0609 and -730.8875 at $\tau_{\mathrm{c}} = 3\times10^{-2}$, 3×10^{-3}, 3×10^{-4}, 3×10^{-5} and 3×10^{-6}, falling by 112.8266 in the last decade, against $(n_{\mathrm{coil}} - 1)\ln 10 = 112.8267$: the 50 tie rows' $\ln\tau_{\mathrm{c}}$ terms less the hyperprior's change-of-variable term, while the evidence $\ln p(\mathbf{y} \mid \boldsymbol{\ell})$ levels off, from 283.0186 to 279.2926. Along $\tau_{\mathrm{c}} = 0.03153$ it falls as well, to -308.0813, -392.1507 and -504.5188 at $\tau_{\gamma} = 3\times10^{-3}$, 3×10^{-4} and $3\times10^{-5}\,\mathrm{yr}^{-1}$. This is the funnel of Section 4.1.3: the density of the state grows without bound as a spread shrinks with its coils pulled to the mean, because the tie rows' residuals stay finite while their normalizing terms ℓ go to minus infinity.

The non-centered joint mode. The non-centered joint objective has a mode, and it lies far out: Newton's iteration from the grid's minimum of its profile reaches it in 5 iterations, with a gradient of 1.6×10^{-12} and Hessian eigenvalues from 0.9994 to 4.499×10^{6}, at $\tau_{\mathrm{c}} = 0.31232$ and $\tau_{\gamma} = 0.10146\,\mathrm{yr}^{-1}$, with $\mu_{\mathrm{c}} = -0.01467$ and $\mu_{\gamma} = -0.02176\,\mathrm{yr}^{-1}$. The spreads are 9.9 and 8.5 times the marginal's medians, set by the hyperpriors' scales, 0.15 and 0.05, which alone hold them in: the $\ln\tau_{\bullet}$ terms left with the change of variable, and nothing in $\mathcal{L}_{\mathrm{nc}}$ penalizes a spread wide enough to let every coil fit its own tests.

The missing term. On the same wide grid the minimum of $-\ln p(\boldsymbol{\ell} \mid \mathbf{y})$ sits between the two, at $\tau_{\mathrm{c}} = 0.0328$ and $\tau_{\gamma} = 0.0119$, the grid node beside the marginal's mode (0.03300, 0.01162). The term that each joint objective lacks is the evidence's $\frac{1}{2}\sum_i \ln\det(\boldsymbol{\Omega}\boldsymbol{\Lambda}_i)$ of equation (9.12). It is zero as the spreads shrink and grows as they widen, so it is the penalty for a wide population that the non-centered profile misses; the centered profile carries the tie rows' $n_{\mathrm{coil}}(\ell_{\mathrm{c}} + \ell_{\gamma})$ in its place, a term that falls without bound. Eliminating the linear entries exactly puts the missing term back.

The three objects are plotted in Figure 9.3. Both axes are logarithmic: the horizontal axis is τ_{c}, the vertical axis τ_{γ}. The three closed curves are the contours of the marginal (9.10) on the grid that enclose 50, 90 and 99 % of its mass; the dot is its mode; the dashed curves are the contours of the normal on the spreads' marginal of Section 9.2.2 that enclose 90 and 99 % of its mass; the square is the non-centered joint mode; the arrow points along the centered profile's descent, toward the corner where it has no minimum. The marginal's contours reach far toward a small commissioning spread, along the horizontal axis, and not toward a small spread of rates; the normal's contours stops short of that tail; and the non-centered mode lies outside every contour, up and to the right.

9.2.4 What the coils and the fleet tell each other

Which tie row moves coil 0. At the spreads' medians, with both tie rows, coil 0's rate is $-2.84\,\%/\mathrm{yr}$, against the chart's $-0.17\,\%/\mathrm{yr}$. With the rate tied and the commissioning value freed ($\tau_{\mathrm{c}} \to 10^{3}$) it is $-2.48\,\%/\mathrm{yr}$; with the commissioning value tied and the rate freed

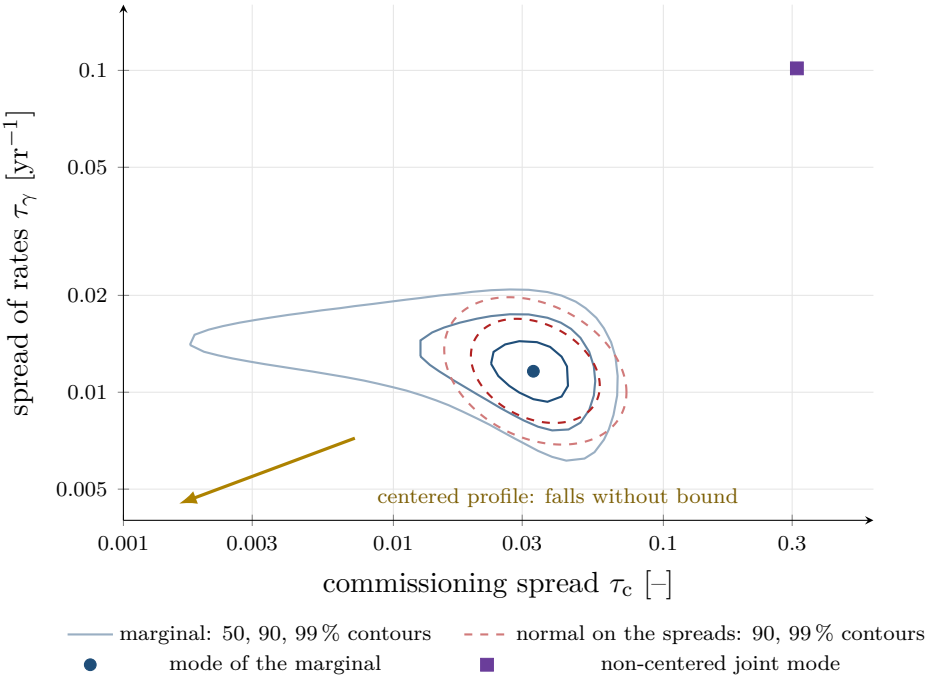


Figure 9.3: The marginal of the two spreads, equation (9.10), on logarithmic axes. Blue: the contours of the grid marginal enclosing 50, 90 and 99 % of its mass (darker inside), and its mode. Red dashed: the contours enclosing 90 and 99 % of the normal fitted to $\ln p(\boldsymbol{\ell} \mid \mathbf{y})$ at its mode, drawn as $\ell_c = \ln 0.03356 + n_\sigma\,(0.2561)\cos\phi$ and $\ell_\gamma = \ln 0.01165 + n_\sigma\,(0.1737)\big(-0.2911\cos\phi + (1-0.2911^2)^{1/2}\sin\phi\big)$ for ϕ around the circle, with $n_\sigma = (-2\ln(1-p))^{1/2} = 2.146$ and 3.035 for $p = 0.90$ and 0.99. Purple square: the non-centered joint mode. Gold arrow: the direction in which the centered profile falls without bound, toward its least value at the corner $\tau_c = \tau_\gamma = 10^{-6}$, off the axes.

($\tau_\gamma \to 10^3$) it is $-2.11\,\%$/yr. Either tie alone carries most of the move from the chart's value to the value with both: the rate tie 86.6 % of it, the commissioning tie 72.9 %, with nothing said to coil 0 about rates. The fleet says coil 0 was commissioned near the design's mean, not at its own line's -0.0717, so its low first test is already fouling. The commissioning tie moves the rate through the correlation -0.5879 within coil 0's block: a higher commissioning value, at the same three tests, is a faster fall.

The information cap. Coil i's term in the sum of equation (9.8) is its information on the design's means,

$$\boldsymbol{\Omega}^{-1} - \boldsymbol{\Omega}^{-1}\boldsymbol{\Lambda}_i^{-1}\boldsymbol{\Omega}^{-1} \preceq \boldsymbol{\Omega}^{-1}, \qquad (\boldsymbol{\Omega}^{-1} - \boldsymbol{\Omega}^{-1}\boldsymbol{\Lambda}_i^{-1}\boldsymbol{\Omega}^{-1})_{\gamma\gamma} < \frac{1}{\tau_\gamma^2}, \tag{9.13}$$

since $\boldsymbol{\Omega}^{-1}\boldsymbol{\Lambda}_i^{-1}\boldsymbol{\Omega}^{-1}$ is positive definite for any record. However long a coil's record, it cannot tell the fleet's mean rate more than $1/\tau_\gamma^2$: the coils are exchangeable draws of the population, and one draw says at most one draw's worth about where the population is centered. At the medians the cap is 6954.3. Coil 0 gives the fleet 2485.7 of information on μ_γ, 35.7 % of the cap; a five-test coil, coil 4, gives 6412.6, 92.2 %; a one-test coil, coil 6, gives 365.2, 5.3 %. All fifty give 198 618.9, against 347 713.7 for fifty coils at the cap, and the prior adds 400.0. The standard deviation of μ_γ at the medians is 0.00283, and would be 0.00224 if μ_c were known: the two means are correlated -0.6114, because every commissioning value is extrapolated.

Where the spreads' uncertainty matters. Over the mixture the variance of coil 0's rate splits by the law of total variance (3.34) into the variance within a fixed pair of spreads, 9.7602×10^{-5}, and the variance of the conditional mean between the pairs, 1.1319×10^{-6}, 1.1 % of the total. Coil 0's rate depends on the spreads' uncertainty only a little; the spreads matter through their values.

9.2.5 Exact draws and the route of each answer

The mixture (9.11) gives a linear quantity of the state directly, as a mixture of normals. A quantity that is not linear, the rank of a coil or its percentile, is read from draws of the state, and the mixture's form gives exact draws by composition.

1. Draw a grid node $\boldsymbol{\ell}_m$ with probability ϖ_m.
2. Draw the means $\boldsymbol{\mu}$ from their conditional at that node, equation (9.8).
3. Draw each coil's $\mathbf{x}_i$ from its conditional given the node and the means, equation (9.9), each through the Cholesky factor of its 2×2 covariance.
4. Read a probability as the fraction of draws in which the event holds and a quantile from the draws' empirical distribution, with the Monte Carlo standard error by batch means: the draws split into 20 batches, as in Section 5.2.4.

Each draw is an independent draw of the posterior up to the grid, not a step of a chain, and the draws carry the rank because every coil's entries are drawn in the same state. The chapter uses $N_\mathrm{d} = 200\,000$ draws from a generator seeded with 9. The normal on the spreads' marginal of Section 9.2.2 is drawn the same way, its Gauss–Hermite nodes in place of the grid's.

The routes. Where the normal does not represent the state, the route that does is reported, the rule of *Reliability Engineering by Derivation*, Chapter 33, for the funnel. Table 9.4 says which route each answer uses: each row is a kind of answer, and the columns give the reported route and the route beside it.

Table 9.4: The route of each answer of the chapter.

Answer	Reported	Beside it
Linear quantities: a coil's rate, commissioning value, derate and health index; the fleet's means	quantiles of the mixture (9.11)	the normal on the spreads' marginal
The two spreads	the grid marginal (9.10)	the normal on $(\ell_c,\ \ell_\gamma)$
Rank, percentile and their probabilities; a new coil's predictive	composition draws of the mixture	draws of the normal on the spreads' marginal
Evidence of the hierarchical drawing	the grid sum of equation (9.10)	– (separate and pooled drawings in closed form)
The next measurement	simulated outcomes, the spreads' weights updated	the weights held fixed; the within-node part alone
The joint normal of the state	not reported: no mode (centered), spreads 9.9 and 8.5 times the medians (non-centered)	shown failing in step 5

9.3 The answers (step 4)

Every answer is read from the posterior of step 3: a linear quantity of the state from the quantiles of the mixture (9.11), exact up to the grid, and the rank and the percentile from the composition draws of Section 9.2.5. The normal on the spreads' marginal stands beside each value, a median beside a median.

9.3.1 Coil 0 among its siblings

Asset manager's questions. Where does this unit sit among its fifty siblings? Is it fouling faster than the design's typical coil, or as slowly as the chart says? The first asks for coil 0's place; the second whether that place is demonstrably fast or slow, and whether the chart's slope survives.

General form. An estimation question about a derived quantity asks for its posterior distribution, equation (2.9),

$$\Pr(f(\mathbf{z}) \le f^* \mid \mathbf{y}) = \int \mathbf{1}\big[f(\mathbf{z}) \le f^*\big]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z},$$

summarized by its median and its credible interval, equation (2.11), at the levels 0.05 and 0.95; a prediction under a threshold is the tail of equation (2.12), the same integral with the indicator of the event.

Specialized form. The questions are equations (2.9), (2.11) and (2.12) with $\mathbf{z}$ the state (9.2), $p(\mathbf{z} \mid \mathbf{y})$ the mixture (9.11) with the spreads' grid, and f one of

$$\mathcal{K}_0(\mathbf{z}) = 1 + \sum_{i=1}^{49} \mathbf{1}[\gamma_i < \gamma_0], \qquad \pi_0(\mathbf{z}) = \Phi\left(\frac{\gamma_0 - \mu_\gamma}{\tau_\gamma}\right), \tag{9.14}$$

the rank of coil 0 among the fifty, 1 the fastest-fouling, and its percentile, the share of the design's population fouling faster. The rank counts the coils whose rate is faster in the same state, so it uses all fifty coils' entries and carries their correlations through the fleet node. The percentile is coil 0's place in the design's population, not among the fifty drawn; its gradient in

the pattern (9.5) reaches ℓ_γ, whose posterior is not normal, and the rank has none, so both are read from the composition draws,

$$\Pr(f(\mathbf{z}) \le f^* \mid \mathbf{y}) = \lim_{N_\mathrm{d}\to\infty} \frac{1}{N_\mathrm{d}} \sum_{j=1}^{N_\mathrm{d}} \mathbf{1}\big[f(\mathbf{z}^{(j)}) \le f^*\big], \tag{9.15}$$

with $\mathbf{z}^{(j)}$ the draws of Section 9.2.5 and the finite sum's standard error by batch means. One choice enters here: position by the rate, not by today's derate. Coils of different ages are not comparable by derate, because an older coil has had longer to foul; by rate they are. Today's derate is the next answer's. For the second question the events are $\gamma_0 < \mu_\gamma$, faster than the design's typical coil; $\pi_0 > 0.9$, in the design's slowest tenth; and $\gamma_0 > \hat{\gamma}_0^{\mathrm{chart}}$, at least as slow as the chart's slope. The first is linear in the state, $\gamma_0 - \mu_\gamma < 0$, and its probability is the mixture's $\sum_m \varpi_m \, \Phi(-\hat{e}_m/\mathrm{sd}_m)$, with $\hat{e}_m$ and sd_m the mean and standard deviation of $\gamma_0 - \mu_\gamma$ at node m from the covariances of equation (9.9); the other two are read by equation (9.15).

Building blocks. The mixture (9.11) and its composition draws; the rank and the percentile (9.14); the fraction of draws (9.15); for the linear event, the conditional moments of equation (9.9) and Φ.

Coil 0's position. Coil 0's rank has median 28 of 50 and 5–95 % band [6, 48], the median's Monte Carlo standard error 0.08; the normal on the spreads' marginal gives 28 and [6, 48]. Its percentile has median 0.553, standard error 0.0010, and band [0.122, 0.929]; the normal on the spreads gives 0.562 and [0.119, 0.935]. The probability that coil 0 ranks among the ten fastest is 0.140, among the ten slowest 0.203, and at the chart's 39th place or slower 0.251.

Faster or slower than typical. The probability that coil 0 fouls faster than the design's typical coil is $\Pr(\gamma_0 < \mu_\gamma \mid \mathbf{y}) = 0.432$ from the mixture, and 0.434 from the draws (standard error 0.0012): the record gives no evidence either way. The deviation $\gamma_0 - \mu_\gamma$ has median $0.00161\,\mathrm{yr}^{-1}$ and band $[-0.01404,\ 0.01750]$. The probability that coil 0 is in the design's slowest tenth is 0.077 (standard error 0.0005), and the probability that it fouls as slowly as the chart's $-0.17\,\%/\mathrm{yr}$ or more slowly is 0.0047 (standard error 0.0002).

Against the other drawings. Read alone, the separate drawing with the same priors put on the coil itself, coil 0 ranks with median 38 and band [11, 48] among the fifty separate posteriors. On the fully pooled drawing every coil has the fleet's one rate, and no coil has a position. The rank band narrows with the length of the record: its mean width is 44.0 places for the one-test coils, 42.2, 38.4 and 31.2 for two to four tests, and 22.7 for five; coil 0's is 42. The narrowest is coil 15's, two places: rank 2, band [1, 3].

The posterior distribution of coil 0's rank is plotted in Figure 9.4. The horizontal axis is the rank, 1 the fastest-fouling, and each bar the posterior probability of that rank; the shaded region is the 5–95 % band, the dashed line the chart's 39th place, the bracket above the bars the separate drawing's band, and the dotted line the rank of coil 0's stated true rate among the fifty true rates, a check marker for Section 9.4.5. The distribution is nearly flat from the sixth place to the forty-eighth: the record places coil 0 almost anywhere in the fleet except at its very ends.

All fifty coils' rates are set side by side in Figure 9.5, the coils sorted by their pooled median, fastest first. The horizontal axis is the coil's place in that order and the vertical axis the rate; for each coil the thick bar is the pooled 5–95 % band with its median as a dot, the thin bar beside it the separate drawing's band, the open circle the chart's slope and the small cross the stated true rate; coil 0 is in red. The shaded strip is the design's 90 % range of rates at the posterior medians, $\mu_\gamma \mp 1.645\,\tau_\gamma$, from -0.0499 to $-0.0105\,\mathrm{yr}^{-1}$, and the line in it μ_γ. The

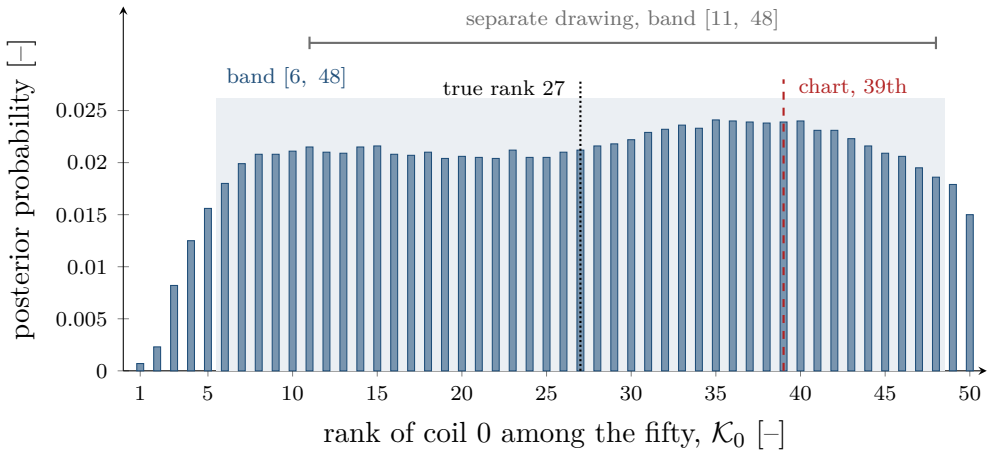


Figure 9.4: The posterior distribution of coil 0's rank among the fifty coils by rate, from the 200 000 composition draws of the mixture. Shaded: the 5–95 % band. Dashed red: the fleet chart's place. Gray bracket: the separate drawing's band. Dotted: the rank of coil 0's stated true rate among the fifty, a check marker (Section 9.4.5).

pooled bands are short and lie inside or near the strip; the separate bands and the chart's slopes scatter far beyond it, most of all for the coils with few tests.

9.3.2 Coil 0's rate and health index

Asset manager's question. How fast is it fouling, and what is its health index today, read with the fleet? Today is coil 0's annual test at age 3.

General form. An estimation question about a derived quantity, equation (2.9),

$$\Pr(f(\mathbf{z}) \le f^* \mid \mathbf{y}) = \int \mathbf{1}[f(\mathbf{z}) \le f^*]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z},$$

summarized by its median, its moments, equation (2.10), and its band, equation (2.11), at the levels 0.05 and 0.95.

Specialized form. The question is equations (2.9), (2.10) and (2.11) with $f(\mathbf{z}) = \mathbf{g}^\mathsf{T}\mathbf{x}_0$ linear in coil 0's block, its gradient $\mathbf{g} = (0,\ 1)$ for the rate γ_0, $(1,\ 0)$ for the commissioning value $g_{0,0}$, and $(1,\ 3)$ for today's log-derate $g_0(3) = g_{0,0} + 3\gamma_0$. Under each node m of the mixture f is normal, with the mean $\mathbf{g}^\mathsf{T}\hat{\mathbf{x}}_{0,m}$ and the variance $\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}_{0,m}\mathbf{g}$, from coil 0's mean and covariance at that node from equation (9.9) with the means integrated, so

$$\Pr(\mathbf{g}^\mathsf{T}\mathbf{x}_0 \le f^* \mid \mathbf{y}) = \sum_m \varpi_m\, \Phi\left(\frac{f^* - \mathbf{g}^\mathsf{T}\hat{\mathbf{x}}_{0,m}}{\sqrt{\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}_{0,m}\mathbf{g}}}\right), \tag{9.16}$$

a quantile is the f^* at which the sum takes its level, found by a one-dimensional root search, and the variance is the mixture's, within nodes and between them. Quantiles pass through the increasing map e^f, so the derate against rating $d_0(3) = \mathrm{e}^{g_0(3)}$ and the health index against commissioning $h_0(3) = \mathrm{e}^{3\gamma_0}$ have as quantiles the exponentials of those of $g_0(3)$ and $3\gamma_0$. The health index is a function of the rate alone, because the commissioning value is an entry of the state and not a test: Chapter 5's joint baseline, one posterior that carries the commissioning value and today, is met here by the state itself. The separate and fully pooled drawings stand beside the answer; both are linear-Gaussian fits of the same tests, in closed form.

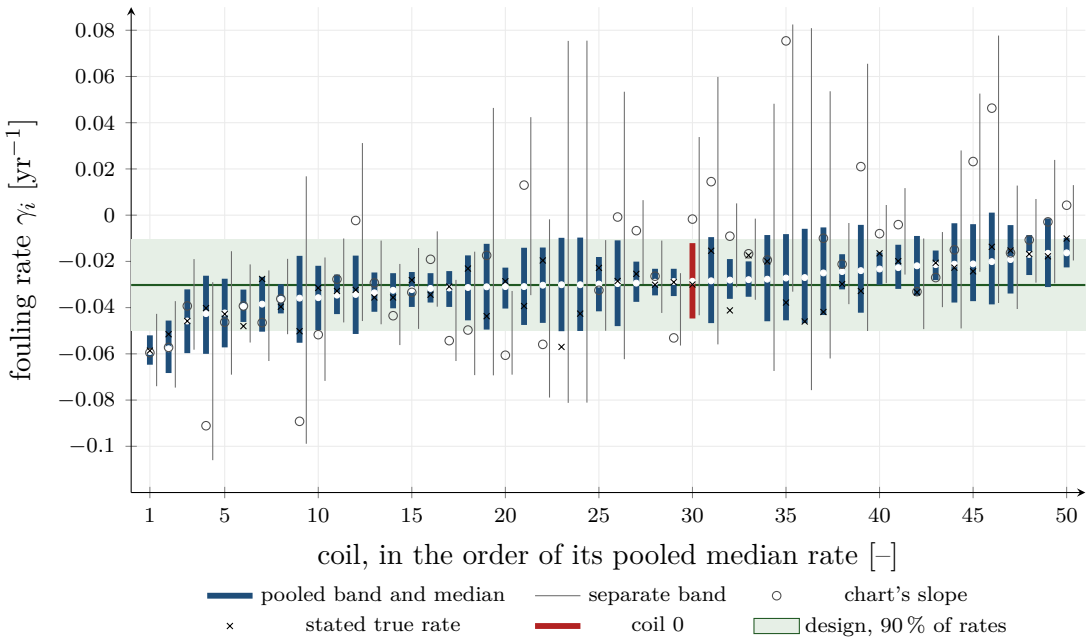


Figure 9.5: The fifty coils' fouling rates, sorted by the pooled median. Thick bars and dots: the pooled 5–95 % bands and medians from the mixture (9.11); coil 0 in red. Thin gray bars: the separate drawing's bands, cut at the axis where they run beyond it. Open circles: the fleet chart's slopes, for the 47 coils with two tests or more. Crosses: the stated true rates, check markers (Section 9.4.5). Green: μ_γ and the design's 90 % range at the posterior medians.

Building blocks. The mixture (9.11); coil 0's moments at each node, equation (9.9); the sum (9.16) and its root; the exponential map.

Coil 0's rate. Read with the fleet, coil 0's rate has median $-0.02861\,\mathrm{yr}^{-1}$ and band $[-0.04471,\ -0.01207]$, a rate of $-2.86\,\%/\mathrm{yr}$, from -4.47 to $-1.21\,\%/\mathrm{yr}$, with standard deviation 0.00994; the normal on the spreads' marginal gives -0.02836 and $[-0.04447,\ -0.01182]$. Its commissioning value has median -0.01158, band $[-0.05129,\ 0.02451]$ and standard deviation 0.02305, and the two are correlated -0.5785 over the mixture.

Today, at age 3. The log-derate has median -0.09795 and band $[-0.13892,\ -0.05671]$, so the derate factor against rating has median 0.9067 and band $[0.8703,\ 0.9449]$, a loss of 9.33 % against rating. The health index against commissioning has median 0.9178 and band $[0.8745,\ 0.9644]$, a loss of 8.22 % against the coil's commissioned value.

Against the other drawings. Read alone, with $\mathcal{N}(0,\ 0.15^2)$ and $\mathcal{N}(0,\ 0.05^2)$ on its own two entries, coil 0's rate is $-0.47\,\%/\mathrm{yr}$, band $[-4.33,\ 3.38]\,\%/\mathrm{yr}$, standard deviation 0.02342, and its commissioning value -0.06407. On the fully pooled drawing, one line for all fifty coils, the rate is $-3.27\,\%/\mathrm{yr}$, band $[-3.42,\ -3.12]$, standard deviation 0.00092, the same for every coil. The fleet informs the unit: the rate's standard deviation falls from 0.02342 to 0.00994, by 57.6 %, and coil 0 keeps 0.065 of its separate departure from the fleet mean, a shrinkage of 0.935. The move runs along the correlation within the coil: read alone, coil 0's commissioning value and rate are correlated -0.8948, and the fleet, by lifting the commissioning value from -0.06407 to the design's neighborhood, lowers the rate with it.

9.3.3 The design and the unit's share in it

Asset manager's question. How fast does this design foul, how alike are its coils, and what does one coil tell the fleet?

General form. The fleet entries are estimation questions, equation (2.9), summarized by medians and bands, equation (2.11); a new coil of the design is a prediction, the predictive (2.13),

$$p(y_{\text{new}} \mid \mathbf{y}) = \int p(y_{\text{new}} \mid \mathbf{z})\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}.$$

Specialized form. The fleet entries are equations (2.9) and (2.11) with f an entry of the fleet node: the means from the mixture (9.11) through the conditional of equation (9.8), the spreads from the grid marginal (9.10). A new coil is equation (2.13) with $y_{\text{new}} = \gamma_{\text{new}}$, the rate of a coil of the design not yet in the fleet, and $p(\gamma_{\text{new}} \mid \mathbf{z}) = \mathcal{N}(\gamma_{\text{new}};\ \mu_\gamma,\ \tau_\gamma^2)$, its tie row read as a density,

$$p(\gamma_{\text{new}} \mid \mathbf{y}) = \int \mathcal{N}(\gamma_{\text{new}};\ \mu_\gamma,\ \tau_\gamma^2)\, p(\mu_\gamma, \ell_\gamma \mid \mathbf{y})\, \mathrm{d}\mu_\gamma\, \mathrm{d}\ell_\gamma, \tag{9.17}$$

drawn by composition, a draw of $(\mu_\gamma,\ \tau_\gamma)$ from Section 9.2.5 and then γ_{new} from the normal, with a generator seeded with 11; the commissioning value of a new coil the same with μ_{c} and τ_{c}. One choice enters here: what one coil tells the fleet is measured by leaving it out, the same fit on forty-nine coils. This is what turns the information of equation (9.13) into a move of the answers, and it measures one coil's effect at the posterior the others give.

Building blocks. The grid marginal (9.10) and the conditional of the means (9.8); the predictive (9.17); the same computation without one coil.

The design. The design's mean rate has median $-3.02\,\%$/yr and band $[-3.49,\ -2.54]$; its spread of rates has median $1.20\,\%$/yr and band $[0.89,\ 1.60]$. The mean commissioning value is -0.00281, band $[-0.01882,\ 0.01289]$, and the commissioning spread 0.03153, band $[0.01361,\ 0.04752]$. A new coil of the design would foul at a median rate of $-3.03\,\%$/yr, with a band from -5.10 to $-0.93\,\%$/yr, and be commissioned at a log-derate of median -0.00260, with a band from -0.05913 to 0.05334. At the posterior medians, 90 % of the design's coils foul at rates between -0.0499 and $-0.0105\,\mathrm{yr}^{-1}$.

What coil 0 tells the fleet. Without coil 0 the design's mean rate is $-0.03047\,\mathrm{yr}^{-1}$, band $[-0.03517,\ -0.02560]$: coil 0 moves μ_γ by $+0.00025$, 0.087 of its standard deviation, moves τ_γ by -0.00003, and narrows the standard deviation of μ_γ from 0.00291 to 0.00290, by 0.62 %, as its share of the cap in Section 9.2.4 leads one to expect.

What a long record tells the fleet. Coil 15 is 12 years old, has five tests, and is the fastest coil by its pooled median, at $-5.90\,\%$/yr, band $[-6.47,\ -5.20]$, rank 2, band $[1,\ 3]$; read alone it is $-0.05834\,\mathrm{yr}^{-1}$, standard deviation 0.00952, and its chart slope is -0.05953. Without coil 15 the design's mean rate is -0.02905, band $[-0.03328,\ -0.02474]$: coil 15 moves μ_γ by -0.00116, -0.402 of its standard deviation. And the spread of rates without it is 0.00953, band $[0.00612,\ 0.01366]$, against 0.01199 with it. One long record of an unusual coil moves the design's spread; coil 0's short one does not.

9.3.4 One line, fifty lines or a population

Asset manager's question. Are the coils alike enough to share one line, or different enough to need their own?

General form. A hypothesis comparison is the ratio of the evidences of two drawings fitted to the same record, equation (2.14),

$$B_{AB} = \frac{p(\mathbf{y} \mid M_A)}{p(\mathbf{y} \mid M_B)}.$$

Specialized form. The question is equation (2.14) for three drawings: the separate drawing M_{S}, fifty independent lines, each with $\mathcal{N}(0,\ 0.15^2)$ and $\mathcal{N}(0,\ 0.05^2)$ on its own two entries; the fully pooled drawing M_{P}, one line $(g_0,\ \gamma)$ for every coil with the same two priors; and the hierarchical drawing M_{H} of this chapter. The first two are linear-Gaussian, and their evidences are equation (4.19), exact. The hierarchical drawing's evidence is the integral over the two spreads of the conditional evidence,

$$p(\mathbf{y} \mid M_{\mathrm{H}}) = \int p(\mathbf{y} \mid \boldsymbol{\ell})\, p(\boldsymbol{\ell})\, \mathrm{d}\boldsymbol{\ell} \approx \sum_m p(\mathbf{y} \mid \boldsymbol{\ell}_m)\, p(\boldsymbol{\ell}_m)\, \Delta\ell_{\mathrm{c}}\, \Delta\ell_{\gamma}, \tag{9.18}$$

the normalizing sum of the grid weights of equation (9.11), with $\Delta\ell_{\mathrm{c}}$ and $\Delta\ell_{\gamma}$ the grid's steps: it costs nothing beyond step 3, and no separate evidence route is needed, as it was for the funnel of *Reliability Engineering by Derivation*, Chapter 33. One choice enters here: the separate drawing puts the hierarchy's priors on the means onto each coil, so that the three drawings differ only in the tie rows. This is what makes the comparison one of pooling and nothing else, and it means that the separate drawing's evidence moves with those priors, which the checks vary.

Building blocks. The closed-form evidences of the separate and pooled drawings; the grid sum (9.18).

The evidences. Table 9.5 sets the three drawings beside the chart for coil 0. Each row is a drawing; the columns give coil 0's rate with its band and standard deviation, its commissioning value, the correlation of the two, and the drawing's log evidence. The log evidences are 228.2136 for the separate drawing, 278.3694 for the hierarchical and −5.8061 for the fully pooled. The hierarchical drawing is preferred to one shared line by $\ln B = 284.18$ and to fifty separate lines by $\ln B = 50.16$. Without coil 15, the hierarchical drawing still beats the pooled one by 136.21. The pooled line's residuals give $\chi^2 = 878.48$ on 190 degrees of freedom: one line cannot carry the fleet.

Table 9.5: Coil 0 under the three drawings and the fleet chart, with the log evidence of each drawing. Bands are 5–95 %; the chart's standard deviation is the standard error of a least-squares slope through three tests, each with standard deviation 0.04.

Drawing	γ_0 [yr^{-1}]	band [yr^{-1}]	sd γ_0 [yr^{-1}]	$g_{0,0}$ [–]	correlation [–]	$\ln p(\mathbf{y} \mid M)$ [–]
Separate, coil 0 alone	−0.00474	−0.04326 to 0.03378	0.02342	−0.06407	−0.8948	228.2136
Hierarchical (reported)	−0.02861	−0.04471 to −0.01207	0.00994	−0.01158	−0.5785	278.3694
Fully pooled, one line	−0.03267	−0.03419 to −0.03115	0.00092	0.00386	−0.8567	−5.8061
Fleet chart, coil 0's line	−0.0017	–	0.0283	−0.0717	–	–

9.4 The checks (step 5)

Each answer of Section 9.3 rests on a grid, on draws, on priors and on the straight lines of step 0, and each of these can be set against the calculation it stands for.

9.4.1 The draws and the grid

The composition draws are exact draws of the mixture, so the only error of a draws answer is the Monte Carlo error, and the linear answers give a test of it. Coil 0's rate from the draws has median -0.02863 (standard error 0.00003) and band $[-0.04472,\ -0.01207]$ (standard errors 0.00004 and 0.00005), against the mixture's -0.02861 and $[-0.04471,\ -0.01207]$ from equation (9.16): the gaps lie within one standard error. The rank's median has standard error 0.08 and the percentile's 0.0010. The grid was tested by its size and its range. With 100, 200 and 300 nodes on the axis of τ_γ, and the axis of τ_c at the matching spacing, coil 0's rate has the same median and band to five decimals, τ_γ the same median, τ_c the median 0.03152, 0.03153 and 0.03154, and the hierarchical log evidence is 278.3694 in all three. A grid with τ_γ down to 10^{-6} and both spreads up to 1.0 and 0.5 gives $\tau_c = 0.03153$ and the same log evidence.

9.4.2 The normal routes against the mixture

Table 9.6 collects the routes. Each row is a route; the columns give coil 0's rate with its band, its percentile and its rank with their bands, and the two spreads. The last row is the stated truth of the record (Section 9.4.5).

Table 9.6: The routes to coil 0's answers and the fleet's spreads, with the stated truth. Bands are 5–95 %.

Route	γ_0 [yr^{-1}]	band [yr^{-1}]	π_0 and band [–]	rank and band [–]	τ_c [–]	τ_γ [yr^{-1}]
Mixture over the spreads (reported)	-0.02861	-0.04471 to -0.01207	0.553, 0.122–0.929	28, 6–48	0.0315	0.01199
Normal on the spreads' marginal, 15×15 nodes	-0.02836	-0.04447 to -0.01182	0.562, 0.119–0.935	28, 6–48	0.0336	0.01165
Spreads fixed at the marginal's mode (empirical Bayes)	-0.02835	-0.04454 to -0.01215	0.565, 0.118–0.935	28, 6–48	0.0330	0.01162
Normal at the non-centered joint mode	-0.00394	-0.04817 to 0.04028	0.570, 0.374–0.749	38, 11–47	0.3123	0.10146
Normal at the centered joint mode	no mode: the objective falls without bound as $\tau_c \to 0$					
Stated truth	-0.0300	–	0.500	27	0.0400	0.0120

The normal on the spreads. It agrees with the mixture on coil 0: the rate -0.02836 against -0.02861, the same rank band, the percentile 0.562 against 0.553, and the design's mean rate -0.03021 against -0.03022. It misses the commissioning spread's tail toward zero: its band of τ_c is $[0.02202,\ 0.05114]$ against the mixture's $[0.01361,\ 0.04752]$, and its $\Pr(\tau_c < 0.01)$ is 0.0000 against 0.0295. That is why it stands beside the mixture and is not reported. The plug-in spreads of empirical Bayes give nearly the same coil answers, -0.02835 and the same rank band, because the spreads' uncertainty is only 1.1 % of the variance of coil 0's rate (Section 9.2.4).

The joint normals. The normal at the non-centered joint mode reports nearly the separate answer, a rate of -0.00394 with a band from -0.04817 to 0.04028 and a rank of 38, because its spreads are 9.9 and 8.5 times the marginal's medians: its draws put τ_c in $[0.20349,\ 0.48328]$ and never below 0.01. Reweighted to the posterior by equation (5.13), its 50 000 draws have an effective sample size of 5 and a largest weight of 0.3461: the full-posterior route of Chapter 5 fails here, and its weighted median rate, 0.00066, is meaningless. The centered joint normal has no mode to sit at.

9.4.3 The limits of the tie rows

The fleet has no physical rows, so the check of the limit of exact rows is made on the tie rows: their two limits are the other two drawings. Table 9.7 gives coil 0's rate with the spreads fixed at common values. Each row is a value of $\tau_c = \tau_\gamma$; the columns give the rate and its standard deviation, and the two reference rows give the drawings the limits should reach. At 10^{-5} coil 0's rate is the fully pooled drawing's to six decimals; at 100, with the tie rows loosened until they say nothing, it is coil 0's own least-squares line, the chart's slope with its standard error. The tie rows interpolate between the two drawings.

Table 9.7: Coil 0's rate with both spreads fixed at a common value, against the fully pooled drawing and coil 0's own least squares.

Spreads $\tau_c = \tau_\gamma$ [–, yr^{-1}]	γ_0 [yr^{-1}]	sd [yr^{-1}]
10^{-5}	−0.032670	0.000923
10^{-3}	−0.032471	0.001403
10^{-2}	−0.031219	0.007972
1	−0.001799	0.028228
10	−0.001701	0.028284
100	−0.001700	0.028284
fully pooled drawing	−0.032670	0.000923
coil 0's own least squares, no prior	−0.001700	0.028284

9.4.4 Sensitivity to the priors and to process noise

Table 9.8 repeats the fit under other hyperpriors. Each row is one prior; the columns give coil 0's rate with its band and its shift from the chapter's median in units of its posterior standard deviation, its percentile and rank with their bands, and the fleet's μ_γ, τ_γ and τ_c. Every row moves coil 0's median by 0.01 standard deviations or less and leaves the rank band at [6, 47] or [6, 48], and τ_γ ranges from 0.01177 to 0.01207: the record sets the spreads. The prior effect on each answer is small for one reason: fifty coils inform the two spreads and the two means, and a hyperprior is one row against them. The evidence against the separate drawing moves with the separate drawing's own priors, as Section 9.3.4 said: $\ln B = 50.16$, 89.69 and 76.59 for coil priors of standard deviations (0.15, 0.05), (0.15, 0.15) and (0.30, 0.05), wider priors spreading the separate drawing's predictions thinner. Against the pooled drawing the hierarchy keeps $\ln B = 283.10$ with τ_γ half-normal of scale 0.15.

Table 9.8: Coil 0's answers and the fleet's entries under other hyperpriors; HN(s_{hn}) is the half-normal prior of scale s_{hn}. The shift is the change of γ_0's median from the chapter's in units of its posterior standard deviation.

Prior	γ_0 [yr^{-1}]	band [yr^{-1}]	shift [sd]	π_0 and band [–]	rank and band [–]	μ_γ [yr^{-1}]	τ_γ [yr^{-1}]	τ_c [–]
τ_γ HN(0.015)	−0.02864	−0.04460, −0.01227	−0.00	0.555, 0.120–0.929	28, 6–48	−0.03025	0.01177	0.03189
τ_γ HN(0.05), the chapter's	−0.02861	−0.04471, −0.01207	+0.00	0.554, 0.122–0.926	28, 6–47	−0.03022	0.01199	0.03153
τ_γ HN(0.15)	−0.02860	−0.04472, −0.01205	+0.00	0.554, 0.122–0.927	28, 6–47	−0.03021	0.01201	0.03150
τ_c HN(0.05)	−0.02868	−0.04478, −0.01218	−0.01	0.551, 0.122–0.925	28, 6–47	−0.03022	0.01207	0.03052
τ_c HN(0.5)	−0.02860	−0.04471, −0.01206	+0.00	0.553, 0.123–0.927	28, 6–47	−0.03021	0.01198	0.03165
$\mu_\gamma \sim \mathcal{N}(0,\ 0.15^2)$	−0.02868	−0.04479, −0.01216	−0.01	0.554, 0.122–0.927	28, 6–47	−0.03030	0.01199	0.03153
$\mu_c \sim \mathcal{N}(0,\ 0.05^2)$	−0.02863	−0.04472, −0.01211	−0.00	0.554, 0.121–0.927	28, 6–47	−0.03023	0.01199	0.03150

The table's percentiles and ranks come from a smaller set of draws of each fit, so the chapter's own row differs from the reported 0.553, [0.122, 0.929] and [6, 48] in the last place; that is Monte Carlo error.

Process noise. The straight line of step 0 is the trend rows of Chapter 4 at $\sigma_q \to 0$. Freed to that chapter's $\sigma_q = 0.005$ per month, 0.0173 per square root of a year, about each coil's line, the fit gives coil 0's rate −0.02968, band [−0.04509, −0.01386], standard deviation 0.00951, against −0.02861 and 0.00994; the spread of rates 0.01057, the commissioning spread 0.02346, and $\mu_\gamma = -0.03047$. Coil 0's answer moves little, while both spreads shrink: part of what the straight lines call spread is wander about the line.

9.4.5 The answers against the stated truth

The last check sets the answers against the stated truth of the record (Section 9.6.5), the last row of Table 9.6. Coil 0's true rate is −0.0300, and the posterior median lies +0.14 standard deviations from it, inside the band; its true commissioning value is −0.0200, and the median lies +0.37 standard deviations from it. Its true derate today is 0.8958 and its true health index 0.9139, both inside their bands. Among the fifty true rates coil 0 ranks 27th, and its true percentile in the design is 0.500; the posterior's 28 and 0.553 are close, and the bands, [6, 48] and [0.122, 0.929], say how little that closeness is worth. The design's stated mean rate is −0.0300, its spread of rates 0.0120 and its commissioning spread 0.0400; the fifty coils drawn from it have −0.0316, 0.0117 and 0.0367. Over the fifty coils the root-mean-square error of the rate is 0.00844 for the pooled medians, 0.02311 for the separate drawing, 0.02970 for the chart's slopes over its 47 coils, and 0.01166 for the one pooled line; 92 % of the true rates lie inside the pooled 5–95 % bands.

9.5 The report (step 6)

The chart's "−0.17 %/yr, 39th of 47" is not coil 0's position. Read with its siblings, coil 0 fouls at −2.86 %/yr, from −4.47 to −1.21 %/yr, the design's typical rate, and the chart's rate has a probability of 0.005. Its place among the fifty is 28th, with a band from 6th to 48th: three tests over two years cannot place a coil in a fleet whose rates differ by 1.2 %/yr, and coil 0 is neither demonstrably fast nor slow, faster than the design's typical coil with probability 0.43. Either of coil 0's tie rows alone carries most of the move from the chart's value, the rate tie 87 % and the commissioning tie 73 %: the line through coil 0's tests put its commissioning value at −0.0717, and the fleet puts it near the design's mean. And the coils are a population: one shared line is rejected by 284 nats, fifty separate lines by 50.

9.5.1 The answers and their bands

Table 9.9 gives the answers in the order of Table 9.2. Each row is a question; the second column gives the reported value, from the mixture over the spreads; the third the normal on the spreads' marginal beside it, where it gives one; and the fourth the band, or for a probability its Monte Carlo standard error.

9.5.2 The entry in the fleet register

The fleet's assessment is entered in the fleet register under coil 0, at its annual test at age 3. It uses nothing later than that test, and in the values of Table 9.9 it reads:

> *Coil 0 (the hall's coil 1), age 3, annual condition test; fleet assessment of 50 coils (192 tests).*

Table 9.9: The report for coil 0 and its fleet at coil 0's annual test at age 3. Reported values from the mixture over the spreads; bands are 5–95 %.

Asset manager's question	Reported value	Normal on the spreads	Band or standard error
Where does this unit sit among its fifty siblings?	rank 28 of 50 [–]; percentile 0.553 [–]	28; 0.562 [–]	rank 6 to 48; percentile 0.122 to 0.929 [–]
Is it fouling faster than the design's typical coil, or as slowly as the chart says?	faster than typical, 0.432, exact up to the grid (draws 0.434, se 0.0012) [–]; slowest tenth, 0.077 [–]; as slow as the chart, 0.0047 [–]	0.422; 0.084; 0.0049 [–]	standard errors 0.0005 and 0.0002 on the last two [–]
How fast is it fouling, and what is its health index today, read with the fleet?	$\gamma_0 = -0.02861$ [yr^{-1}]; derate 0.9067 [–]; health index 0.9178 [–]	$\gamma_0 = -0.02836$ [yr^{-1}]; 0.9063; 0.9184 [–]	-0.04471 to -0.01207 [yr^{-1}]; 0.8703 to 0.9449; 0.8745 to 0.9644 [–]
How fast does this design foul, how alike are its coils, and what does one coil tell the fleet?	$\mu_\gamma = -0.03022$, $\tau_\gamma = 0.01199$, a new coil -0.03028 [yr^{-1}]; coil 0 moves μ_γ by 0.087 sd [–]	-0.03021, 0.01165 [yr^{-1}]	-0.03488 to -0.02537; 0.00887 to 0.01600; -0.05095 to -0.00935 [yr^{-1}]
Are the coils alike enough to share one line, or different enough to need their own?	a population: $\ln B = 284.18$ against one line, 50.16 against fifty [–]	–	against fifty, 50.16 to 89.69 over the separate drawing's priors [–]

Rate: -2.86 %/yr (5–95 % band -4.47 to -1.21); the design's mean -3.02 %/yr, spread 1.20 %/yr.
Position: rank 28 of 50 (band 6–48); faster than the design's typical coil with $P = 0.43$; in the design's slowest tenth with $P = 0.08$.
Today: derate 0.907 against rating (0.870–0.945); health index 0.918 against commissioning (0.875–0.964).
The chart's -0.17 %/yr, 39th of 47: not supported ($P = 0.005$).
Action: coil 0 is not moved to the back of the cleaning queue; its order is set by Chapter 20's ranking, not by its slope.
Next measurement: carry the hall's monthly condition record of coil 1 into the fleet fit.

The entry records what the band rests on as well as the values: three annual tests of coil 0, the forty-nine other coils' tests through the design's mean and spread, and a commissioning value that no test read. Which coil the crew cleans first, and when, is not settled by a rank; it needs the expected consequence of waiting that Chapter 20 computes for each coil.

Common mistake

Ranking units by their own slopes. A short record's slope is mostly its noise, and its intercept is an extrapolation; read with the fleet, the unit's rate and its commissioning value move together, and its position is a band.

9.6 The next measurement (step 7)

The report leaves coil 0's rate between -4.47 and -1.21 %/yr and its rank anywhere from 6th to 48th. A further record is a set of rows on coil 0's two entries, or on other coils', and what it would take off the band is known before its values are.

9.6.1 The preposterior over the spreads

A candidate record is a set of rows reading coil 0's log-derate at ages a_j, each with the Jacobian line $(1,\ a_j)/\sigma_j$ in the columns of $(g_{0,0},\ \gamma_0)$, with $\mathbf{J}_{\text{cand}}$ the stacked lines. At each grid node the

conditional update is exact, the preposterior covariance of equation (4.65) on coil 0's block,

$$\boldsymbol{\Sigma}_{\mathrm{cand},m} = \left(\boldsymbol{\Sigma}_{0,m}^{-1} + \mathbf{J}_{\mathrm{cand}}^{\mathsf{T}}\mathbf{J}_{\mathrm{cand}}\right)^{-1}, \tag{9.19}$$

with $\boldsymbol{\Sigma}_{0,m}$ coil 0's covariance at node m. The spreads' weights also move with the outcome $\tilde{\mathbf{y}}_{\mathrm{cand}}$, $\varpi_m(\tilde{\mathbf{y}}_{\mathrm{cand}}) \propto \varpi_m\, p(\tilde{\mathbf{y}}_{\mathrm{cand}} \mid \boldsymbol{\ell}_m, \mathbf{y})$, so the candidate is valued by the design form (2.16) with U_{p} the posterior variance of γ_0: the outcomes are simulated, 2000 of them from a generator seeded with 16, each from a draw of the state, the weights updated for each, and the variances averaged; the standard error of each resulting standard deviation is at most 0.000002. Two values stand beside it. With the weights held fixed, the variance is the within-node variance of equation (9.19) averaged over the present weights plus the present between-node variance. And the within-node part alone, with the between-node variance dropped, is the most that any update of the weights could take off. The width a candidate takes off is $1 - \mathrm{sd}_{\mathrm{cand}}/\mathrm{sd}$ of γ_0, from the present 0.00994. Chapter 5's fit supplies two standard deviations: the hall's coil 1 at commissioning, $g_{1,0}^{\mathrm{hall}} = -0.02423$ with standard deviation 0.03962, and one monthly window's standard deviation of g, 0.03476.

9.6.2 The candidates

The candidates are of three kinds. More tests of coil 0: one or two more annual tests, or a reference test of half the standard deviation. Coil 0's own records at the hall: its commissioning test, carried in as a row on $g_{0,0}$; the hall's condition monitoring of coil 1, the permanent instrument at the place of the question, over the next twelve monthly windows; and the 36 monthly windows the hall has already logged, at ages 1/12 to 3. And the fleet: one more annual test of each of the other forty-nine coils, and, as a limit, the fleet's means known exactly with the spreads as now. A last limit is coil 0's commissioning value known exactly.

One choice enters with the monthly windows: each is taken as one reading of coil 0's log-derate, independent of the others and of the annual tests, with Chapter 5's standard deviation of one window's g, 0.03476. This is what lets the hall's record be valued without reading its values, which are Chapter 10's. It overstates the record's worth in two ways. That standard deviation is a posterior one, which already carries information that the windows share, through the recirculation fraction that the airflow survey and the commissioning test's readings inform, here counted once per window; and the annual tests may be windows of the same record, counted twice. The widths of the monthly candidates are therefore upper limits of what the hall's record adds, and their ranking above the annual tests, not their size, is what the recommendation rests on.

Table 9.10 ranks the candidates. Each row is a candidate; the columns give the standard deviation of γ_0 after it by simulated outcomes, and the width it takes off by simulated outcomes, with the weights held fixed, and by the within-node part alone. The simulated widths are compared as bars in Figure 9.6: the horizontal axis is the width taken off coil 0's band in percent, each bar a candidate, ranked, and the dashed line the fleet's floor.

The spreads' weights. Updating the weights moves each simulated width away from the fixed-weight value by less than the gap between that value and the within-node part alone, in either direction: a reading on coil 0 tells little about the fleet's spreads, and the spreads' uncertainty is 1.1 % of the variance of coil 0's rate (Section 9.2.4). The candidates rank in the same order in all three columns.

The fleet's floor. The fleet has told coil 0 all it can. A year of tests on every other coil takes 0.8 % off coil 0's band, and the fleet's means known exactly 3.8 %. The fleet's information reaches coil 0 through its two tie rows, whose tolerances are the spreads, and the spreads are already well known; only coil 0's own record narrows its rate further. The commissioning test is

Table 9.10: Candidate records for coil 0's rate, by the preposterior (9.19) over the spreads; widths taken off the present $\mathrm{sd}(\gamma_0 \mid \mathbf{y}) = 0.00994\,\mathrm{yr}^{-1}$.

Candidate	$\mathrm{sd}_{\mathrm{cand}}(\gamma_0)$, simulated [$\mathrm{yr}^{-1}$]	width off, simulated [%]	fixed weights [%]	within-node part alone [%]
One more annual test of coil 0 (age 4), sd 0.04	0.00853	14.2	14.5	15.1
Two more annual tests (ages 4 and 5)	0.00724	27.2	27.3	28.1
A reference test of coil 0 at age 4, sd 0.02	0.00716	27.9	28.2	29.0
Coil 0's commissioning test carried in (sd 0.0396)	0.00951	4.3	4.4	5.0
The hall's condition monitoring of coil 1, next 12 windows (sd 0.0348 a window)	0.00680	31.5	32.0	32.8
The hall's 36 logged windows of coil 1, carried in	0.00554	44.3	44.0	45.0
The logged windows and the commissioning test together	0.00542	45.4	45.1	46.1
One more annual test of each of the other 49 coils	0.00985 (fixed)	–	0.8	–
The fleet's means known exactly, spreads as now (the fleet's floor)	0.00956 (fixed)	–	3.8	–
Coil 0's commissioning value known exactly (a floor)	0.00811	18.4	18.3	19.0

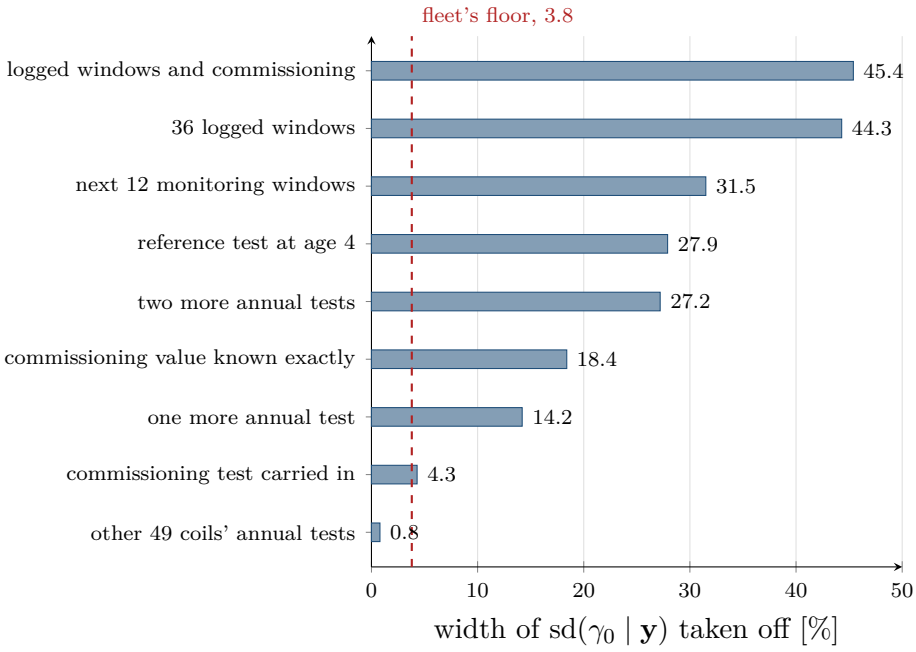


Figure 9.6: The candidate records of Table 9.10, ranked by the width they take off coil 0's rate, by simulated outcomes with the spreads' weights updated (the other coils' tests with the weights held fixed). Dashed: the width taken off if the fleet's means were known exactly, the most the fleet's entries can give coil 0.

worth less here than it was in Chapter 5: the fleet already places $g_{0,0}$ to a standard deviation of 0.02305, the hall's commissioning fit, at 0.0396, adds little, and even an exact commissioning value takes off under a fifth.

9.6.3 The recommendation

The hall's own monthly record of coil 1 is already logged, and carried into the fleet fit as coil 0's rows it takes off the most, at most 44.3 % of the band, and 45.4 % with the commissioning test, upper limits by the choice above. Whether one straight line holds across those windows is the question of Chapters 10 and 11, and this chapter reads only their rows' standard deviation, not their values. Looking ahead, next year's twelve monitoring windows take off at most 31.5 %, against 14.2 % for next year's annual test. A further test of every other coil is not worth ordering for coil 0's sake.

9.6.4 What the analysis showed

The chart said coil 0 fouled at -0.17 %/yr, 39th of 47, among the slowest of its design; read with its siblings it fouls at -2.86 %/yr, 28th of 50 with a band from 6th to 48th. The tie rows with a learned tolerance read each coil with its siblings, and either of coil 0's tie rows alone carries most of the move, the rate tie 87 % and the commissioning tie 73 %: a three-test line misplaces its own start, and the slope pays for it. The joint normal of the state, the book's usual route, fails in both parametrizations, with no mode in one and spreads 9.9 and 8.5 times too wide in the other; eliminating the linear entries exactly leaves two spreads, and their quadrature is the posterior. The fleet then has little more to give: a short record is placed in a fleet only to a wide band, and the band, not the rank, is the answer, until the coil's own record grows.

9.6.5 Limits of the model

A constructed record. The fleet record is generated from a stated truth, with the stated test errors, and step 5 compares the answers with that truth (Section 9.4.5). It is not field data, and the agreement is a check of the method on its own model, not of the model against a real fleet.

Straight lines. Each coil follows one straight line from commissioning. A coil whose rate changes needs the curvature of Chapter 11, and the process noise of step 5 already moves the spreads.

Exchangeable coils. The coils are taken as draws of one population, with no site covariates and no population correlation of commissioning value and rate. A fleet whose sites differ in water quality or duty needs those as known inputs of each coil's rate.

Unbiased tests. Each test is taken as one unbiased reading of g; a site's rig offset would read as its coil's commissioning value.

No failures. The fleet's failure log by age is not read. Whether condition and failures share a cause is the question of Chapters 16 and 20.

The next chapter. Chapter 10 turns to the hall's own monthly record of coil 1 and reads the degradation rate as the trend rows of Chapter 4 across its 36 windows.

9.7 What to retain

Definition 9.1 (Partial pooling by tie rows with a free tolerance)**.** Partial pooling reads the units of a fleet in one state that carries each unit's parameters and the design's means and spreads. Each unit is tied to the means by rows whose standard deviations are the spreads, so that the tie rows' terms in the objective include $\ln \tau_\bullet$. At fixed spreads the rows are linear, the units are eliminated exactly by their Schur complements, $\mathbf{\Lambda}_\mu = \mathbf{\Lambda}_{\mathrm{pr}} + \sum_i (\mathbf{\Omega}^{-1} - \mathbf{\Omega}^{-1}\mathbf{\Lambda}_i^{-1}\mathbf{\Omega}^{-1})$, and the spreads' marginal is the evidence of that linear fit times their prior, integrated on a grid; the posterior is the mixture of the exact conditionals over the grid, with exact draws by composition. A unit cannot tell the fleet's mean more than $1/\tau_\bullet^2$, and the joint normal of the state has no mode in the centered parametrization and the wrong one in the non-centered.

Three conclusions carry beyond the fleet. A unit's position in its fleet is a derived quantity of the whole state, read as a band from draws, and for a short record the band is wide whatever the point estimate says. The fleet informs a unit through every tie row, including the one on its commissioning value, which alone can carry most of the move of the unit's rate. And where the joint normal of a state fails, eliminating the entries the rows are linear in leaves a state small enough to integrate exactly.

Table 9.11 lists the terms of the chapter with their expressions, units and the subsection that defines each.

Table 9.11: The terms of fleet pooling and where each is defined. Units for the fleet of coils.

Term	Expression	Units	Meaning	Defined in
Partial pooling	tie rows $(\gamma_i - \mu_\gamma)/\tau_\gamma$, eq. (9.3), with μ_γ, ℓ_γ entries	–	the fleet informs the unit and the unit the fleet	9.1.1
The population's spread	τ_{c}, τ_γ; $\mathbf{\Omega} = \mathrm{diag}(\tau_{\mathrm{c}}^2, \tau_\gamma^2)$	–, yr^{-1}	the tolerance of the tie rows	9.1.2
Separate and fully pooled drawings	the limits $\tau_\bullet \to \infty$ and $\tau_\bullet \to 0$	–	each unit alone; one shared line	9.1.4
Rows and posterior				
Arrow pattern	eq. (9.5), (9.7)	–	the coils meet only in the fleet node	9.1.4
Shrinkage	$\mathbf{\Lambda}_i^{-1}(\mathbf{q}_i + \mathbf{\Omega}^{-1}\boldsymbol{\mu})$, eq. (9.9)	–	the fleet tells the coil	9.2.1
Marginal of the spreads	eq. (9.10)	–	the evidence of the linear fit times the hyperpriors	9.2.2
Funnel	$\mathcal{L}_{\mathrm{cen}}$, eq. (9.12)	–	no mode as a spread shrinks	9.2.3
Centered and non-centered parametrizations	γ_i against $\mu_\gamma + \tau_\gamma \varepsilon_{\gamma,i}$	–	which joint mode exists, and where	9.2.3
Information cap	eq. (9.13)	yr^2	the most one unit tells the fleet's mean	9.2.4
Composition draws	a node by ϖ_m, then $\boldsymbol{\mu}$, then each $\mathbf{x}_i$	–	exact draws of the mixture	9.2.5
Questions and answers				
Fleet position	$\mathcal{K}_0$, π_0, eq. (9.14)	–	rank among the fifty; percentile in the design	9.3.1
Posterior predictive of a new unit	eq. (9.17)	yr^{-1}	the rate of a coil not yet in the fleet	9.3.3
Evidence of the hierarchy	eq. (9.18)	–	the grid's normalizing sum	9.3.4
Preposterior over the spreads	eq. (9.19)	mixed	per node, with the weights updated	9.6.1

9.8 Exercises

Exercise 9.1. *Formulation.* Write the question "is coil 0 in the design's fastest quarter?" in the form of Table 9.2 and evaluate it from the fit of this chapter. (A prediction under a threshold, equation (2.12) with $f = -\pi_0$ and $f^* = -0.25$, π_0 of equation (9.14), read from the composition draws by equation (9.15): $\Pr(\pi_0 < 0.25 \mid \mathbf{y}) = 0.1526$.)

Exercise 9.2. *Derivation.* Show that a coil's information on μ_γ, $(\mathbf{\Omega}^{-1} - \mathbf{\Omega}^{-1}\mathbf{\Lambda}_i^{-1}\mathbf{\Omega}^{-1})_{\gamma\gamma}$, is below $1/\tau_\gamma^2$ for any record, and evaluate it for coil 0 at the spreads' medians. ($\mathbf{\Lambda}_i = \mathbf{\Omega}^{-1} + \mathbf{J}_i^\top \mathbf{J}_i$ is positive definite, so is its inverse, and so is $\mathbf{\Omega}^{-1}\mathbf{\Lambda}_i^{-1}\mathbf{\Omega}^{-1}$, whose diagonal entries are therefore positive; subtracted from $(\mathbf{\Omega}^{-1})_{\gamma\gamma} = 1/\tau_\gamma^2$ it leaves less. For coil 0, 2485.7 of 6954.3, 35.7 %.)

Exercise 9.3. *Derivation.* Show that the centered joint objective, minimized over the linear entries, falls by $(n_{\text{coil}} - 1)\ln 10$ per decade of τ_{c} as $\tau_{\text{c}} \to 0$, while the evidence levels off. (By equation (9.12), $\mathcal{L}_{\text{cen}} = \frac{1}{2}\Psi + (n_{\text{coil}} - 1)(\ell_{\text{c}} + \ell_\gamma) + \tau_{\text{c}}^2/(2 \times 0.15^2) + \tau_\gamma^2/(2 \times 0.05^2)$; as $\tau_{\text{c}} \to 0$ the tie rows pull every commissioning value to μ_{c}, Ψ tends to the sum of squares with them held there, and the hyperprior term to zero, so the profile falls by $(n_{\text{coil}} - 1)$ per unit of ℓ_{c}; in the evidence (9.10) the $n_{\text{coil}} \ln \tau_\bullet$ terms have joined $\ln\det(\mathbf{\Omega}\mathbf{\Lambda}_i)$, which tends to a constant, $\ln\!\big(1 + \tau_\gamma^2(\mathbf{J}_i^\top \mathbf{J}_i)_{\gamma\gamma}\big)$. On the fleet, 112.8266 in the last decade against 112.8267.)

Exercise 9.4. *Application.* Carry the hall's commissioning fit of coil 1, $g_{1,0}^{\text{hall}} = -0.02423$ with standard deviation 0.03962 (Chapter 5), into coil 0 as a row on $g_{0,0}$ and refit. (The row adds $(1,\ 0)/0.03962$ to $\mathbf{J}_0$; $\gamma_0 = -0.02793\,\text{yr}^{-1}$, band $[-0.04336,\ -0.01215]$, standard deviation 0.00950; $g_{0,0} = -0.01441$, standard deviation 0.01987.)

Exercise 9.5. *Application.* Coil 0's test at age 4 reads -0.2060, two test standard deviations below its median line. Refit and give its rate and its rank. ($\gamma_0 = -0.03634\,\text{yr}^{-1}$, band $[-0.05037,\ -0.02231]$; rank 14, band $[4,\ 38]$.)

Part III

Estimating Degradation Rates

Chapter 10

Degradation Rates from Condition Records

10.1 The dashboard's trend line and the coil's rate

The hall is three years past its handover. Today is December of year 3, window 36, the last window of the condition record, early summer in the southern hemisphere where the hall stands. The condition-monitoring dashboard now holds 37 points per coil: each window's health-index tile, the commissioning approach over that window's approach, and the least-squares line drawn through them. Coil 1's line falls by 15.76 % of the commissioned value per year and coil 2's by 14.82 %. Today's tiles read 0.49 for coil 1, "lost 51.3 %", and 0.59 for coil 2, "lost 40.5 %", and the tiles scatter about each line with a standard deviation of 0.082, coil 1's from −0.10 to +0.26. The conclusion the chart invites is that both coils are losing about 15 % of their conductance a year, coil 1 slightly faster, and that both should be cleaned at the October shutdown of year 4.

The chart cannot say which part of that slope is the coils. The IT load has grown, the hall meter's line rising by 3.21 kW a year, 2.53 % of its mean, and the season swings the other heat and the chilled-water supply. Each tile reads its own coil and, through the return air the two share, the other one; and unit 1's tile reads through the discharge sensor that Chapter 7 left under suspicion. Nor can the chart say how well 37 windows know a rate of a few percent a year, or which coil is losing faster. The question of this chapter is

how fast each coil's conductance is falling, read as the trend entry of one fit of the whole condition record, with its band, and how much of the dashboard's slope is the coils.

It is answered with the method of Chapter 5, on the stacked state of all 37 windows, tied by the trend rows of Chapter 4.

Jargon: condition record, stacked state, dashboard trend line, degradation rate, wander, wander floor, drift rate of a sensor, as-of fit, pseudo-true value

The *condition record* is the commissioning test and every later window of an asset read in one fit; carried as one *stacked state*, it holds each window's entries, the entries that do not change from window to window, and the trend entries that tie the windows together. A dashboard's *trend line* is the least-squares straight line through the windows' tiles, its slope reported as a loss per year. The *degradation rate* of a coil is the trend entry of its log-derate's path, a change per year, read in percent of conductance per year. The *wander* of the path is its departure from a straight line from one month to the next, the process noise of the path; the *wander floor* is the spread a rate keeps when the parameter of every window is known exactly, which no measurement on those windows removes. The *drift rate* of a sensor is the change of its offset per year from the window at which its drift begins. An *as-of fit* is the fit of the record cut at a window, using nothing later. The *pseudo-true value* of an entry is the value a drawing reports from the truth's noise-free readings: the like-with-like reference when the truth lies outside the drawing.

Key idea

A degradation rate is the trend entry of one fit over every window. The physics takes the load and the season out of each window's parameter, and the trend rows tie the windows, so the rate is read where the dashboard's slope cannot be read: on the coil's own conductance. Its band has two parts: what the readings leave unknown of the path's two ends, and a floor set by the path's wander from month to month. With 37 windows the floor is about half of each variance. No sensor lowers the floor; only a longer record does.

10.1.1 A rate read from every window

Chapter 4 read a rate from two condition tests run at the same load and supply air (Section 4.6.5), and its band was the two tests' errors with a small share of wander. Chapter 5 read one window against commissioning, and Chapters 6 and 7 read later windows the same way. Here every window of the record carries the coils' parameters at its own load and inputs, and the trend entry is read from the whole record at once. The same trend rows that carry the rate also tie each window to its neighbors, so every window informs the commissioning baseline and today's value as well.

The trend entry γ_u of coil u is the change per year of its log-derate $g_{u,t}$ along the path, in yr^{-1}. Over one year on that path the conductance is multiplied by e^{γ_u}, so the degradation rate in percent of conductance per year is

$$100\,(\mathrm{e}^{\gamma_u} - 1)\ \%/\mathrm{yr}, \tag{10.1}$$

negative for a loss, Chapter 4's reading of a trend entry in conductance terms. The map is increasing, so the median and the band ends of γ_u pass through it unchanged in order, and the rate in percent needs no further computation than the posterior of γ_u; $100\,\gamma_u$ is the same rate to first order. The subscript is the coil: γ_1 and γ_2 are the two coils' rates, and a curvature of the path, when one is freed, is written $\dot{\gamma}_u$. Chapter 9's γ_i indexes the fleet's coils instead, so its γ_1 is fleet coil 1, not the hall's coil 1, which is the fleet's coil 0.

Principle 10.1. *A degradation rate is the posterior of the trend entry of the parameter's path, after the physics has taken the inputs out of every window. A trend of a performance figure is a trend of everything that figure depends on.*

10.1.2 The two-unit hall through its record (step 0)

The plant is the two-unit hall of Chapter 5, Section 5.1.2, cited and not derived again: in each window the five physical rows (5.2) fix the return air, the two discharge temperatures, the cold aisle and rack 1's inlet, and the forward equations (5.3) give them at that window's known inputs $T_s(t)$ and $Q_o(t)$. Racks 2 and 3 are not drawn, and Section 10.4.2 shows that their block still separates on the whole record. The modeling choices of the graph are that chapter's paragraph: steady state within each window, the airflow pinned at the units' fixed fan speed with nothing controlling the discharge air, the chilled-water supply and the other heat as known inputs, and one effective conductance per coil. The choices the record adds are named at the rows, in step 2.

The record. The record is the commissioning test, window 0, and the 36 monthly windows from January of year 1 to December of year 3, five readings each. Table 10.1 shows four of the 37 windows, a year apart. Each row is a window; the columns give its two known inputs, the hall meter's reading, the approach of each unit (its discharge reading less the supply) and the dashboard's tile of each coil, the commissioning approach over the window's. Across the record the supply runs from 12.55 to 15.45°C and the other heat from 5.10 to 10.90 kW with the season, the coldest supply and the least other heat in July and August, the southern winter. The three later rows of the table share their known inputs, because each is a December; the meter and both approaches rise between them.

Table 10.1: Four of the 37 windows of the two-unit hall's condition record, with the dashboard's raw-performance tiles $h_{u,t}^{\mathrm{raw}} = A_{u,0}/A_{u,t}$.

Window [–]	$T_s(t)$ [°C]	$Q_o(t)$ [kW]	y_Q [kW]	$A_{1,t}$ [K]	$A_{2,t}$ [K]	$h_{1,t}^{\mathrm{raw}}$ [–]	$h_{2,t}^{\mathrm{raw}}$ [–]
0 (commissioning)	14.00	8.00	110.92	2.86	3.08	1.0000	1.0000
12 (Dec Y1)	15.06	10.12	124.28	3.57	3.91	0.8013	0.7879
24 (Dec Y2)	15.06	10.12	122.75	3.69	4.51	0.7752	0.6830
36 (Dec Y3), today	15.06	10.12	128.56	5.87	5.18	0.4873	0.5947

The maintenance log so far. The two-unit hall's log holds the commissioning test of window 0; Chapter 5's condition assessment and Chapter 6's entry on the effort at the set point, both at window 18 (Jun Y2); and Chapter 7's assessment at window 25 (Jan Y3), which held coil 1's cleaning order and asked for a side-by-side check of unit 1's discharge sensor (Section 7.5.2). That check has not been made. The record read here therefore carries the drift of that sensor as an entry with Chapter 7's prior, and the check is a candidate of step 7.

What the record adds to the graph. The graph of one window is copied 37 times, each copy with its own free load Q_t and its two log-derates $g_{1,t}$, $g_{2,t}$. Between consecutive windows one trend row per coil ties the copies, and above them the two trend entries γ_1 and γ_2 are linked to every trend row of their coil. The recirculation fraction β_1 is one entry shared by all the copies, because the racks do not age (Chapter 5's choice). The rating prior is attached to window 0's coils only. And unit 1's dashed discharge reading edge carries a drift node from window 20 on: the sensor's offset, calibrated at commissioning, grows from window $t_b = 20$ at the drift rate $\dot{b}_{d,1}$, in K/yr, so that the drift in window t is

$$b_{d,1}(t) = \begin{cases} 0, & t < t_b, \\ \dot{b}_{d,1}\,(t - t_b + 1)\,\Delta t, & t \geq t_b, \end{cases} \qquad t_b = 20, \tag{10.2}$$

with $\Delta t = 1/12\,\mathrm{yr}$ the spacing of the windows: $\dot{b}_{d,1}\,\Delta t$ in window 20, $\dot{b}_{d,1} \times 0.5000\,\mathrm{yr}$ in window 25 and $\dot{b}_{d,1} \times 1.4167\,\mathrm{yr}$ in window 36. It is Chapter 7's bias $b_{d,1}$, carried now through every window from the onset as one entry, the rate.

The graph. The record's graph is drawn in Figure 10.1, Chapter 4's chain of windows (Figure 4.7) grown to the two-unit hall. The middle row is the chain of windows: window 0, a dashed box for windows 1 to 19, window 20, a dashed box for windows 21 to 35, and window 36, each box standing for one copy of Figure 5.1's window with its three free entries. The black squares between neighbors are the trend rows, two per pair of windows, one per coil; the node above holds the two trend entries, linked to every square. Below each window are its five readings. The shared fraction β_1 (bottom left) is linked to the rack-1 inlet reading of every window, and the drift rate $\dot{b}_{d,1}$ (bottom right) to unit 1's discharge reading from window 20 on. The rating prior sits on window 0, the survey prior on β_1 and the drift prior on $\dot{b}_{d,1}$.

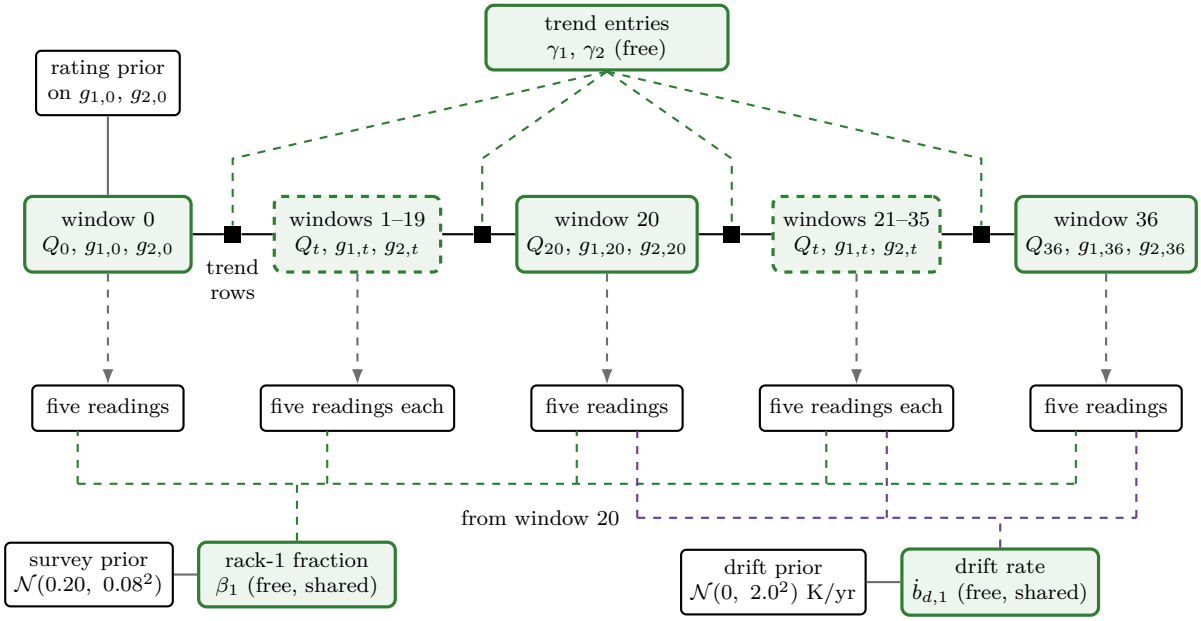


Figure 10.1: The plant graph of the condition record: the two-unit hall's window, Figure 5.1, copied for windows 0 to 36 and tied by trend rows. Free (green): each window's IT load Q_t and log-derates $g_{1,t}$, $g_{2,t}$ (one box per window, dashed boxes for the runs of windows 1–19 and 21–35); the trend entries γ_1, γ_2, linked to every trend row (black squares, two between each pair of neighbors, one per coil); the recirculation fraction β_1, shared by every window and linked to its rack-1 inlet reading; and the drift rate $\dot{b}_{d,1}$ of unit 1's discharge sensor, linked to that reading from window 20 on (purple). Pinned in every window, not drawn: the known inputs $T_s(t)$ and $Q_o(t)$, the ratings $\bar{G}_u = 7.0\,\mathrm{kW/K}$ and the capacity rates $C_u = 10.0\,\mathrm{kW/K}$. Observed (dashed arrows): each window's return air, two discharge temperatures, rack-1 inlet and load meter. Priors (solid gray): the rating prior on window 0's coils, the survey prior on β_1 and the drift prior on $\dot{b}_{d,1}$.

The stacked state. In the limit of exact physical rows each window's five temperatures are functions of its free entries, and the fit runs on the reduced state of Section 4.2.8, stacked over the record,

$$\mathbf{z} = (Q_0,\ g_{1,0},\ g_{2,0},\ \ Q_1,\ g_{1,1},\ g_{2,1},\ \ \ldots,\ \ Q_{36},\ g_{1,36},\ g_{2,36},\ \ \beta_1,\ \ \gamma_1,\ \gamma_2,\ \ \dot{b}_{d,1})^{\mathsf{T}}, \qquad (10.3)$$

115 entries: z_{3t+1} to z_{3t+3} the entries $(Q_t, g_{1,t}, g_{2,t})$ of window t, $z_{112} = \beta_1$, $z_{113} = \gamma_1$, $z_{114} = \gamma_2$ and $z_{115} = \dot{b}_{d,1}$; the loads in kW, the trend entries in yr^{-1}, the drift rate in K/yr and the others

dimensionless. The full state adds each window's five temperatures, 300 entries under 185 physical rows; its point on the set where the rows hold is equation (5.3) applied in each window at that window's inputs.

10.1.3 The asset manager's questions (step 1)

The asset manager asks six questions of the record, and step 1 writes each as a probability over the stacked state (10.3), with the unknown on the left of the bar and the 185 readings, the four prior rows and the 72 trend rows on the right. Table 10.2 lists them. The first column is the question in the asset manager's words, the first of them the drill question; the second the general form of its category from Chapter 2; the third its mathematical form for the hall; the fourth its category; the fifth the computation of Section 2.3.6 that answers it; and the last the subsection where it is answered. Both coils are answered, because the state carries both. The first two questions share a general form and one answer. This is the first chapter of the book whose lead computation is the trend. Whether the rate is itself changing (Chapter 11), whether the path stepped (Chapter 12), how much of a summer drop is the season (Chapter 13) and when the capacity limit is crossed (Chapter 14) are not asked here. Which measurement would narrow the rate asks nothing new of the record; it is step 7, Section 10.6.

Table 10.2: The questions asked of the two-unit hall's condition record at window 36 (Dec Y3). $\mathbf{y}$ is the 185 readings, four prior rows and 72 trend rows; γ_u the trend entry (10.3); $h_{u,t}^{\text{raw}}$ the dashboard's tile; $b_{d,1}(t)$ the drift (10.2).

Asset manager's question	General form	Mathematical form	Category	Computation	Answered in
How many percent of conductance per year is it losing?	sign and estimation of a trend entry, (2.21) (first line), (2.9), (2.11)	$p(\gamma_u \mid \mathbf{y})$, read as $100(\mathrm{e}^{\gamma_u} - 1)$ %/yr: median and 5–95 % band, $u = 1, 2$	estimation	trend	10.3.1
Is it losing conductance at all, beyond three years of scatter?	(2.21) (first line)	$\Pr(\gamma_u < 0 \mid \mathbf{y})$	prediction under a threshold	trend	10.3.1
Which coil is losing conductance faster, and how sure is that?	$\Pr(f(\mathbf{z}) > f^* \mid \mathbf{y})$, (2.12)	$\Pr(\gamma_2 < \gamma_1 \mid \mathbf{y})$	prediction under a threshold	trend	10.3.2
Is the dashboard's 15 % a year the coils, or the load, the season and the sensor?	moments of a derived quantity, (2.10)	$s_u^{\text{raw}}(\mathbf{z})$, the slope of the line through $h_{u,t}^{\text{raw}}(\mathbf{z})$, split into load and season, coils, sensor and reading errors	estimation	trend	10.3.3
How far has each coil fallen since commissioning, read with the whole record?	(2.9), (2.11)	$p(h_{u,36} \mid \mathbf{y})$, $\ln h_{u,36} = g_{u,36} - g_{u,0}$	estimation	posterior	10.3.4
Does the record still need the drift of unit 1's sensor, and how large is it now?	B_{AB}, (2.14); (2.9)	$B(\text{drift} : \text{none})$; $p(b_{d,1}(36) \mid \mathbf{y})$	hypothesis comparison	evidence	10.3.5

10.1.4 The evidence as rows (step 2)

Step 2 attaches the record to the graph as rows on the stacked state (10.3). Each window contributes Chapter 5's five reading rows, equation (5.7), on its own entries $\mathbf{z}_t = (Q_t, g_{1,t}, g_{2,t}, \beta_1)$ at its own inputs, with standard deviations 0.3 K for the four temperatures and 3.0 kW for the load meter, and with one change: unit 1's discharge row adds the drift (10.2),

$$r_{d,1,t}(\mathbf{z}) = \big(T_{d,1}(\mathbf{z}_t) + b_{d,1}(t) - y_{T_{d,1},t}\big)/0.3, \qquad t = 0, \dots, 36, \tag{10.4}$$

which is Chapter 7's bias row (7.2) with the bias a function of the window. The record gives 185 reading rows. Three priors carry what was known before the readings, as in Chapter 5: the rating prior $g_{u,0} \sim \mathcal{N}(0,\ 0.15^2)$ on each commissioning entry and the survey prior $\beta_1 \sim \mathcal{N}(0.20,\ 0.08^2)$. A fourth is the drift prior, a row $\dot{b}_{d,1}/\sigma_{\dot{b}}$ with $\sigma_{\dot{b}} = 2.0$ K/yr. And between consecutive windows each coil has a trend row, equation (4.68) written for coil u,

$$r_{q,u,t}(\mathbf{z}) = \frac{g_{u,t+1} - g_{u,t} - \gamma_u\,\Delta t}{\sigma_q}, \qquad u = 1, 2, \quad t = 0, \dots, 35, \qquad \sigma_q = 0.005, \tag{10.5}$$

72 rows with Chapter 4's spread per monthly window. The loads and the trend entries have no prior row: Chapter 4's decision that the record alone says what the rate is. In all there are $185 + 2 + 1 + 72 + 1 = 261$ rows on 115 entries, 146 degrees of freedom.

Which row acts on which entry. A reading row of window t sees the free entries through the row of $\mathbf{N}_\mathrm{b}$ of the temperature it reads, equation (4.47), so it touches only that window's $(Q_t, g_{1,t}, g_{2,t})$, the rack-1 row also β_1, and unit 1's discharge row also $\dot{b}_{d,1}$ from window 20 on. A trend row touches $g_{u,t}$, $g_{u,t+1}$ and γ_u only. With the columns grouped as in equation (10.3), window by window and then β_1, (γ_1, γ_2) and $\dot{b}_{d,1}$, and $*$ for an entry that can be nonzero, the 261×115 Jacobian has the block pattern

$$\mathbf{J} = \begin{bmatrix} \mathbf{R}_0 & & & & \mathbf{v}_\beta & \mathbf{0} & \mathbf{0} \\ & \ddots & & & \vdots & \vdots & \vdots \\ & & \mathbf{R}_t & & \mathbf{v}_\beta & \mathbf{0} & \mathbf{v}_{b,t} \\ & & & \ddots & \vdots & \vdots & \vdots \\ \mathbf{P}_0 & & & & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ & & & & * & \mathbf{0} & 0 \\ \mathbf{D}_0 & \cdots & \cdots & \mathbf{D}_{36} & \mathbf{0} & \mathbf{E} & \mathbf{0} \\ & & & & 0 & \mathbf{0} & * \end{bmatrix} \begin{array}{l} \text{window 0: five readings} \\ \\ \text{window } t\text{: five readings} \\ \\ \text{rating priors on } g_{1,0}, g_{2,0}\text{,} \\ \text{survey prior on } \beta_1 \\ \text{72 trend rows} \\ \text{drift prior on } \dot{b}_{d,1} \end{array}$$

$$\mathbf{R}_t = \begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \\ * & 0 & 0 \\ * & * & * \end{bmatrix}, \quad \mathbf{v}_\beta = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ * \end{bmatrix}, \quad \mathbf{v}_{b,t} = \begin{bmatrix} 0 \\ * \\ 0 \\ 0 \\ 0 \end{bmatrix} \ (t \geq 20), \tag{10.6}$$

$$\mathbf{E} = -\frac{\Delta t}{\sigma_q} \begin{bmatrix} \mathbf{1}_{36} & \mathbf{0} \\ \mathbf{0} & \mathbf{1}_{36} \end{bmatrix},$$

$$\mathbf{g}_{\gamma_1}^\mathsf{T} = \mathbf{e}_{113}^\mathsf{T}, \qquad \mathbf{g}_{\gamma_2}^\mathsf{T} = \mathbf{e}_{114}^\mathsf{T}, \qquad \mathbf{g}_{\gamma_1 - \gamma_2}^\mathsf{T} = \mathbf{e}_{113}^\mathsf{T} - \mathbf{e}_{114}^\mathsf{T}.$$

Here $\mathbf{R}_t$ is window t's 5×3 block in the columns of $(Q_t, g_{1,t}, g_{2,t})$, rows in the order of equation (5.7); $\mathbf{v}_\beta$ its column in β_1, the rack-1 row; $\mathbf{v}_{b,t}$ its column in $\dot{b}_{d,1}$, nonzero only from window 20; $\mathbf{P}_0$ the rating priors' 2×3 block on window 0's coils; and $\mathbf{D}_t$ the trend rows' block in window t's columns, $\pm 1/\sigma_q$ in the columns of $g_{u,t}$ for the rows that tie window t to its neighbors,

so that the 72 trend rows form a banded block along the windows. **E** is the trend rows' block in (γ_1, γ_2), $-\Delta t/\sigma_q$ in coil 1's 36 rows of the first column and coil 2's 36 of the second. Of the $261 \times 115 = 30\,015$ places, 755 can be nonzero, 2.52 %. The last line is not a row: it gives the gradients of the questions' trend quantities, which are entries of the state.

Two facts of the graph. Two facts are read here, before any number. The first: no reading row sees γ_u. The trend entry's column is **E**'s, 36 entries all in coil u's trend rows, as in Chapter 4 with 36 rows in place of one, so the rate is read from the path the readings fix, through the trend rows alone. At the MAP state the objective's slope in γ_u vanishes, and only the trend rows contribute to it:

$$\frac{\partial \mathcal{L}}{\partial \gamma_u} = \sum_{t=0}^{35} r_{q,u,t} \frac{\partial r_{q,u,t}}{\partial \gamma_u} = -\frac{\Delta t}{\sigma_q^2} \sum_{t=0}^{35} (\hat{g}_{u,t+1} - \hat{g}_{u,t} - \hat{\gamma}_u \, \Delta t) = 0 \quad \Longrightarrow \quad \hat{\gamma}_u = \frac{\hat{g}_{u,36} - \hat{g}_{u,0}}{36 \, \Delta t}, \tag{10.7}$$

because the sum of the increments telescopes. The trend entry at the MAP state is the fitted path's change from commissioning to today over the three years, the telescoped sum of equation (4.69) read at the fitted path; the windows in between act on it through where they put the path's two ends. The second: the drift's column touches unit 1's discharge rows of windows 20 to 36 and the drift prior, and nothing else. Only the other readings of those windows tell the drift from coil 1, Chapter 7's principle, now over 17 windows instead of one.

Modeling choices at the rows. Six choices make these rows, each named once. One path per coil, a constant rate with wander, equation (10.5) at $\sigma_q = 0.005$, Chapter 4's value: this buys a rate read from every window and a band that allows the path its month-to-month departures, and it fails if a coil's rate changes or its path steps; the record's departures are reported in step 3 as the record shows them, and whether they are either is the question of Chapters 11 and 12. The spread σ_q is not re-chosen from this record; Section 10.4.4 sets the evidence over it beside the choice. The drift from window 20 as a rate, equation (10.2): one entry for the whole record, linear in time, and it fails if the drift began elsewhere or is not linear; Section 10.4.4 shows that the record cannot place the onset and that the rates do not depend on it. Chapter 7's prior on the drift's size, $\mathcal{N}(0,\ 1.0^2)$ K at window 25, carried to the rate: $\sigma_{\dot{b}} = 1.0/((25 - t_b + 1)\,\Delta t) = 2.0$ K/yr, which Section 10.4.3 varies. The sensors calibrated at commissioning, the rating prior on window 0 only and β_1 shared, Chapter 5's choices carried to 37 windows. The readings independent across windows: one record's errors are not correlated from month to month, which is what lets each window add its own information, and it fails for errors that persist across months, Chapter 24's question. And commissioning on the same path: the trend row from window 0 to window 1 ties the commissioning test to the first month, which buys every window informing the baseline, and it fails if the coil changed between handover and January of year 1, by a bedding-in or a first clean.

10.2 The posterior of the state (step 3)

All six questions of Table 10.2 are read from one posterior, that of the stacked state (10.3) under the 261 rows of step 2, with the known inputs of every window and the pinned quantities of Section 10.1.2. Part III reads this one fit, the base fit of the condition record, and the chapters that follow add to it only what their questions need.

10.2.1 The rows on the stacked state

The rows are those of Chapter 5's joint state, Section 5.2.1, written for every window, with the trend rows of Section 4.6.5 between them. Window t's five reading rows are $\mathbf{r}_t(\mathbf{z}_t)$ of

equation (5.7) with unit 1's discharge row replaced by equation (10.4), and the residual vector stacks the 37 windows, the priors and the trend rows in the order of the pattern (10.6),

$$\mathbf{r}(\mathbf{z}) = \begin{bmatrix} \mathbf{r}_0(\mathbf{z}_0) \\ \vdots \\ \mathbf{r}_{36}(\mathbf{z}_{36}) \\ g_{1,0}/0.15 \\ g_{2,0}/0.15 \\ (\beta_1 - 0.20)/0.08 \\ r_{q,1,0}(\mathbf{z}) \\ \vdots \\ r_{q,2,35}(\mathbf{z}) \\ \dot{b}_{d,1}/\sigma_{\dot{b}} \end{bmatrix}, \qquad \mathcal{L}(\mathbf{z}) = \tfrac{1}{2}\,\mathbf{r}(\mathbf{z})^{\mathsf{T}}\mathbf{r}(\mathbf{z}), \tag{10.8}$$

261 dimensionless rows on 115 entries. The reduced Jacobian $\mathbf{J} = \mathbf{J}_{\text{rows}}\mathbf{N}$ of equation (4.47) is formed window by window, as in Chapter 5: each reading row is its temperature's row of that window's $\mathbf{N}_{\text{b}}$ over the sensor's 0.3 K, the meter row is 1/3.0 in the column of Q_t, and unit 1's discharge row adds $(t - t_b + 1)\,\Delta t/0.3$ in the column of $\dot{b}_{d,1}$ from window 20 on. Each trend row has the constant line of Section 4.6.5, $-1/\sigma_q$ and $+1/\sigma_q$ in the columns of $g_{u,t}$ and $g_{u,t+1}$ and $-\Delta t/\sigma_q$ in the column of γ_u; each prior row is the reciprocal of its standard deviation in its own column. The reading rows are nonlinear in the state through $\mathrm{e}^{g_{u,t}}$ and $\beta_1(T_r - T_c)$, so their lines are evaluated at each iterate; the trend rows are linear and add nothing to the curvature term $\mathbf{C}$ of equation (4.44).

The Jacobian at the MAP state. The 261 lines are 37 copies of Chapter 5's window lines at each window's own state and inputs, plus the constant lines. Window 25's lines show both kinds. With the columns $(Q_{25}, g_{1,25}, g_{2,25}, \beta_1, \dot{b}_{d,1}, g_{1,24}, g_{1,26}, \gamma_1, \gamma_2)$ and the rows window 25's five readings, the two trend rows of coil 1 that touch $g_{1,25}$, and the drift prior,

$$\begin{bmatrix} 0.2682 & -19.8850 & -18.3188 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.0947 & -31.7336 & -6.4702 & 0 & 1.6667 & 0 & 0 & 0 & 0 \\ 0.1084 & -8.0364 & -30.1674 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.3333 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.1370 & -19.8850 & -18.3188 & 23.7371 & 0 & 0 & 0 & 0 & 0 \\ 0 & 200 & 0 & 0 & 0 & -200 & 0 & -16.6667 & 0 \\ 0 & -200 & 0 & 0 & 0 & 0 & 200 & -16.6667 & 0 \\ 0 & 0 & 0 & 0 & 0.5 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{matrix} y_{T_r} \\ y_{T_{d,1}} \\ y_{T_{d,2}} \\ y_Q \\ y_{T_1} \\ \text{trend 24–25} \\ \text{trend 25–26} \\ \text{drift prior} \end{matrix}, \tag{10.9}$$

the load column in kW^{-1}, the trend entries' columns in yr, the drift's in yr/K and the others dimensionless. The window's lines in equation (10.9) are those of Chapter 5's window 18, equation (5.12), with larger slopes, because January's heat to remove is larger: each discharge row has its largest entry in its own coil's column, the return-air row sees both coils nearly alike, and the rack-1 row repeats it in the log-derate columns. Unit 1's discharge row carries $1.6667 = (25 - 19)\,\Delta t/0.3$ in the drift's column, the only reading entry there. The trend lines are $\pm 1/\sigma_q = \pm 200$ in the path's columns and $-\Delta t/\sigma_q = -16.6667$ in the column of γ_1: a trend row weighs a month's change of the path at 200 per unit, which is what makes the path's increments, not its level, the trend rows' business.

10.2.2 The normal equations and the covariance

The MAP state $\hat{\mathbf{z}}$ minimizes the objective (10.8). Each Gauss–Newton iteration solves the normal equations (5.9) on the stacked rows, with every window's temperatures solved again by equation (5.3) at the iterate. The iteration starts from the state the record suggests, the meter's

loads, the coils at their ratings, the survey's fraction, no trend and no drift: $Q_t = y_{Q,t}$, $g_{u,t} = 0$, $\beta_1 = 0.20$, $\gamma_u = 0$ and $\dot{b}_{d,1} = 0$. With every row normalized the information matrix is $\mathbf{\Lambda} = \mathbf{J}^\mathsf{T}\mathbf{J}$; with unnormalized rows and their weights it is $\mathbf{J}^\mathsf{T}\mathbf{W}\mathbf{J}$. At the minimum the Laplace covariance is $\mathbf{\Sigma} = (\mathbf{J}^\mathsf{T}\mathbf{J})^{-1}$, the Gauss–Newton form of equation (4.38), as in Chapter 5's equation (5.10), and by equation (4.48) it is the free-entry block of the 300-entry full state's covariance in the limit of exact physical rows; Section 10.4.2 sets the full Hessian and the full state beside it.

The block structure. The pattern (10.6) fixes the shape of $\mathbf{\Lambda}$ before any number. Two columns meet only where a row touches both. The reading rows give each window a 3×3 block on the diagonal and no coupling between windows. The trend rows couple each window's coil entries to the next window's, and every window's coil entries to the trend entries: one trend row, with the line $\mathbf{j}_q = (-1,\ 1,\ -\Delta t)/\sigma_q$ in the columns of $(g_{u,t}, g_{u,t+1}, \gamma_u)$, adds

$$\mathbf{j}_q^\mathsf{T}\mathbf{j}_q = \frac{1}{\sigma_q^2}\begin{bmatrix} 1 & -1 & \Delta t \\ -1 & 1 & -\Delta t \\ \Delta t & -\Delta t & \Delta t^2 \end{bmatrix} \tag{10.10}$$

to those three entries. So $\mathbf{\Lambda}$ is block-tridiagonal along the windows, with a border: the columns of β_1 (every window's rack-1 row), of γ_1 and γ_2 (every trend row) and of $\dot{b}_{d,1}$ (windows 20 to 36). The trend entries' own information is the sum of the corner entries of equation (10.10) over 36 rows, $\Lambda_{\gamma_u\gamma_u} = 36\,\Delta t^2/\sigma_q^2$, the trend rows' alone. Chapter 5's two windows, with no trend rows, were independent given β_1; here each window is tied to its neighbors, and through γ_u to all of them, whatever β_1 and the drift. That is how every window comes to inform the commissioning baseline.

10.2.3 The fit of the record

The iterations. Table 10.3 runs the normal equations from the starting state. Each row is an iteration, with the objective and the norm of its gradient $\mathbf{J}^\mathsf{T}\mathbf{r}$. The objective falls from 1250.311469 to 82.074870 in the first step, to 77.488380 in the second and to 77.487758 in the third, where it stays; the gradient falls to 1.1×10^{-7} by iteration 7, the round-off of the Jacobian.

Table 10.3: Gauss–Newton iterations of the stacked state of the condition record, from the meter's loads, the rated coils, the survey's fraction, no trend and no drift.

Iteration l [–]	Objective $\mathcal{L}(\mathbf{z}^{(l)})$ [–]	$\lVert\mathbf{J}^\mathsf{T}\mathbf{r}\rVert$ [–]
0	1250.311469	3.649×10^{3}
1	82.074870	2.248×10^{2}
2	77.488380	1.825
3	77.487758	4.006×10^{-2}
4	77.487758	9.853×10^{-4}
5	77.487758	2.318×10^{-5}
6	77.487758	1.144×10^{-6}
7	77.487758	1.126×10^{-7}

The residuals. At the MAP state twice the objective is $2\mathcal{L}(\hat{\mathbf{z}}) = 154.9755$ on $261 - 115 = 146$ degrees of freedom, and $\Pr(\chi^2_{146} > 155.0) = 0.290$: the record sits within its errors. The readings carry 150.10 of it on 185 rows and the trend rows 4.70 on 72. The largest normalized residual of a reading, Chapter 7's $r_k/\sqrt{1 - K_{kk}}$ of equation (7.20), is window 20's discharge of unit 2, -3.94, a reading above its prediction; the next are window 8's return air, -2.83, and window 1's meter, $+2.63$. A largest value of 3.94 in absolute value over 185 independent normalized

rows has probability 0.015. The reading is kept in the fit, and Section 10.4.3 sets the fit without it beside the reported one.

The MAP state. Table 10.4 gives the eight entries the answers use, with their posterior spreads $\sqrt{\Sigma_{ii}}$ and prior spreads; the last column names the rows that inform each. The 37 loads have spreads of 2.17 to 2.37 kW, and they and the 72 inner log-derates are drawn in the path figure of step 4. The commissioning entries are known to 0.01543 and 0.01558, against 0.03962 and 0.04074 in Chapter 5's fit of windows 0 and 18: the trend rows bring every window's readings to bear on them. Today's entries are known less well, because the last window has neighbors on one side only, and coil 1's the least, because the drift shares unit 1's discharge reading with coil 1 in the last 17 windows.

Table 10.4: The MAP state of the condition record: the entries the answers use, with their posterior and prior spreads.

Entry [unit]	$\hat{z}_i$ [entry's]	$\sqrt{\Sigma_{ii}}$ [entry's]	Prior sd [entry's]	Informed by
$g_{1,0}$ [–]	−0.00594	0.01543	0.15	every window, through the trend rows; the rating prior
$g_{2,0}$ [–]	−0.02448	0.01558	0.15	every window, through the trend rows; the rating prior
$g_{1,36}$ [–]	−0.13314	0.02386	none	every window; the drift shares its discharge reading
$g_{2,36}$ [–]	−0.21208	0.01690	none	every window, through the trend rows
β_1 [–]	0.21272	0.00893	0.08	37 rack-1 inlet readings; the survey
γ_1 [yr^{-1}]	−0.04240	0.01463	none	coil 1's 36 trend rows
γ_2 [yr^{-1}]	−0.06253	0.01323	none	coil 2's 36 trend rows
$\dot{b}_{d,1}$ [K/yr]	0.6837	0.1681	2.0	windows 20–36; the drift prior

The covariance. The block of $\mathbf{\Sigma}$ on these eight entries, rows and columns in the order of Table 10.4, is

$$\mathbf{\Sigma}_8 = 10^{-4} \times \begin{bmatrix} 2.3812 & -0.86017 & -1.0913 & 0.58355 & 0.027359 & -1.1575 & 0.48124 & -5.6793 \\ -0.86017 & 2.4273 & 0.86837 & -0.73606 & 0.21693 & 0.57618 & -1.0545 & 4.9732 \\ -1.0913 & 0.86837 & 5.6927 & -2.3487 & 0.68981 & 2.2613 & -1.0724 & 32.510 \\ 0.58355 & -0.73606 & -2.3487 & 2.8577 & -0.13410 & -0.97743 & 1.1979 & -13.910 \\ 0.027359 & 0.21693 & 0.68981 & -0.13410 & 0.79678 & 0.22082 & -0.11701 & 5.0246 \\ -1.1575 & 0.57618 & 2.2613 & -0.97743 & 0.22082 & 2.1396 & -0.51787 & 12.730 \\ 0.48124 & -1.0545 & -1.0724 & 1.1979 & -0.11701 & -0.51787 & 1.7508 & -6.2945 \\ -5.6793 & 4.9732 & 32.510 & -13.910 & 5.0246 & 12.730 & -6.2945 & 282.65 \end{bmatrix},$$

with the trend entries' rows and columns in yr^{-1} and the drift's in K/yr. The correlations read by sign, and the rows fix each sign. The two rates are correlated by −0.2676: the return air sees the hall's total conductance, so a faster loss on one coil is paid for by a slower one on the other, Chapter 5's within-window anticorrelation carried to the slopes. The drift rate is correlated with γ_1 by +0.5176 and with $g_{1,36}$ by +0.8104: a larger drift leaves less of the rise of unit 1's discharge reading to coil 1, so coil 1 stands higher today and loses less, Chapter 7's ridge read along 17 windows; with γ_2 it is correlated by −0.2830, through the same total. Each commissioning entry is correlated with its coil's rate by −0.5128 (coil 1) and −0.5115 (coil 2): a higher commissioned value, with the later windows held, is a faster loss, Chapter 4's −0.71 of two windows, now shared with 36 more. Today's entries are correlated with the rates by +0.6479 and +0.5356. And β_1 is correlated with γ_1 by +0.1691, through rack 1's inlet and the cold aisle unit 1 feeds.

10.2.4 The path as the record shows it

The fit returns a path for each coil, one log-derate per window, and its departures from the line are part of the posterior. The MAP path changes by -0.0387, -0.0317 and -0.0568 in years 1, 2 and 3 for coil 1, and by -0.0607, -0.0724 and -0.0545 for coil 2. Coil 2's steepest monthly increments are those from window 14 to 15 (-0.00846), 15 to 16 (-0.00829) and 13 to 14 (-0.00774), against a mean increment of -0.00521, and its trend rows there carry the record's largest trend residuals, -0.649, -0.616 and -0.507. Coil 1's steepest increments lie in year 3, from window 31 to 32 (-0.00604), 33 to 34 (-0.00596) and 30 to 31 (-0.00560), against a mean of -0.00353, with trend residuals of -0.501 and -0.486 at the first two. These are the record's departures as it shows them, within what σ_q allows a month, and this chapter does not test them: whether coil 2's path changed in a single event is the question of Chapter 12, and whether coil 1's rate is changing that of Chapter 11. The base fit is the constant-rate drawing with wander and no step.

10.2.5 The full posterior by reweighted draws

The Laplace normal $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ is the reported posterior. Beside it stands the full posterior of the same rows, computed by the reweighted draws of Chapter 5, Section 5.2.4: draws from the Laplace normal, each weighted by equation (5.13) with the objective (10.8) evaluated through the exact forward equations of every window, and Monte Carlo standard errors by batch means over 20 batches. For this record, $N_{\mathrm{d}} = 100\,000$ draws from a seeded generator give an effective sample size of 98 849, 98.8 % of N_{d}, in 115 dimensions, and a largest weight of 2.15×10^{-5}: the Laplace normal is close to the posterior, and no draw dominates. Section 10.4.1 states the standard errors of each answer.

10.3 The answers (step 4)

Every answer is read from the pair $(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ of Section 10.2.3, and that value is the one reported. A derived quantity f is carried by its gradient $\mathbf{g} = \partial f / \partial \mathbf{z}$ at $\hat{\mathbf{z}}$, with variance $\mathbf{g}^{\mathsf{T}} \boldsymbol{\Sigma} \mathbf{g}$. The full posterior of Section 10.2.5 stands beside each value, a median beside a median. Rates are reported in percent of conductance per year by equation (10.1), negative for a loss.

10.3.1 Each coil's degradation rate

Asset manager's questions. How many percent of conductance per year is it losing? Is it losing conductance at all, beyond three years of scatter? Both are asked of each coil in turn.

General form. The sign of a trend is the first line of equation (2.21), here for a falling quantity, and the size of the trend is the estimation form (2.9) with its credible interval (2.11) at the levels 0.05 and 0.95:

$$\Pr(\gamma < 0 \mid \mathbf{y}) = \int \mathbf{1}[\gamma < 0]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}, \qquad \Pr(\gamma \le \gamma^* \mid \mathbf{y}) = \int \mathbf{1}[\gamma \le \gamma^*]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z},$$

with γ the trend entry of a state that carries the parameter's path.

Specialized form. The questions are the first line of equation (2.21) and equations (2.9) and (2.11) with $\mathbf{z}$ the stacked state (10.3), $\gamma = \gamma_u$, the entry z_{113} for coil 1 and z_{114} for coil 2, and $p(\mathbf{z} \mid \mathbf{y})$ the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ of Section 10.2.2. The gradient of γ_u is the unit vector of the pattern's second line (10.6), so the quantity's variance is the diagonal entry $\Sigma_{\gamma_u \gamma_u}$, and

by equation (2.22)

$$\Pr(\gamma_u < 0 \mid \mathbf{y}) \approx \Phi\left(\frac{-\hat{\gamma}_u}{\sqrt{\Sigma_{\gamma_u\gamma_u}}}\right), \qquad \operatorname{med}(\gamma_u \mid \mathbf{y}) = \hat{\gamma}_u, \qquad [\hat{\gamma}_u - 1.6449\sqrt{\Sigma_{\gamma_u\gamma_u}},\ \hat{\gamma}_u + 1.6449\sqrt{\Sigma_{\gamma_u\gamma_u}}], \tag{10.11}$$

the median and the 5–95 % band carried to the rate in percent by the increasing map (10.1). One choice enters here: the band is split by what a measurement could remove. Given the path, every window's log-derates known exactly, the trend rows are the only rows left on γ_u, with no prior on it; by equation (10.10) their information on it is $36\,\Delta t^2/\sigma_q^2$, and by equation (10.7) their mean is the path's end-to-end change, so

$$\gamma_u \mid g_{u,0}, \ldots, g_{u,36} \sim \mathcal{N}\left(\frac{g_{u,36} - g_{u,0}}{36\,\Delta t},\ \frac{\sigma_q^2}{36\,\Delta t^2}\right), \qquad \text{sd} = \frac{\sigma_q}{\sqrt{36}\,\Delta t}, \tag{10.12}$$

the wander floor: the spread of the rate of a path known at every window, which no reading of those windows can lower. The law of total variance over the path then splits the posterior variance exactly, with $g_{u,36} - g_{u,0} = \ln h_{u,36}$ and $36\,\Delta t = 3\,\text{yr}$:

$$\operatorname{Var}(\gamma_u \mid \mathbf{y}) = \underbrace{\frac{\operatorname{Var}(\ln h_{u,36} \mid \mathbf{y})}{(3\,\text{yr})^2}}_{\text{the path's ends: readings can narrow it}} + \underbrace{\frac{\sigma_q^2}{36\,\Delta t^2}}_{\text{the floor: no sensor removes it}}. \tag{10.13}$$

This is what lets the band be read by what would narrow it. A second split, by the straight line, is often quoted and is a different thing: the variance in the limit $\sigma_q \to 0$, equation (4.74), is what the readings' errors give a path with no wander, and the rest of the variance is what the wander adds to it, part of which a better reading of the path would still remove; the straight line also moves the MAP value.

Building blocks. The MAP state $\hat{\mathbf{z}}$ and the covariance $\boldsymbol{\Sigma}$ of Section 10.2.3, of which $\hat{\gamma}_u$ and $\Sigma_{\gamma_u\gamma_u}$ are entries, and its block on $(g_{u,0}, g_{u,36})$; the normal tail and the band (10.11); the map (10.1); the floor (10.12) and the budget (10.13); the fit of the same rows in the straight-line limit.

Each coil's rate and its loss. For coil 1, $\hat{\gamma}_1 = -0.04240\,\text{yr}^{-1}$ with $\sqrt{\Sigma_{\gamma_1\gamma_1}} = 0.01463$, a 5–95 % band of $[-0.06646,\ -0.01834]\,\text{yr}^{-1}$ and, by equation (10.1), a rate of -4.151 %/yr with band $[-6.430,\ -1.817]$ %/yr. By equation (10.11),

$$\Pr(\gamma_1 < 0 \mid \mathbf{y}) = \Phi(0.04240/0.01463) = \Phi(2.8986) = 0.998126.$$

For coil 2, $\hat{\gamma}_2 = -0.06253\,\text{yr}^{-1}$ with $\sqrt{\Sigma_{\gamma_2\gamma_2}} = 0.01323$, band $[-0.08430,\ -0.04077]\,\text{yr}^{-1}$, a rate of -6.062 %/yr with band $[-8.084,\ -3.995]$ %/yr, and $\Pr(\gamma_2 < 0 \mid \mathbf{y}) = \Phi(4.7261)$, with $1 - \Pr = 1.144 \times 10^{-6}$. Both coils are losing conductance with near certainty. The full posterior gives a median of -4.121 %/yr and a band of $[-6.388,\ -1.803]$ for coil 1, with $\Pr(\text{loss}) = 0.998186$, and a median of -6.087 %/yr and a band of $[-8.109,\ -4.020]$ for coil 2. Read as $100\,\gamma_u$, the two rates are -4.240 and -6.253 %/yr; the chapter reports the form of equation (10.1).

The two posteriors of the rate are plotted in Figure 10.2. The horizontal axis is the rate in percent of conductance per year and the vertical axis its posterior density; the shaded areas are the 5–95 % bands, the solid gray line marks no change, the short dashed ticks the rates the readings give when taken as exact through the physics (Section 10.3.3), the arrow at the left edge the dashboard's two slopes, which lie off the axis, and the dotted lines the drawing's pseudo-true values, check markers for Section 10.4.7. Both densities lie wholly to the left of zero, coil 2's to the left of coil 1's with a large overlap, and both lie far to the right of the dashboard's slopes.

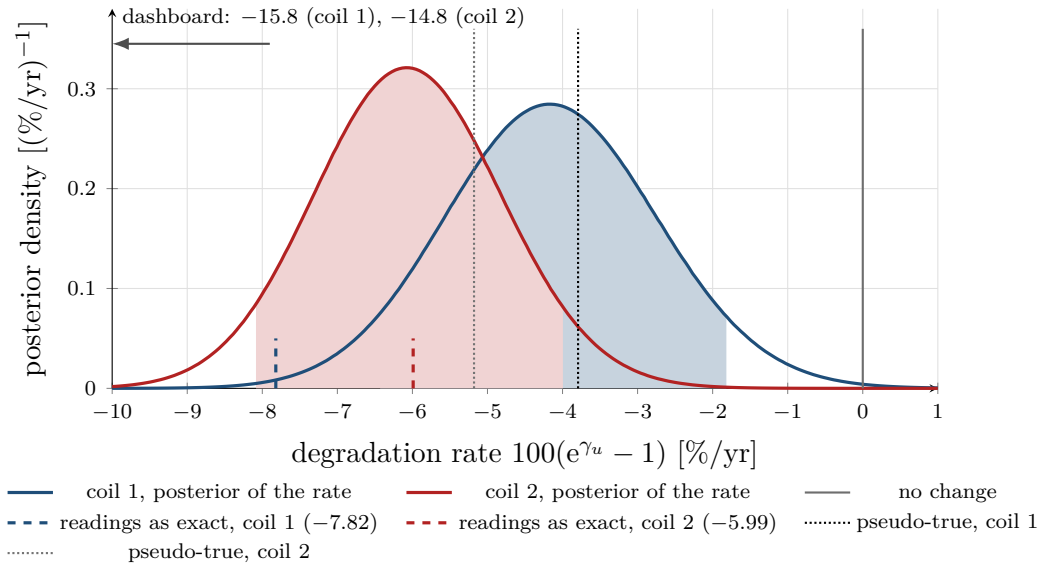


Figure 10.2: The posterior densities of each coil's degradation rate under the Laplace normal of the stacked state, carried through equation (10.1): $\ln(1 + \text{rate}/100) = \gamma_u \sim \mathcal{N}(\hat{\gamma}_u, \Sigma_{\gamma_u\gamma_u})$ with $(-0.04240,\ 0.01463^2)$ for coil 1 and $(-0.06253,\ 0.01323^2)$ for coil 2. Shaded: the 5–95 % bands. Solid gray: no change. Short dashed ticks: the readings taken as exact through the physics. Arrow: the dashboard's slopes, off the axis to the left. Dotted: the drawing's pseudo-true values, check markers (Section 10.4.7).

The band by its sources. The floor (10.12) is $0.005/(\sqrt{36}/12) = 0.01000\,\text{yr}^{-1}$, a variance of $1.0000 \times 10^{-4}\,\text{yr}^{-2}$. By equation (10.13), coil 1's variance of 2.1396×10^{-4} is 1.1396×10^{-4} from the path's ends, 53.3 %, and the floor, 46.7 %; coil 2's 1.7508×10^{-4} is 7.5080×10^{-5} from the ends, 42.9 %, and the floor, 57.1 %. The two parts add to the printed digits. In the straight-line limit the same rows give $\hat{\gamma}_1 = -0.03861$ with standard deviation 0.00970 and $\hat{\gamma}_2 = -0.06637$ with 0.00777, so that by the straight-line budget coil 1's variance is 9.4068×10^{-5} of readings, 44.0 %, and 1.1989×10^{-4} added by the wander, 56.0 %, and coil 2's 6.0303×10^{-5}, 34.4 %, and 1.1478×10^{-4}, 65.6 %; the wander's part there is larger than the floor, because it also holds what the wander does to the reading of the path's ends. The surprise of the step is the width. Three years of monthly readings know each coil's rate to about ±2.3 and ±2.0 %/yr, half of bands 4.613 and 4.089 points wide, and nearly half of coil 1's variance and more than half of coil 2's is the floor, which no sensor removes.

The path. The two paths behind the rates are drawn in Figure 10.3, one panel per coil. The horizontal axis is the window and the vertical axis the log-derate $g_{u,t}$; the shaded band is the MAP path with one standard deviation either side, the solid line through it the MAP path, the dashed line the fitted straight line $\hat{g}_{u,0} + \hat{\gamma}_u\, t\, \Delta t$, the open circles each window's readings taken as exact through the physics, $\ln(G_{u,\text{test}}/\bar{G}_u)$, the vertical line in coil 1's panel window 20, from which the drift entry acts, and the crosses the stated truth at windows 0 and 36, check markers for Section 10.4.7. The circles scatter about the path by much more than its band, because each is one window's readings alone; the path follows their run, not their scatter. From window 20 on, coil 1's circles fall increasingly below its path, most in year 3: they carry the drift, which a sensor reading high turns into a lower conductance, and the fit gives that part to the drift entry. Coil 2's lowest circle is window 20's, the reading with the largest normalized residual of Section 10.2.3. The fitted line and the MAP path end at the same two values, by equation (10.7).

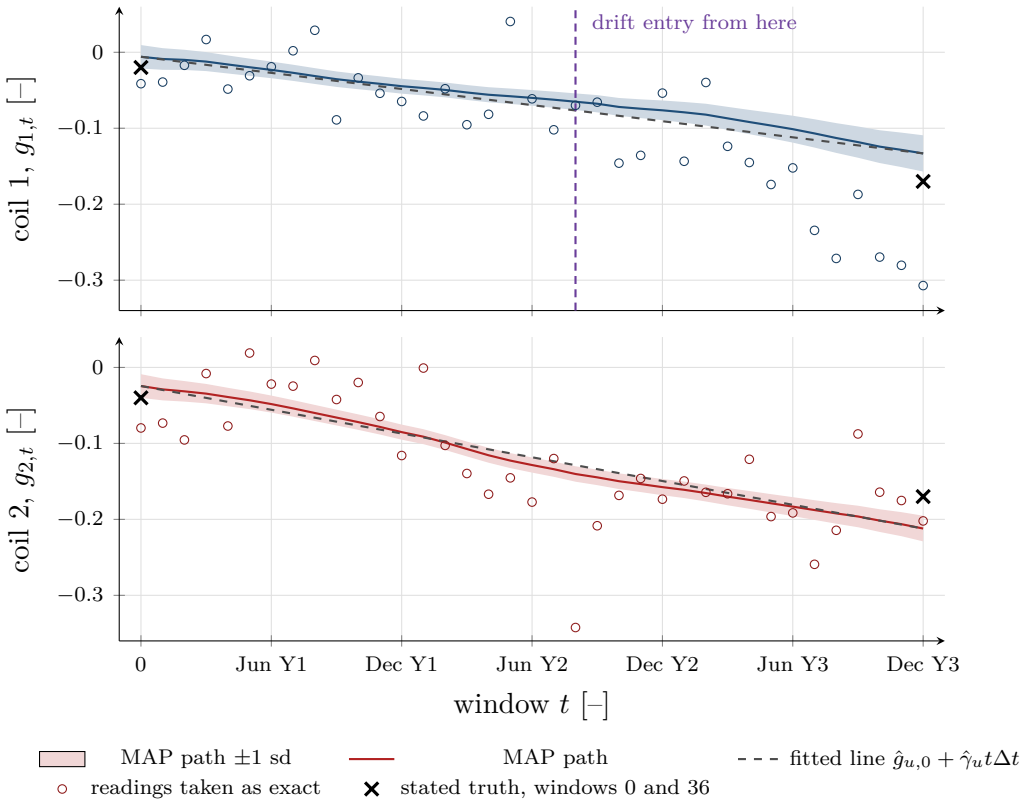


Figure 10.3: Each coil's path of log-derates across the record, from the stacked state: coil 1 (top, blue) and coil 2 (bottom, red); the legend applies to both panels. Shaded: the MAP path with one posterior standard deviation either side. Solid: the MAP path. Dashed: the fitted straight line through the trend entry, with the MAP commissioning entry, $-0.00594 - 0.04240\,t/12$ and $-0.02448 - 0.06253\,t/12$. Open circles: each window's readings taken as exact through the physics, $\ln(G_{u,\text{test}}/\bar{G}_u)$, with the logged sensors as they read. Purple line: window 20, the drift's first window. Crosses: the stated truth at windows 0 and 36, check markers (Section 10.4.7).

10.3.2 The coil losing conductance faster

Asset manager's question. Which coil is losing conductance faster, and how sure is that? The dashboard ranks coil 1 first, by its steeper line.

General form. A prediction under a threshold is the posterior probability of an event on a derived quantity, equation (2.12),

$$\Pr(f(\mathbf{z}) > f^* \mid \mathbf{y}) = \int \mathbf{1}\big[f(\mathbf{z}) > f^*\big]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}.$$

Specialized form. The question is equation (2.12) with $f(\mathbf{z}) = \gamma_1 - \gamma_2$ and $f^* = 0$: the event $\gamma_2 < \gamma_1$, coil 2's trend entry the more negative, so that coil 2 is losing conductance faster. One choice enters here: "faster" is read on the trend entries, each coil's rate over the whole record; a coil whose rate changed during the record is compared by that rate, and the record's departures of Section 10.2.4 stay inside it. The quantity is linear, with the gradient $\mathbf{e}_{113} - \mathbf{e}_{114}$ of the pattern's second line (10.6), and under the Laplace normal

$$\mathrm{Var}(\gamma_1 - \gamma_2 \mid \mathbf{y}) = \Sigma_{\gamma_1\gamma_1} + \Sigma_{\gamma_2\gamma_2} - 2\,\Sigma_{\gamma_1\gamma_2}, \qquad \Pr(\gamma_2 < \gamma_1 \mid \mathbf{y}) = \Phi\left(\frac{\hat\gamma_1 - \hat\gamma_2}{\sqrt{\mathrm{Var}(\gamma_1 - \gamma_2 \mid \mathbf{y})}}\right). \tag{10.14}$$

Building blocks. The trend entries and their block of $\boldsymbol{\Sigma}$ from Section 10.2.3; the variance (10.14); Φ.

The order of the two coils. The difference is $\hat\gamma_1 - \hat\gamma_2 = 0.02014\,\mathrm{yr}^{-1}$, and by equation (10.14)

$$\mathrm{Var}(\gamma_1 - \gamma_2 \mid \mathbf{y}) = 2.1396\times10^{-4} + 1.7508\times10^{-4} - 2\,(-5.1787\times10^{-5}) = 4.9261\times10^{-4} = 0.02219^2,$$

so that

$$\Pr(\gamma_2 < \gamma_1 \mid \mathbf{y}) = \Phi(0.02014/0.02219) = \Phi(0.9072) = 0.8179 :$$

coil 2 is losing conductance faster, with probability 0.82. The full posterior gives 0.8263. The cross term sets this answer. The two rates are correlated negatively, because the return air sees their sum, and the gradient takes them with opposite signs, so the negative covariance adds to the variance of the difference: 0.02219 against 0.01463 and 0.01323 for the rates themselves. The record knows each coil's rate better than it knows which is faster. The dashboard's order is the reverse, coil 1's line the steeper of the two. On the straight line the same question gives 0.9605. Whether coil 2's lead is a faster rate or a single event in its path is the question of Chapter 12.

10.3.3 The dashboard's slope as a quantity of the state

Asset manager's question. Is the dashboard's 15 % a year the coils, or the load, the season and the sensor? The question asks what the dashboard's line measures, and how much of its slope would be there if the coils had not aged and the sensor had not drifted.

General form. An estimation question about a derived quantity is summarized by its moments, equation (2.10),

$$\mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}] = \int f(\mathbf{z})\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z},$$

and by its credible interval, equation (2.11), at the levels 0.05 and 0.95.

Specialized form. The question is equations (2.10) and (2.11) with $f(\mathbf{z}) = s_u^{\text{raw}}(\mathbf{z})$, the slope of the dashboard's line computed from the state, and $p(\mathbf{z} \mid \mathbf{y})$ the Laplace normal. Window t's tile is Chapter 5's $h_{u,t}^{\text{raw}} = A_{u,0}/A_{u,t}$, with the approach $A_{u,t}$ of equation (5.18) at window t's entries and inputs and, for unit 1, the drift added to its discharge reading, $A_{1,t} = T_{d,1}(\mathbf{z}_t) + b_{d,1}(t) - T_s(t)$. The dashboard's line is the least-squares line through the 37 tiles against the time $t\,\Delta t$, so its slope is a fixed linear combination of them,

$$s_u^{\text{raw}}(\mathbf{z}) = \frac{\sum_{t=0}^{36} (t-18)\,\Delta t\; h_{u,t}^{\text{raw}}(\mathbf{z})}{\sum_{t'=0}^{36} \big((t'-18)\,\Delta t\big)^2}, \tag{10.15}$$

the slope row of the line's design applied to the tiles, with window 18 the record's center; with the readings' approaches in place of the state's it is the dashboard's own number. Its gradient is the same weighted sum of the tiles' gradients $\partial h_{u,t}^{\text{raw}}/\partial \mathbf{z}$, each carrying the approach's slopes (5.19) at window t in that window's columns, and its variance is $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$. One choice enters here: the split of the dashboard's number is sequential. With $\tilde{\mathbf{z}}^0$ the state with every window's log-derates set to the commissioned ones and no drift, $\tilde{\mathbf{z}}^g$ the fitted coils with no drift, and $s_u^{\text{raw}}(\mathbf{y})$ the slope from the readings,

$$s_u^{\text{raw}}(\mathbf{y}) = \underbrace{s_u^{\text{raw}}(\tilde{\mathbf{z}}^0)}_{\text{load and season}} + \underbrace{\big[s_u^{\text{raw}}(\tilde{\mathbf{z}}^g) - s_u^{\text{raw}}(\tilde{\mathbf{z}}^0)\big]}_{\text{coils}} + \underbrace{\big[s_u^{\text{raw}}(\hat{\mathbf{z}}) - s_u^{\text{raw}}(\tilde{\mathbf{z}}^g)\big]}_{\text{sensor}} + \underbrace{\big[s_u^{\text{raw}}(\mathbf{y}) - s_u^{\text{raw}}(\hat{\mathbf{z}})\big]}_{\text{reading errors}}, \tag{10.16}$$

the first term Chapter 5's unaged counterfactual in every window, by equation (5.20) the heat ratio of window 0 to window t. The order of the steps moves the terms by second-order amounts; the split is used only to take the dashboard's number apart, as Chapter 5's unaged counterfactual and Chapter 7's split were.

Building blocks. The pair $(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ of Section 10.2.3; the drift (10.2); the approach (5.18) and its slopes (5.19); the slope row (10.15); the split (10.16).

The dashboard's slope from the state. Carried through its gradient, the slope of coil 1's line is -0.1546 per year, with standard deviation 0.0073 and band $[-0.1667,\ -0.1425]$, against the readings' -0.1576; coil 2's is -0.1349, with standard deviation 0.0074 and band $[-0.1470,\ -0.1227]$, against -0.1482. Table 10.5 takes the readings' slopes apart by equation (10.16). Each row is a coil; the columns give the slope, its four terms per year in units of the commissioned value, and each term's share of the slope.

Table 10.5: The dashboard's slope of each coil, from the readings, split by equation (10.16); per year, in units of the commissioned value, with each term's share of the slope.

Coil [–]	Readings' slope [yr^{-1}]	Load and season [yr^{-1}] (%)	Coils [yr^{-1}] (%)	Sensor [yr^{-1}] (%)	Reading errors [yr^{-1}] (%)
1	−0.1576	−0.0242 (15.3)	−0.0920 (58.4)	−0.0384 (24.4)	−0.0030 (1.9)
2	−0.1482	−0.0242 (16.3)	−0.1107 (74.7)	0	−0.0133 (9.0)

Load and season alone pull each tile down by 0.0242 a year, with standard deviation 0.00335: at the MAP state the heat to remove grew from 120.300 kW at window 0 to 141.186 kW at window 36, and the line through it rises by 3.562 kW a year, 2.96 % of window 0's. Within year 3 the unaged tile swings between 0.8310 and 0.9130 with the months. The coils' term is 2.17 times coil 1's own rate on the same scale, $\hat{\gamma}_1 = -0.0424$, and 1.77 times coil 2's, -0.0625: by equation (5.19) the approach's slope on its own coil is -2.1020 for coil 1 and -1.7856 for coil 2 at window 36 (-2.7925 and -2.6507 at window 0), and on the other coil -0.4803 and -0.5197 (-0.4954 and -0.5046). The tile magnifies its coil's own change about twice and charges it with

about half of the other's, Chapter 5's finding carried along the record. And a quarter of coil 1's slope, 24.4 %, is its drifting sensor.

The readings taken as exact through the physics, Chapter 5's $G_{u,\text{test}} = C_u(y_{T_r} - y_{T_{d,u}})/(y_{T_r} - T_s)$ in every window with no probability, have a log-slope of −0.08144 and −0.06180 per year, −7.82 and −5.99 %/yr: the physics takes out the load, the season and the other coil, and leaves coil 1's line carrying the drift and both carrying the readings' scatter.

The tiles and the coils' own index are set side by side in Figure 10.4, one panel per coil. The horizontal axis is the window and the vertical axis an index against commissioning; the dots are the dashboard's 37 tiles from the readings and the gray line its least-squares line; the shaded band and the solid line are the health index $h_{u,t} = \mathrm{e}^{g_{u,t} - g_{u,0}}$ of the path, its median and 5–95 % band in every window; the dashed curve is the tile that unaged coils and a sound sensor would show, load and season alone. By window 36 the tiles lie far below the coil's own index, 0.4873 against 0.8806 for coil 1 and 0.5947 against 0.8289 for coil 2, and the dashed curve, which carries no aging at all, swings with the load and the season between 0.8105 and 1: the line's slope is mostly the coils' change, magnified; not their rate.

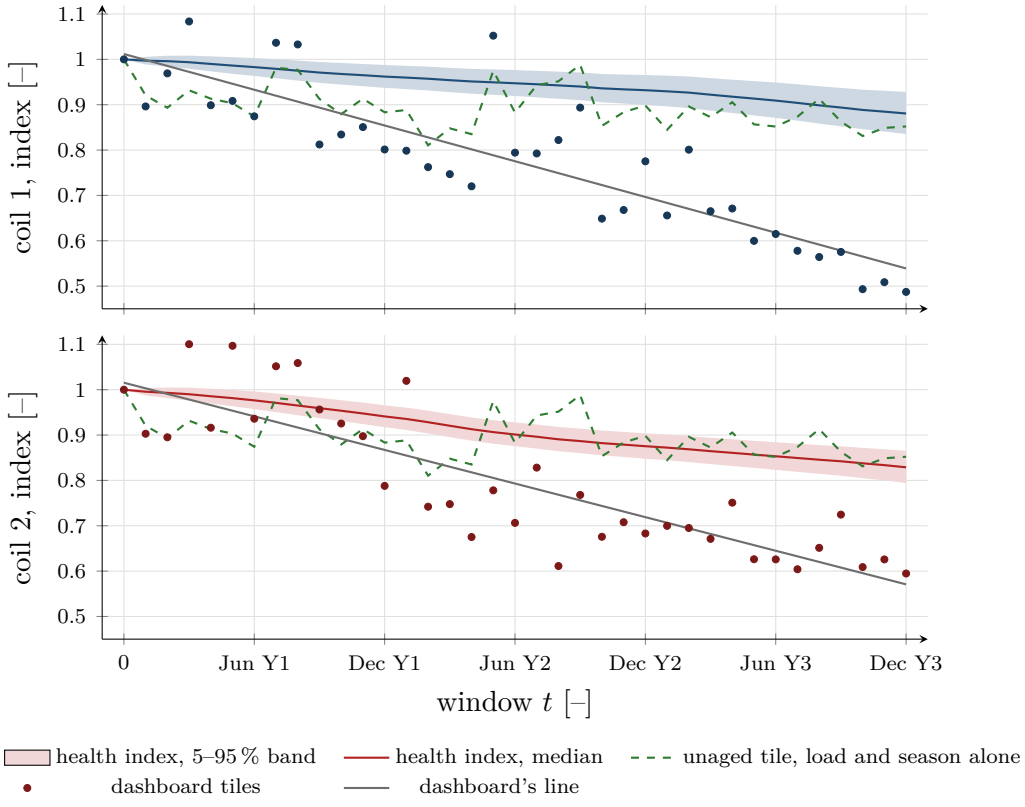


Figure 10.4: The dashboard against the coil, per coil: coil 1 (top, blue) and coil 2 (bottom, red); the legend applies to both panels. Dots: the dashboard's tiles $h_{u,t}^{\text{raw}}$ from the readings. Gray line: their least-squares line, $1.0118 - 0.1576\,t/12$ and $1.0154 - 0.1482\,t/12$. Solid line and shaded band: the health index $h_{u,t} = \mathrm{e}^{g_{u,t} - g_{u,0}}$ of the stacked state, median and 5–95 % band. Dashed green: the tile of unaged coils and a sound sensor at the MAP loads and the known inputs, the heat ratio of equation (5.20), the same for both coils.

10.3.4 Each coil's health index at window 36

Asset manager's question. How far has each coil fallen since commissioning, read with the whole record? It is Chapter 5's question asked today, with every window in the fit.

General form. An estimation question about a derived quantity asks for its posterior distribution, equation (2.9),

$$\Pr(f(\mathbf{z}) \le f^* \mid \mathbf{y}) = \int \mathbf{1}\big[f(\mathbf{z}) \le f^*\big]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z},$$

summarized by its median and its credible interval, equation (2.11), at the levels 0.05 and 0.95.

Specialized form. The question is equations (2.9) and (2.11) with $f(\mathbf{z}) = \ln h_{u,36} = g_{u,36} - g_{u,0}$, Chapter 5's index (5.1) on the two entries of the stacked state, and $p(\mathbf{z} \mid \mathbf{y})$ the Laplace normal. Its gradient has -1 in the column of $g_{u,0}$ and $+1$ in that of $g_{u,36}$, and its variance is the change-since-commissioning variance (3.38),

$$\mathrm{Var}(\ln h_{u,36} \mid \mathbf{y}) = \Sigma_{g_{u,0}g_{u,0}} + \Sigma_{g_{u,36}g_{u,36}} - 2\,\Sigma_{g_{u,0}g_{u,36}}, \tag{10.17}$$

with the median and the band of equation (5.15). By equation (10.7) the same difference is $36\,\Delta t\,\hat{\gamma}_u = 3\,\mathrm{yr} \times \hat{\gamma}_u$ at the MAP state, so the index today and the rate are one fact read two ways; their spreads differ, because the index does not carry the wander of the trend rows the rate is read through.

Building blocks. The entries $\hat{g}_{u,0}$, $\hat{g}_{u,36}$ and their block of $\boldsymbol{\Sigma}$ from Section 10.2.3; the variance (10.17); the quantiles (5.15).

The index of each coil. For coil 1,

$$\begin{aligned}
\widehat{\ln h}_{1,36} &= -0.13314 - (-0.00594) = -0.12719,\\
\mathrm{Var}(\ln h_{1,36} \mid \mathbf{y}) &= 2.3812 \times 10^{-4} + 5.6927 \times 10^{-4} - 2\,(-1.0913 \times 10^{-4})\\
&= 1.0256 \times 10^{-3} = 0.03203^2,
\end{aligned}$$

so the index has median 0.8806 and 5–95 % band [0.8354, 0.9282]; the full posterior gives a median of 0.8814. For coil 2,

$$\begin{aligned}
\widehat{\ln h}_{2,36} &= -0.21208 - (-0.02448) = -0.18760,\\
\mathrm{Var}(\ln h_{2,36} \mid \mathbf{y}) &= 2.4273 \times 10^{-4} + 2.8577 \times 10^{-4} - 2\,(-7.3606 \times 10^{-5})\\
&= 6.7572 \times 10^{-4} = 0.02599^2,
\end{aligned}$$

median 0.8289 and band [0.7942, 0.8652], with 0.8284 from the full posterior. The three years' change of the line, $3\,\mathrm{yr} \times \hat{\gamma}_u$, is -0.12719 and -0.18760, equal to the indices' logarithms to the fifth decimal, as equation (10.7) requires. The cross term is now positive in the budget: a coil's two entries are anticorrelated, by -0.2964 and -0.2795, through the rate that ties them, so their difference is the wider for it. In Chapter 5's two windows the two entries were correlated positively, through β_1 alone.

The commissioning floor lifted. Chapter 5's two-window state, read at windows 0 and 36 with the drift as one bias entry whose prior is the base fit's carried to window 36, 2.8333 K, gives the commissioning entries spreads of 0.03970 and 0.04082 (its fit of windows 0 and 36), and the indices $h_{1,36} = 0.8372$ with band [0.7290, 0.9615] ($\mathrm{sd}\ln h = 0.08415$) and $h_{2,36} = 0.8676$ with band [0.7838, 0.9604] (0.06177), with the drift left unresolved at 0.6726 ± 0.6697 K. The whole record gives the commissioning entries spreads of 0.01543 and 0.01558, and index bands 0.381 and 0.421 times as wide. Every window now informs the baseline through the trend rows, the sixth choice of step 2, and the floor that Chapter 5 found under every later index (Section 5.6.3) is lower when the windows between are read.

The capacity margin today. By equation (5.21) on window 36's entries, the hall's total conductance is $G_1 + G_2 = 11.7897\,\mathrm{kW/K}$ with standard deviation 0.1193, a margin of $1.0288\,\mathrm{kW/K}$ above $G^* = 10.761\,\mathrm{kW/K}$, 8.62 standard deviations. When the margin closes is Chapter 14's question.

10.3.5 The drift of unit 1's sensor

Asset manager's question. Does the record still need the drift of unit 1's sensor, and how large is it now? Chapter 7 could not decide at window 25 whether coil 1 or its sensor had moved; the question asks it again of 37 windows.

General form. A hypothesis comparison is the ratio of the evidences of two drawings of the same record, the Bayes factor of equation (2.14),

$$B_{AB} = \frac{p(\mathbf{y} \mid M_A)}{p(\mathbf{y} \mid M_B)},$$

and the size of a quantity under the chosen drawing is the estimation form (2.9) with its credible interval.

Specialized form. The comparison is equation (2.14) with M_A the base drawing and M_B the same drawing without the drift entry, $\dot{b}_{d,1}$ pinned at zero. Each evidence is Chapter 7's Laplace evidence (7.7) on its own drawing, at its own MAP state and with its own information $\mathbf{J}^\mathsf{T}\mathbf{J}$, the loads and the trend entries entering with the flat priors' common constant, which cancels. The two drawings are nested, so their log ratio is equation (7.8) with $\sigma = \sigma_{\dot{b}}$: the gain in fit, half the drop in $2\mathcal{L}$, less the Occam factor. The size is the estimation form (2.9) with $f(\mathbf{z}) = b_{d,1}(36) = 17\,\Delta t\,\dot{b}_{d,1}$, by equation (10.2), whose gradient is $17\,\Delta t$ in the column of $\dot{b}_{d,1}$, so that its standard deviation is $17\,\Delta t\sqrt{\Sigma_{\dot{b}\dot{b}}}$. One choice enters here: the two drawings have equal prior probabilities, which lets the evidence alone decide, and it fails where the asset manager's experience favors a sound sensor.

Building blocks. The base fit of Section 10.2.3 and the fit of the same rows without $\dot{b}_{d,1}$; their objectives and $\ln\det \mathbf{J}^\mathsf{T}\mathbf{J}$ in equations (7.7) and (7.8); the entry $\hat{\dot{b}}_{d,1}$ and $\Sigma_{\dot{b}\dot{b}}$; the drift (10.2).

The evidence for the drift. The two log evidences are -68.9335 with the drift and -75.0049 without it, so

$$\ln B(\text{drift} : \text{none}) = -68.9335 - (-75.0049) = +6.0714, \qquad B = 433,$$

and at even prior odds the drift has posterior probability 0.9977. The gain in fit is 8.3648 and the Occam factor 2.2934. Importance sampling gives $\ln B = +6.0725$. What Chapter 7's one window could not decide, the record decides: cut at window 25, the same rows give $\ln B = -0.4477$ and a drift of $0.3500\,\mathrm{K}$ at window 25 with standard deviation 0.2722. The windows after 25, and the readings of every window that do not pass through unit 1's sensor, are what separate the drift from coil 1.

The size of the drift. The drift rate is $\hat{\dot{b}}_{d,1} = 0.6837\,\mathrm{K/yr}$ with standard deviation 0.1681 and band $[0.4072,\ 0.9602]\,\mathrm{K/yr}$, and $\Pr(\dot{b}_{d,1} > 0 \mid \mathbf{y}) = 0.99998$; the full posterior's median is 0.6916. By equation (10.2) the drift at window 36 is $1.4167\,\mathrm{yr} \times 0.6837 = 0.9686\,\mathrm{K}$, with standard deviation 0.2382 and band $[0.5768,\ 1.3603]\,\mathrm{K}$, and at window 25 it is $0.3419\,\mathrm{K}$ (0.0841). Without the drift entry the record reads coil 1 at $-7.068\,\%/\mathrm{yr}$, band $[-8.973,\ -5.124]$, and coil 2 at $-4.669\,\%/\mathrm{yr}$, band $[-6.627,\ -2.669]$: the drift is 2.917 points of coil 1's apparent rate,

and through the return air the drift entry moves 1.393 points onto coil 2's. The order of the two coils turns with it: without the drift entry, coil 2 is the faster with probability 0.0911 only, against 0.8179 with it. The dashboard's order is the order of the drawing that takes the sensor as sound.

10.4 The checks (step 5)

Each answer of Section 10.3 rests on approximations, priors and modeling choices that this record lets be tested against the calculation they stand for, and on one constructed record whose truth is known.

10.4.1 The full posterior against the Laplace answers

The reweighted draws of Section 10.2.5 have been set beside every answer of step 4. Their Monte Carlo standard errors, by batch means, are at most $0.00005\,\mathrm{yr}^{-1}$ on the medians of the trend entries, at most 0.00009 on their band ends, 0.000128 on coil 1's probability of a loss, 0.00142 on the probability that coil 1 is the faster, 0.0001 on the medians of the indices, 0.0006 on the drift rate's median, and 0.00027 and 0.00030 on the two log evidences. Every median and band end of the trend entries lies within $0.00045\,\mathrm{yr}^{-1}$ of the Laplace value, $0.045\,\%/\mathrm{yr}$ on the rate, and the posterior standard deviations are 0.01453 and 0.01323 against 0.01463 and 0.01323. One gap is resolved and real: the full posterior's $\Pr(\gamma_1 < \gamma_2 \mid \mathbf{y})$ is 0.17370, against the Laplace 0.18214, a gap of 0.00844, 5.9 standard errors. The full posterior puts slightly more weight on coil 2 being the faster, 0.8263 against 0.8179, and the report gives both. The medians of the indices, 0.8814 and 0.8284 against 0.8806 and 0.8289, and of the drift rate, 0.6916 against $0.6837\,\mathrm{K/yr}$, lie further from the Laplace values than their standard errors: the full posterior is slightly skewed, and the report gives these medians beside the Laplace values. The importance-sampling evidences are -68.9375 against the Laplace -68.9335 with the drift and -75.0099 against -75.0049 without it (effective sample size $99.2\,\%$), so $\ln B$ is $+6.0725$ against $+6.0714$. The Laplace normal is the full posterior to the precision of the report, except in the probability of the coils' order, 0.818 against 0.826.

10.4.2 The limit of exact rows and the curvature term

The fit runs on the reduced state, which equation (4.48) makes the limit $\sigma_p \to 0$ of the full state. The check fits the chapter's own full state of 300 entries, the 115 of equation (10.3) and each window's $(T_r, T_{d,1}, T_{d,2}, T_c, T_1)$, with the 185 physical rows (5.2) at a common finite tolerance beside the readings, the trend rows and the priors, and lets the tolerance shrink. Table 10.6 gives the result: each row is a tolerance, the columns the two trend entries and the drift rate with their standard deviations and the iterations the fit took, and the last row the reduced state. At 0.1 the spreads are wider by the discrepancy the rows allow and the entries move in the fourth decimal; at 0.01 and 0.001 the full state reproduces the reduced state to the printed digits.

The covariance is the Gauss–Newton form, without the curvature term $\mathbf{C} = \mathbf{H} - \mathbf{J}^{\mathsf{T}}\mathbf{J}$ of equation (4.44). The full Hessian of the objective at $\hat{\mathbf{z}}$, computed by differences, gives the Newton covariance, in which the standard deviations are 0.01458 and $0.01321\,\mathrm{yr}^{-1}$ for the trend entries and $0.16778\,\mathrm{K/yr}$ for the drift rate, against Gauss–Newton's 0.01463, 0.01323 and 0.16812, ratios of 0.9965, 0.9986 and 0.9980. The largest entry of $\mathbf{C}$ is 8.81×10^{-4} of the largest entry of $\mathbf{J}^{\mathsf{T}}\mathbf{J}$, and $\ln\det\mathbf{H} - \ln\det\mathbf{J}^{\mathsf{T}}\mathbf{J} = 0.0119$, which would move a log evidence by -0.0060. The Gauss–Newton covariance is the one reported. The state that also carries racks 2 and 3, with their survey priors, moves no entry of the base fit by more than 1.0×10^{-7} and no entry of

Table 10.6: The limit of exact physical rows on the 300-entry full state of the record: the trend entries and the drift rate with their standard deviations at three tolerances, against the reduced state.

Tolerance σ_p [kW or K]	γ_1 [yr^{-1}]	sd [yr^{-1}]	γ_2 [yr^{-1}]	sd [yr^{-1}]	$\dot{b}_{d,1}$ [K/yr]	sd [K/yr]	Iterations [–]
0.1	−0.04219	0.01469	−0.06254	0.01324	0.6868	0.1694	7
0.01	−0.04240	0.01463	−0.06253	0.01323	0.6837	0.1681	5
0.001	−0.04240	0.01463	−0.06253	0.01323	0.6837	0.1681	13
reduced state	−0.04240	0.01463	−0.06253	0.01323	0.6837	0.1681	–

its $\mathbf{\Sigma}$ by more than 2.3×10^{-7}; both fractions keep the survey's spread 0.0800 and are correlated with no base entry. Their block is the prior's, as Section 10.1.2 stated.

10.4.3 Sensitivity to the priors and the choices

Table 10.7 repeats the fit under other priors and other choices of step 2. Each row is one variant; the columns give each coil's rate and the standard deviation of its trend entry, the probability that coil 2 is the faster, and the drift rate.

Table 10.7: The rates under other priors and choices. Rates by equation (10.1); $P_{2>1}$ is $\Pr(\gamma_2 < \gamma_1 \mid \mathbf{y})$.

Variant	rate 1 [%/yr]	sd γ_1 [yr^{-1}]	rate 2 [%/yr]	sd γ_2 [yr^{-1}]	$P_{2>1}$ [–]	$\dot{b}_{d,1}$ [K/yr]
the base fit	−4.151	0.01463	−6.062	0.01323	0.8179	0.6837
rating prior sd 0.05	−4.131	0.01444	−6.132	0.01307	0.8332	0.6864
rating prior sd 0.30	−4.153	0.01465	−6.055	0.01325	0.8162	0.6834
drift prior 0.7059 K/yr (1.0 K at window 36)	−4.289	0.01454	−5.995	0.01320	0.7926	0.6515
drift prior 4.0 K/yr	−4.136	0.01464	−6.069	0.01324	0.8206	0.6873
drift prior 100 K/yr (flat)	−4.130	0.01464	−6.072	0.01324	0.8215	0.6886
survey sd 0.04 on β_1	−4.163	0.01462	−6.056	0.01323	0.8157	0.6808
$\sigma_q = 0.0025$	−3.894	0.01118	−6.324	0.00947	0.9263	0.7121
$\sigma_q = 0.01$	−4.673	0.02338	−5.406	0.02239	0.5892	0.6219
straight line, $\sigma_q \to 0$	−3.787	0.00970	−6.421	0.00777	0.9605	0.7227
drift onset at window 16	−4.412	0.01479	−5.961	0.01329	0.7664	0.4738
drift onset at window 24	−4.207	0.01432	−5.990	0.01311	0.8070	0.9759
drift from commissioning (onset at window 1)	−5.564	0.01415	−5.405	0.01310	0.4689	0.1689
no drift entry	−7.068	0.01259	−4.669	0.01262	0.0911	–
window 20's unit-2 discharge reading left out	−4.285	0.01465	−5.975	0.01322	0.7888	0.6573
the fleet's prior on γ_u, $\mathcal{N}(-0.0302,\ 0.0120^2)$ (Chapter 9)	−3.669	0.00920	−4.493	0.00879	0.7382	0.6786

The priors' effects read coil by coil. The rating and survey priors leave coil 1 between −4.131 and −4.163 %/yr and coil 2 between −6.132 and −6.055: the record, not the priors, sets the path. A tighter drift prior gives more of unit 1's rising reading to coil 1, −4.289 %/yr, and through the return air less to coil 2, −5.995; a wider one moves both the other way, by less. The wander moves the two coils in opposite directions. The rate is the path's end-to-end change, equation (10.7), and with a larger σ_q each end follows its own windows more closely: coil 1's last windows lie low, so it steepens, −4.673, and coil 2's eases, −5.406, shifts of −0.37 and +0.53 posterior standard deviations. Window 20's unit-2 discharge reading, the largest normalized residual, left out, gives −4.285 and −5.975 %/yr. A drift from commissioning reverses the order

of the coils, but the record rejects it (Section 10.4.4). The fleet's prior of Chapter 9, the design's population of rates, pulls both coils toward its mean of $-0.0302\,\mathrm{yr}^{-1}$, coil 1 by $+0.34$ and coil 2 by $+1.25$ posterior standard deviations, because coil 2's record lies the farther from the fleet; partial pooling of the hall's coils with the fleet is not this chapter's drawing, and the row stands as a sensitivity.

10.4.4 The wander and the drift's onset by the evidence

The wander. Step 2 chose $\sigma_q = 0.005$ and did not re-choose it from the record. Table 10.8 sets the evidence over it beside the choice: the base drawing at ten spreads and on the straight line. Each row is a spread; the columns give the Laplace log evidence, twice the objective at the fit, the two trend entries with their standard deviations, and the drift rate.

Table 10.8: The base drawing at ten spreads of the trend rows and on the straight line: its Laplace log evidence and its fit.

σ_q [–]	$\ln p(\mathbf{y} \mid M)$ [–]	$2\mathcal{L}(\hat{\mathbf{z}})$ [–]	γ_1 [yr^{-1}]	sd [yr^{-1}]	γ_2 [yr^{-1}]	sd [yr^{-1}]	$\dot{b}_{d,1}$ [K/yr]
0.0010	−69.9335	160.4491	−0.03880	0.00995	−0.06620	0.00807	0.7211
0.0020	−69.7750	159.6310	−0.03934	0.01068	−0.06570	0.00890	0.7159
0.0030	−69.5353	158.3702	−0.04018	0.01177	−0.06488	0.01012	0.7075
0.0050 (the base fit)	−68.9335	154.9755	−0.04240	0.01463	−0.06253	0.01323	0.6837
0.0075	−68.2323	150.0728	−0.04535	0.01885	−0.05899	0.01769	0.6506
0.0100	−67.8027	145.2227	−0.04786	0.02338	−0.05557	0.02239	0.6219
0.0125	−67.6877	140.7173	−0.04976	0.02805	−0.05267	0.02719	0.5995
0.0150	−67.8480	136.5938	−0.05114	0.03280	−0.05033	0.03204	0.5826
0.0200	−68.7854	129.3314	−0.05284	0.04244	−0.04714	0.04183	0.5600
0.0300	−72.1356	117.6538	−0.05432	0.06197	−0.04448	0.06154	0.5381
straight line	−69.9890	160.7335	–	–	–	–	–

The evidence peaks at $\sigma_q = 0.0125$, with $\ln B = +1.2458$ ($B = 3.48$) against the base fit's 0.005, which itself leads the straight line by 1.0555. The chapter does not adopt the peak. At the base fit the trend rows' squared residuals sum to 2.0869 for coil 1 and 2.6153 for coil 2, above their expectations under the drawing, the rows' redundancy sums of equation (7.19) restricted to the trend rows, 1.5980 and 1.5288, and the excess sits at the months of Section 10.2.4: coil 2's windows 13 to 16 and coil 1's year 3. A larger wander absorbs those months. Whether they are wander, a change of rate or something else is what Chapters 11 and 12 test, and a σ_q set by them here would answer those questions in advance. Fresh records (Section 10.4.6) show what 0.005 does to the band at this record's truth.

The drift's onset. Step 2 fixed the drift's first window at 20. Table 10.9 moves it: each row is an onset, with the log evidence of the base drawing carrying the drift from there, the two trend entries and the drift rate with its standard deviation.

The log evidence rises from -75.1833 at onset 1, no better than no drift at all, to -68.9335 at window 20 and -65.8724 at window 28, then falls, to -69.0260 at window 32. The record cannot place the onset: a later, steeper drift, 1.5370 K/yr from window 28, and coil 1's steep months of year 3 move the same reading, and the record has no row that tells them apart. The rates do not depend on it: over onsets 16 to 28, γ_1 lies in $[-0.04573,\ -0.04223]$ and γ_2 in $[-0.06253,\ -0.05985]$, within 0.23 and 0.20 posterior standard deviations of the reported values. The drift's size does, from 0.4738 to 1.5370 K/yr. The chapter keeps window 20, and the drift rate it reports is conditional on that onset.

Table 10.9: The base drawing with the drift carried from each onset window t_b: its log evidence and its fit. Without a drift entry the log evidence is −75.0049.

Onset t_b [window]	$\ln p(\mathbf{y} \mid M)$ [–]	γ_1 [yr^{-1}]	γ_2 [yr^{-1}]	$\dot{b}_{d,1}$ [K/yr]	sd [K/yr]
1	−75.1833	−0.05724	−0.05557	0.1689	0.0694
8	−73.9365	−0.05167	−0.05831	0.2557	0.0914
12	−72.9138	−0.04877	−0.05975	0.3362	0.1090
16	−71.1058	−0.04512	−0.06146	0.4738	0.1330
18	−69.7675	−0.04301	−0.06240	0.5785	0.1488
20 (the base fit)	−68.9335	−0.04240	−0.06253	0.6837	0.1681
22	−68.0769	−0.04223	−0.06240	0.8161	0.1924
24	−67.4448	−0.04299	−0.06177	0.9759	0.2239
26	−66.2250	−0.04342	−0.06126	1.2293	0.2675
28	−65.8724	−0.04573	−0.05985	1.5370	0.3317
30	−67.0114	−0.05050	−0.05740	1.8819	0.4345
32	−69.0260	−0.05677	−0.05445	2.3365	0.6216

10.4.5 The rate as the record grew

An as-of fit is the base fit of the record cut at window t_{as}, using nothing later. Table 10.10 gives it at six cuts, half a year apart. Each row is a cut; the columns give each coil's trend entry, its standard deviation and its probability of a loss, and the standard deviations of the straight line on the same windows.

Table 10.10: The base fit as of each half year: the record cut at window t_{as}, with the straight line's standard deviations on the same windows.

t_{as} [window]	γ_1 [yr^{-1}]	sd [yr^{-1}]	P_1 [–]	γ_2 [yr^{-1}]	sd [yr^{-1}]	P_2 [–]	line sd$_1$ [yr^{-1}]	line sd$_2$ [yr^{-1}]
6 (Jun Y1)	−0.01740	0.08769	0.579	+0.11215	0.08916	0.104	0.08391	0.08545
12 (Dec Y1)	−0.06167	0.03778	0.949	−0.03049	0.03817	0.788	0.03303	0.03345
18 (Jun Y2)	−0.03240	0.02372	0.914	−0.07886	0.02417	0.99945	0.01829	0.01881
24 (Dec Y2)	−0.02632	0.01994	0.907	−0.08430	0.01872	1.00000	0.01455	0.01329
30 (Jun Y3)	−0.03136	0.01706	0.967	−0.06973	0.01541	1.00000	0.01193	0.00992
36 (Dec Y3)	−0.04240	0.01463	0.998	−0.06253	0.01323	1.00000	0.00970	0.00777

The standard deviations are plotted against the cut in Figure 10.5. The horizontal axis is the last window of the fit and the vertical axis the standard deviation of the trend entry, on a logarithmic scale; the solid lines are the stacked state, the dashed lines the straight line on the same windows, and the gray curve the wander floor of equation (10.12) for a record of n_{w} monthly increments, $\sigma_q/(\sqrt{n_{\text{w}}}\,\Delta t)$, here with $n_{\text{w}} = t_{\text{as}}$: 0.02449, 0.01732, 0.01414, 0.01225, 0.01095 and 0.01000 yr^{-1} at the six cuts. From $t_{\text{as}} = 12$ to $t_{\text{as}} = 36$ the straight line's standard deviation falls by a factor 3.41 for coil 1 and 4.31 for coil 2, the square law of Chapter 2's equation (2.20) at work, by which each window's information on the slope grows with the square of its distance from the record's center; the stacked state's falls by only 2.58 and 2.88, toward the floor, which falls as $1/\sqrt{n_{\text{w}}}$. The floor bounds the stacked state only: the straight line has no wander, so its curves pass below the floor, coil 2's from window 30 and coil 1's at window 36. By window 36 the floor is 46.7 and 57.1 % of the two variances. Coil 2's trend entry moves between $t_{\text{as}} = 12$ and $t_{\text{as}} = 18$, from −0.03049 to −0.07886, as the record shows it (Chapter 12). Each as-of fit uses nothing later than its window, and none is a time at which a verdict is first entered in the log, so no spread over fresh records is needed for it.

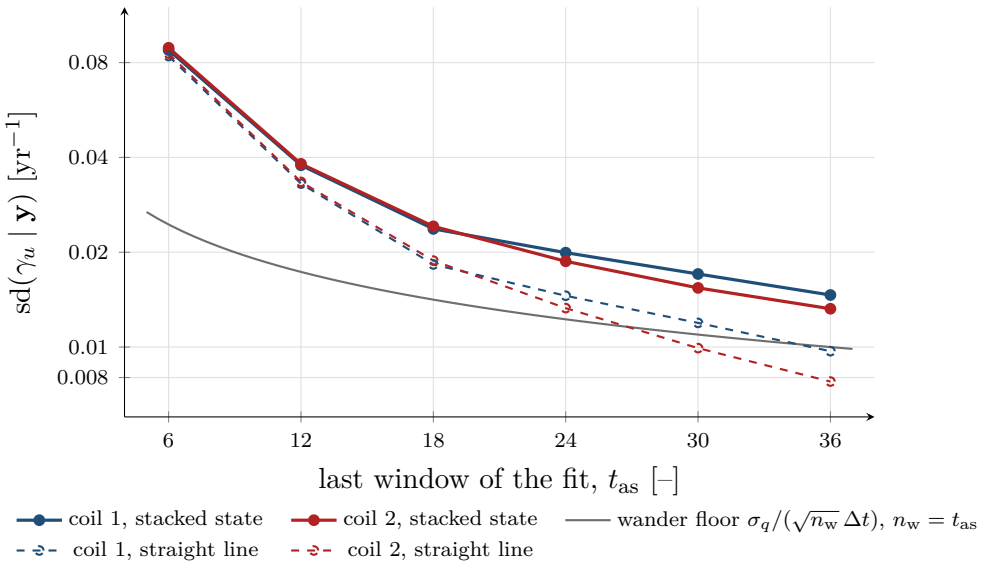


Figure 10.5: The standard deviation of each coil's trend entry as the record grows, from the as-of fits of Table 10.10, on a logarithmic scale. Solid: the stacked state at $\sigma_q = 0.005$. Dashed: the straight line on the same windows. Gray: the wander floor of equation (10.12) for $n_w = t_{as}$ increments, $0.005 \times 12/\sqrt{n_w}$. The floor bounds the stacked state only; the straight line has no wander, which is why its curves cross below the floor.

10.4.6 Fresh records at the stated truth

The record is one draw of the sensors' errors. The base drawing is refitted to 200 fresh records, the same truth and inputs with fresh errors from seeds 1000 to 1199. Its like-with-like reference is not the truth's own change: the truth's path is not a straight line in either coil, so the reference is the drawing's pseudo-true value, the base drawing fitted to the truth's noise-free readings, at which the drawing would land with no sensor error at all. Table 10.11 collects the result. Each row is a quantity over the 200 records, for each coil where it is per coil.

The trend entries scatter over fresh records with standard deviations of 0.00965 and 0.00778, against posterior standard deviations of 0.0147 and 0.0132, and the 5–95 % bands cover the pseudo-true value in 0.990 and 0.995 of the records, against the 0.90 they claim. The truth carries no wander from month to month, so the wander the trend rows allow is extra width at this truth: the chapter's band is conservative for this record, and the report says so. The straight line's bands, whose spread matches the records' scatter, cover in 0.930 and 0.890. This record's trend entries sit at the 0.350 and 0.125 quantiles of the records. Coil 2 is read the faster in 0.820 of the records, close to the posterior's 0.82. And over fresh records the whole record decides the drift, a posterior probability of 0.95 or more at even odds ($\ln B \geq 2.944$), in 0.815 of them; this record's +6.0714 sits near their median, at the 0.485 quantile.

10.4.7 The answers against the stated truth

The last check sets the answers against the stated truth of the record (Section 10.6.5). Table 10.12 collects the routes: each row is a route or a reference, the columns each coil's trend entry, its standard deviation and probability of a loss, the probability that coil 2 is the faster, and the drift rate.

The truth's path is not a straight line in either coil, and Chapters 11 and 12 say how; the base drawing's mean path is. The like-with-like reference for its trend entries is therefore the pseudo-true value, $\gamma_1^\circ = -0.03867$ (−3.793 %/yr) and $\gamma_2^\circ = -0.05318$ (−5.179 %/yr), and the posterior lies −0.25 and −0.71 standard deviations from them, inside both bands. The truth's

Table 10.11: The base drawing over 200 fresh records at the stated truth (seeds 1000–1199), and on the straight line.

Quantity [unit]	Coil 1	Coil 2
Pseudo-true value γ_u° of the drawing [yr^{-1}]	−0.03867	−0.05318
$\hat{\gamma}_u$ over the records: mean [yr^{-1}]	−0.03850	−0.05335
standard deviation [yr^{-1}]	0.00965	0.00778
5–95 % range [yr^{-1}]	−0.05351 to −0.02272	−0.06570 to −0.04006
Mean posterior standard deviation of γ_u [yr^{-1}]	0.01470	0.01320
Share of 5–95 % bands covering γ_u° [–]	0.990	0.995
This record's $\hat{\gamma}_u$ as a quantile of the records [–]	0.350	0.125
Straight line: pseudo-true value [yr^{-1}]	−0.03677	−0.05540
$\hat{\gamma}_u$ standard deviation over the records [yr^{-1}]	0.00939	0.00762
the line's posterior standard deviation [yr^{-1}]	0.00970	0.00777
share of bands covering its pseudo-true value [–]	0.930	0.890
Records reading coil 2 the faster, $\hat{\gamma}_2 < \hat{\gamma}_1$ [–]	0.820 (standard error 0.035)	
$\ln B$(drift : none): median, 5–95 % range [–]	+6.313; +0.119 to +13.644	
Records deciding the drift, $\ln B \geq 2.944$ [–]	0.815	
This record's $\ln B = +6.0714$ as a quantile [–]	0.485	

Table 10.12: The routes to the trend entries and the drift rate, with the drawing's pseudo-true values and the stated truth. P_1 is the probability that coil 1 loses conductance and $P_{2>1}$ that coil 2 is the faster; the Newton row's last entry is a standard deviation. The rates in percent per year follow from each trend entry by equation (10.1).

Route	γ_1 [yr^{-1}]	sd [yr^{-1}]	P_1 [–]	γ_2 [yr^{-1}]	sd [yr^{-1}]	$P_{2>1}$ [–]	$\dot{b}_{d,1}$ [K/yr]
Laplace, stacked state (reported)	−0.04240	0.01463	0.99813	−0.06253	0.01323	0.8179	0.6837
Full posterior, reweighted draws (medians)	−0.04208	0.01453	0.99819	−0.06280	0.01323	0.8263	0.6916
Full state, $\sigma_p = 0.1$	−0.04219	0.01469	–	−0.06254	0.01324	–	0.6868
Newton covariance	–	0.01458	–	–	0.01321	–	sd 0.16778
Straight line, $\sigma_q \to 0$	−0.03861	0.00970	0.99997	−0.06637	0.00777	0.9605	0.7227
The drawing's pseudo-true value	−0.03867	–	–	−0.05318	–	–	0.6963
Stated truth, net change per year	−0.05000	–	–	−0.04333	–	–	0.6000

net change per year, −0.05000 and −0.04333 (−4.877 and −4.241 %/yr), stands beside it: the constant-rate drawing reads a path that is not straight by a constant rate, and the order of the coils' net changes is not the order of the drawing's pseudo-true rates. The true drift rate is 0.6000 K/yr, 0.85 K at window 36, and the posterior lies +0.50 standard deviations from it, with the pseudo-true 0.6963. The true β_1 is 0.22, against 0.21272 with standard deviation 0.00893.

The health indices are where the drawing's limit shows. Coil 1's true index at window 36 is 0.8607 ($\ln h = -0.1500$), and the posterior lies +0.71 standard deviations from it. Coil 2's true index is 0.8781 (−0.1300), and the posterior lies −2.22 standard deviations from it, but −1.08 from the drawing's pseudo-true −0.15953. The difference is the drawing's, not the record's draws: fitted to the truth's noise-free readings, the base drawing itself moves coil 2's window-36 entry from the true −0.1700 to −0.19042 and coil 1's from −0.1700 to −0.13113, because the drift entry and coil 1's steep months of year 3 share unit 1's discharge reading, and the return air hands the difference to coil 2. The total is right: $G_1 + G_2$ at window 36 lies −0.18 standard deviations from the true 11.8113 kW/K. A drawing that frees the form of coil 1's path is Chapter 11's. Window 20's unit-2 discharge reading, the largest normalized residual of Section 10.2.3, reads 17.59°C against the truth's 16.4516°C, an error of +3.79 standard deviations and the largest of all 185 readings: a large draw of the sensor, kept in the fit. The dashboard's slope at the true state is −0.1451 and −0.1216 per year, and that of unaged coils at the true loads −0.0228: the dashboard's distortion is the plant's, not the sensors'.

10.5 The report (step 6)

The dashboard's 15.8 and 14.8 % a year are not rates of the coils. Read on the conductance, coil 1 is losing 4.2 %/yr (1.8 to 6.4) and coil 2 6.1 %/yr (4.0 to 8.1), and both are losing with near certainty. Of the dashboard's slope, about a sixth is the load and the season, a quarter of coil 1's is its drifting sensor, and the rest is the coils' own loss magnified about twice. The order is reversed: under the base fit's constant rates, coil 2, not coil 1, is losing faster, with probability 0.82 (0.83 by the full posterior); with the sensor taken as sound it would have been coil 1. The record now needs the drift that Chapter 7 could not decide, with a Bayes factor of 433: about 0.7 K/yr, 0.97 K by today. And nearly half of coil 1's variance and more than half of coil 2's is the wander floor, which no sensor removes; only more years lower it. At this record's truth the band is conservative.

10.5.1 The answers and their bands

Table 10.13 gives the answers in the order of Table 10.2. Each row is a question; the second column gives the reported value from the Laplace normal or the Laplace evidence, the third the full posterior's value beside it, and the fourth the band: the 5–95 % band of a rate, an index, a slope or the drift, and for a probability its range over the prior rows of Table 10.7 (the rating, drift and survey priors).

The answers rest on the stacked state's posterior and on one choice above all, the wander σ_q: it sets the floor, about half of each rate's variance, and Section 10.4.3 shows that it also sets the order of the two coils more than any prior does.

10.5.2 The entry in the maintenance log

The rates are the next entry of the two-unit hall's maintenance log, after the commissioning test of window 0, the assessments of window 18 and the assessment of window 25. The entry uses nothing later than window 36, and in the values of Table 10.13 it reads:

> *Window 36 (Dec Y3), degradation rates from the condition record (windows 0–36), against the commissioning baseline.*

Table 10.13: The report for the two-unit hall's condition record at window 36 (Dec Y3), windows 0–36. Reported values from the Laplace normal of the stacked state and the Laplace evidence; bands are 5–95 %, or for a probability its range over the prior rows of Table 10.7.

Asset manager's question	Reported value	Full posterior	Band
How many percent of conductance per year is it losing?	coil 1: −4.151; coil 2: −6.062 [%/yr]	medians −4.121; −6.087 [%/yr]	coil 1: −6.430 to −1.817; coil 2: −8.084 to −3.995 [%/yr]
Is it losing conductance at all, beyond three years of scatter?	coil 1: 0.998126; coil 2: $1 - 1.1 \times 10^{-6}$ [–]	0.998186; 1.000000 [–]	coil 1: 0.99802 to 0.99871; coil 2: 1.00000 in every row [–]
Which coil is losing conductance faster, and how sure is that?	coil 2, probability 0.8179 [–]	0.8263 [–]	0.7926 to 0.8332 [–]
Is the dashboard's 15 % a year the coils, or the load, the season and the sensor?	slope from the state −0.1546 and −0.1349 [yr^{-1}]; of the readings' −0.1576: load and season −0.0242, coils −0.0920, sensor −0.0384; of −0.1482: −0.0242, −0.1107, 0 [yr^{-1}]	–	−0.1667 to −0.1425; −0.1470 to −0.1227 [yr^{-1}]
How far has each coil fallen since commissioning, read with the whole record?	coil 1: $h = 0.8806$; coil 2: $h = 0.8289$ [–]	0.8814; 0.8284 [–]	coil 1: 0.8354 to 0.9282; coil 2: 0.7942 to 0.8652 [–]
Does the record still need the drift of unit 1's sensor, and how large is it now?	yes: $B = 433$, $\ln B = +6.0714$, Pr(drift) = 0.9977 [–]; $\dot{b}_{d,1} = 0.6837$ [K/yr], $b_{d,1}(36) = 0.9686$ [K]	$\ln B = +6.0725$ [–]; median $\dot{b}_{d,1} = 0.6916$ [K/yr]	$\dot{b}_{d,1}$: 0.4072 to 0.9602 [K/yr]; $b_{d,1}(36)$: 0.5768 to 1.3603 [K]

Coil 1: −4.2 %/yr (5–95 % band −6.4 to −1.8); loss certain ($P = 0.998$). Health index today 0.88 (0.84 to 0.93).
Coil 2: −6.1 %/yr (−8.1 to −4.0); loss certain. Health index 0.83 (0.79 to 0.87). Under constant rates, coil 2 loses faster than coil 1, $P = 0.82$ (0.83 by the full posterior).
Unit 1's discharge sensor: drift carried from window 20, 0.68 K/yr (0.41 to 0.96), 0.97 K today; the record requires it (Bayes factor 433).
Dashboard trend lines (−15.8, −14.8 %/yr): not a rate of the coils.
The record's path departs from a straight line in coil 2 around windows 13 to 16 and in coil 1 through year 3; not tested here.
Action: no cleaning decision from the rates alone (no decision level is set); the forecast of the capacity limit waits until the form of the path is tested.
Next measurement: the side-by-side check of unit 1's discharge sensor (the entry of window 25, still open), and the record continued with a permanent instrument, a second discharge sensor on unit 1 or water-side heat meters on both coils; neither leads for the rates.

The entry records what the bands rest on as well as the values: 37 windows tied by trend rows of spread 0.005 a month, a drift from window 20, and a constant rate for each coil's path. Which coil to clean, and when, needs the decision level of Chapter 19 and the crossing date of Chapter 14, both in Volume II; this entry sets neither.

Common mistake

Reading a dashboard's trend line as a degradation rate. The slope of a line through raw-performance tiles is a trend of everything the tile depends on: the load, the season, both coils and the sensor. The rate is the trend entry of the coil's own conductance, read from one fit of the whole record with its band.

10.6 The next measurement (step 7)

The report leaves each rate with a 5–95 % band of about ±2 %/yr about its median, a half-band, and the drift rate with a standard deviation of 0.17 K/yr. A further measurement is one more row of the residual vector, or a run of rows in windows not yet read, and what it removes from each band is known before its value is.

10.6.1 The preposterior on the extended record

A candidate that reads the existing windows adds rows in their columns. A candidate that reads future windows needs their entries first: the state is extended by windows 37 to 48, January to December of year 4, each with its load and two log-derates, 151 entries in all, tied to the record and to each other by their own 24 trend rows and carried at the fitted line with the loads at the forecast. A future load carries no row until its window is read. With no new reading rows the extension leaves the trend entries where they are, 0.01463 and 0.01323: the new trend rows add new entries along with them. A candidate a is then rows with Jacobian lines $\mathbf{J}_a$ in the columns of the extended state, and its preposterior covariance is equation (4.65),

$$\boldsymbol{\Sigma}_a = \left(\boldsymbol{\Sigma}_{\text{ext}}^{-1} + \mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a\right)^{-1}, \qquad \text{sd}_a(\gamma_u) = \sqrt{\mathbf{e}_{\gamma_u}^{\mathsf{T}}\,\boldsymbol{\Sigma}_a\,\mathbf{e}_{\gamma_u}}, \tag{10.18}$$

with $\boldsymbol{\Sigma}_{\text{ext}}^{-1}$ the information of the base fit with the new windows' trend rows added. The width a candidate takes off each rate is $1 - \text{sd}_a/\text{sd}$ of γ_u.

The candidates open to the asset manager are of three kinds. Now, in window 36: the side-by-side check of unit 1's discharge sensor that Chapter 7 asked for, ten pairs at 0.1 K, a row on the drift alone by equation (7.27); a reference condition test, reference sensors of 0.1 K on the return air, both discharges and rack 1's inlet; and a second logged hour of every window 0 to 36 from the logger's archive, a new load entry per window with five fresh rows on the same coils. Forward: a fourth year of the record, windows 37 to 48 with their five readings each, alone or with a permanent instrument, water-side heat meters on both coils ($\sigma_W = 3.0$ kW) or a second discharge sensor on unit 1 (0.3 K), which are the permanent instruments at the place of the question, or with the side-by-side check made now. The last two lines are not measurements: every window's log-derates known exactly, over windows 0 to 36 and over windows 0 to 48.

10.6.2 The candidates

Table 10.14 ranks the candidates. Each row is a candidate; the columns give, for each coil, the standard deviation of the trend entry after it by equation (10.18) and the share of the present 0.01463 and 0.01323 that it removes, and the drift rate's standard deviation after it, against the present 0.1681 K/yr. The widths removed are compared as bars in Figure 10.6, two per candidate, ranked by coil 1's, with the 36-window floor dashed.

The ranking follows each candidate's rows. The side-by-side check is a row on $\dot{b}_{d,1}$ alone, with the line $(36 - t_b + 1)\,\Delta t/0.1 = 14.1667$ in its column: it takes the drift's standard deviation from 0.1681 to 0.0651 K/yr and helps coil 1, 12.1 %, through the drift's correlation of +0.5176 with γ_1, and coil 2 hardly at all, 3.5 %. The reference condition test reads window 36's temperatures more precisely and takes 18.0 and 9.8 % off; the second hour of every window narrows the reading of the path's ends, 12.5 and 9.5 %, and leaves the floor. A fourth year takes 23.1 and 20.3 % off, more than any measurement of the existing windows, because it lengthens the record. The coil-1 heat meter's row at window 37 has the slopes +36.4112 on $g_{1,37}$, −36.4112 on $g_{2,37}$ and 0.5201 on Q_{37}, in kW per unit entry, over $\sigma_W = 3.0$ kW: it reads the split of the heat between the coils, as in Chapter 5, which the return air cannot.

The last two rows of the table are the wander floor. With every window's conductance known exactly, the standard deviation of each trend entry falls only to $\sigma_q/(\sqrt{36}\,\Delta t) = 0.01000\,\text{yr}^{-1}$,

Table 10.14: Candidate measurements, from the posterior of the stacked state extended to window 48; the widths removed are taken from the present $\mathrm{sd}(\gamma_1) = 0.01463$ and $\mathrm{sd}(\gamma_2) = 0.01323\,\mathrm{yr}^{-1}$.

Candidate	$\mathrm{sd}_a(\gamma_1)$ [yr^{-1}]	coil 1 [%]	$\mathrm{sd}_a(\gamma_2)$ [yr^{-1}]	coil 2 [%]	$\mathrm{sd}_a(\dot{b}_{d,1})$ [K/yr]
Side-by-side check of unit 1's discharge sensor now, ten pairs (0.1 K)	0.01285	12.1	0.01277	3.5	0.0651
Reference condition test now (0.1 K on return air, both discharges, rack 1)	0.01200	18.0	0.01194	9.8	0.1163
A second logged hour of every window 0–36, from the logger's archive	0.01280	12.5	0.01197	9.5	0.1227
A fourth year of the record, windows 37–48	0.01126	23.1	0.01055	20.3	0.0858
with water-side heat meters on both coils (3.0 kW), permanent	0.01015	30.6	0.01000	24.4	0.0608
with a second discharge sensor on unit 1 (0.3 K), permanent	0.01007	31.1	0.01022	22.8	0.0555
with the side-by-side check now	0.01058	27.7	0.01035	21.8	0.0545
The path known exactly at windows 0–36 (the wander floor)	0.01000	31.6	0.01000	24.4	0.0872
The path known exactly at windows 0–48	0.00866	40.8	0.00866	34.5	0.0872

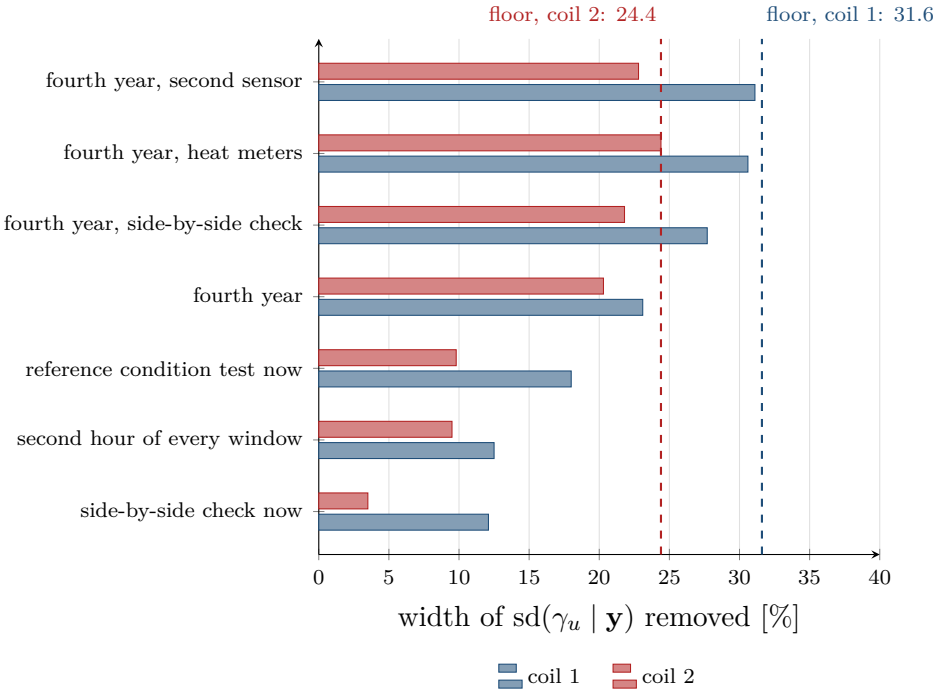


Figure 10.6: The candidate measurements of Table 10.14, ranked by the width they remove from coil 1's trend entry, with coil 2's beside each, by the preposterior (10.18). Dashed: the width removed if every window 0–36 were known exactly, the wander floor of equation (10.12), the most any measurement of the existing windows can remove.

equation (10.12), which takes 31.6 % off coil 1's band and 24.4 % off coil 2's. Every measurement of the existing 36 windows is capped there: the best of them takes 18.0 % off coil 1's. The fourth year with a permanent instrument nearly reaches the 36-window floor: with a second discharge sensor on unit 1, coil 1's standard deviation falls to 0.01007, and with heat meters on both coils, coil 2's falls to 0.01000. Only a fourth year lowers the floor itself, to 0.00866 with 48 windows. The record's length, not its sensors, sets the rate's precision. That holds for the chosen wander: on the straight line, $\sigma_q \to 0$, the floor vanishes, and the readings set the band again (Section 10.4.4 gives the evidence that keeps σ_q at 0.005).

10.6.3 The recommendation

The side-by-side check of unit 1's discharge sensor now, first: it closes the open entry of window 25, takes the drift rate's standard deviation from 0.1681 to 0.0651 K/yr, and is an hour's work with a portable reference thermometer. Then keep the record running: the fourth year takes about a fifth off each band and lowers the floor that no other measurement can. Of the permanent instruments, neither leads for the rates: a second discharge sensor on unit 1 takes 31.1 and 22.8 % off with the fourth year, and water-side heat meters on both coils 30.6 and 24.4 %; what either is worth across later windows, and for the date the capacity limit is crossed, is Chapter 23's question, in Volume II. The ten pairs' errors are taken as independent of each other and of the logged reading's, as Chapter 7 took them.

10.6.4 What the analysis showed

The dashboard said the coils were losing 15.8 and 14.8 % of their conductance a year, coil 1 the faster; read as the trend entries of one fit of the record, they are losing 4.2 and 6.1 %/yr, coil 2 the faster under the base fit's constant rates. The physics takes the load, the season and the other coil out of every window, the drift entry takes the sensor out of unit 1's reading, and the trend rows carry the rate from window to window and, in passing, lift the commissioning floor of every later index. The probability adds the band, and says that about half of it is the floor the path's wander sets, which only a longer record lowers. The record, at three years, decides the drift that one window could not. What it did not test is the path's departures from a straight line, which it shows in both coils and which the base fit folds into a constant rate; the order of the coils rests on that fold, and the stated truth's net change over the three years runs the other way, coil 1 the larger.

10.6.5 Limits of the model

A constructed record. The condition record is generated from a stated truth, with the stated sensor errors, and step 5 compares the answers with that truth (Section 10.4.7). It is not field data, and the agreement is a check of the method on its own model, not of the model against a real hall. The truth's path is not a straight line, and the drawing's mean path is.

A constant rate with wander. The spread σ_q is chosen, not estimated; the evidence leans toward a larger one because of the months a constant rate does not fit, and at this truth the chosen spread gives conservative bands.

The drift from window 20. The drift is taken as linear from a fixed onset. The record does not place the onset, and the drift's size is conditional on it; the rates are not.

Readings independent across windows. One record's errors are taken as uncorrelated from month to month; errors that persist across months would make each window worth less, which is Chapter 24's question.

Chapter 5's limits. Steady state within each window, pinned airflow with no discharge control, known inputs, β_1 unchanged since handover, and the sensors other than unit 1's discharge calibrated at commissioning hold here unchanged.

Commissioning on the same path. The trend row from window 0 to window 1 ties the commissioning test to the first month; a coil that changed between handover and January of year 1 would break that tie.

The next chapter. Chapter 11 asks of the same base fit whether a coil's rate is itself changing, read as a curvature entry in the trend rows.

10.7 What to retain

Definition 10.1 (Degradation rate from a condition record)**.** The degradation rate of a parameter is the trend entry γ_u of its path, read from one fit of the stacked state of every window, in which each window's readings act on that window's entries at its own inputs and trend rows $(g_{u,t+1} - g_{u,t} - \gamma_u \Delta t)/\sigma_q$ tie each window to the next, equation (10.5). With γ_u free, the trend rows are the only rows on it and its MAP value is the fitted path's telescoped change, $\hat{\gamma}_u = (\hat{g}_{u,n_\mathrm{w}} - \hat{g}_{u,0})/(n_\mathrm{w}\,\Delta t)$, equation (10.7). It is reported as $100(\mathrm{e}^{\gamma_u} - 1)\,\%/\mathrm{yr}$, its sign as $\Phi(-\hat{\gamma}_u/\sqrt{\Sigma_{\gamma_u\gamma_u}})$. Its variance is the path's ends' part, $\mathrm{Var}(\ln h_{u,n_\mathrm{w}} \mid \mathbf{y})/(n_\mathrm{w}\,\Delta t)^2$, plus the floor $\sigma_q^2/(n_\mathrm{w}\,\Delta t^2)$ that no measurement of the n_w windows lowers, equation (10.13). A dashboard's trend line through raw-performance tiles is a derived quantity of the same state, not a rate.

Three conclusions carry beyond the hall. A degradation rate is read on the asset's own parameter, after the physics has taken out every input of every window, and a trend of a performance figure carries the load, the season, the other units and the sensors with it. A record of many windows, tied by trend rows, informs its own beginning: the commissioning baseline is known better from the whole record than from the commissioning test and today. And the band of a rate has a part no sensor removes, the floor the path's wander sets, which only the length of the record lowers; the next measurement of a rate is often more time.

Table 10.15 lists the terms of the chapter with their expressions, units and the subsection that defines each.

Table 10.15: The terms of the degradation rate from a condition record and where each is defined. Units for the two-unit hall.

Term	Expression	Units	Meaning	Defined in
Stacked state of the condition record	eq. (10.3)	kW, –, yr^{-1}, K/yr	every window in one posterior	10.1.2
Drift rate of a sensor	$b_{d,1}(t) = \dot{b}_{d,1}(t - t_b + 1)\Delta t$, eq. (10.2)	K/yr	a bias that grows from its onset	10.1.2
Trend rows	eq. (10.5)	–	each window tied to the next	10.1.4
Sparsity pattern	eq. (10.6)	–	block-banded; γ_u seen by trend rows only	10.1.4
Degradation rate	γ_u, read as $100(\mathrm{e}^{\gamma_u} - 1)$, eq. (10.1)	yr^{-1}, %/yr	the trend entry of the path	10.1.1
Rate as the telescoped path	$\hat{\gamma}_u = (\hat{g}_{u,36} - \hat{g}_{u,0})/(36\Delta t)$, eq. (10.7)	yr^{-1}	the fitted path's end-to-end change	10.1.4
Variance budget	eq. (10.13)	yr^{-2}	the path's ends plus the floor	10.3.1
Wander floor	$\sigma_q/(\sqrt{n_\mathrm{w}}\,\Delta t)$, eq. (10.12)	yr^{-1}	no measurement of the windows lowers it	10.3.1
Dashboard's trend line	s_u^{raw}, eq. (10.15), eq. (10.16)	yr^{-1}	a derived quantity, not a rate	10.3.3
Commissioning floor lifted	eq. (10.17) on the stacked state	–	every window informs the baseline	10.3.4
As-of fit	the record cut at window t_{as}	–	the band as the record grows	10.4.5
Pseudo-true value	γ_u°, the drawing on noise-free readings	yr^{-1}	the reference when the truth lies outside the drawing	10.4.6
Preposterior on the extended record	eq. (10.18)	–	the value of a longer record	10.6.1

10.8 Exercises

Exercise 10.1. *Formulation.* Write the question "is coil 2 losing more than 5 % a year?" in the form of Table 10.2 and evaluate it for both coils. (A prediction under a threshold, equation (2.12) with $f = -\gamma_u$ and $f^* = -\ln 0.95$, since a rate below $-5\,\%/\mathrm{yr}$ is $\gamma_u < \ln 0.95$ by equation (10.1): $\Phi\big((\ln 0.95 - \hat{\gamma}_u)/\sqrt{\Sigma_{\gamma_u\gamma_u}}\big)$, 0.2716 for coil 1 and 0.8022 for coil 2.)

Exercise 10.2. *Derivation.* With γ_u free, show from the trend rows (10.5) that $\hat{\gamma}_u = (\hat{g}_{u,36} - \hat{g}_{u,0})/(36\,\Delta t)$, and that a path known exactly at every window gives $\mathrm{Var}(\gamma_u) = \sigma_q^2/(36\,\Delta t^2)$. (Only the trend rows touch γ_u, so $\partial\mathcal{L}/\partial\gamma_u = -(\Delta t/\sigma_q^2)\sum_t(g_{u,t+1} - g_{u,t} - \gamma_u\Delta t) = 0$, and the sum telescopes to $g_{u,36} - g_{u,0} - 36\,\Delta t\,\gamma_u$; with the path known, the 36 lines $-\Delta t/\sigma_q$ give the information $36\,\Delta t^2/\sigma_q^2$, so $\mathrm{Var}(\gamma_u) = 1.0000 \times 10^{-4}\,\mathrm{yr}^{-2}$, standard deviation $0.01000\,\mathrm{yr}^{-1}$.)

Exercise 10.3. *Derivation.* Show that the dashboard's tile of coils that have not aged and a sound sensor is the heat ratio $(Q_0 + Q_o(0))/(Q_t + Q_o(t))$, independent of $T_s(t)$, and give the slope of its line over the record. (By equation (5.18), $A_{u,t} = (Q_t + Q_o(t))(1 - G_{u,t}/C_u)/(G_{1,t} + G_{2,t})$ holds no T_s; with $G_{u,t} = G_{u,0}$ and no drift the last two factors cancel in $A_{u,0}/A_{u,t}$, equation (5.20) in every window; the first term of equation (10.16), through the slope row (10.15) at the MAP loads, is a line falling by 0.02417 per year.)

Exercise 10.4. *Application.* Refit the record with $\sigma_q = 0.01$ and say why coil 1's rate steepens and coil 2's slows. (The rate is the path's end-to-end change, equation (10.7), and a larger wander lets each end of the path follow its own windows more closely, so the rate leans toward

the change between the record's first and last months and away from the line through all of them; coil 1's steepest months lie at the end of its record, in year 3, and coil 2's inside it: -4.673 and $-5.406\,\%/\mathrm{yr}$, with coil 2 the faster at probability 0.5892.)

Exercise 10.5. *Application.* A ten-pair side-by-side check of unit 1's discharge sensor at window 36 reads $0.80\,\mathrm{K}$ (standard deviation $0.1\,\mathrm{K}$), against the predicted drift of $0.9686\,\mathrm{K}$, whose posterior standard deviation is $0.2382\,\mathrm{K}$. Update coil 1's rate by the correction form and by a refit. (The row is equation (7.27) with the drift $b_{d,1}(36) = 17\,\Delta t\,\dot{b}_{d,1}$, gradient $17\,\Delta t$ in the column of $\dot{b}_{d,1}$; by the rank-one update (4.66) with the innovation $0.80 - 0.9686$, the gain on γ_1 is 0.02703 per K, γ_1 moves from -0.04240 to -0.04695 and its standard deviation from 0.01463 to 0.01285: $-4.587\,\%/\mathrm{yr}$, band $[-6.583,\ -2.548]$. The refit with the row gives $-4.584\,\%/\mathrm{yr}$, standard deviation 0.01287, and $\dot{b}_{d,1} = 0.5826\,\mathrm{K/yr}$.)

Chapter 11

Accelerating Degradation

11.1 A bend in the trend chart

The hall is three years past its handover, and today is December of year 3, window 36. Chapter 10 read each coil's rate of loss from the whole condition record as a constant, and the condition-monitoring dashboard drew its straight line through the 37 health-index tiles of each coil. The dashboard also fits a curve, and the planner looks at year 3. Coil 1's tile fell by 0.288 between December of year 2 and December of year 3, against 0.026 the year before; the line through year 3's twelve tiles falls by 0.265 a year, against 0.068 in year 2 and 0.158 in year 1. The quadratic through the 37 tiles bends down, by -0.058 per yr^2, and its slope today is -0.244 a year, 1.55 times the three-year line's -0.158. Coil 2's tile does the opposite: its quadratic bends up, by $+0.049$ per yr^2, and its slope today is -0.075 a year, half its line's -0.148. The conclusion the chart invites is that coil 1's fouling has taken off in year 3, that the coil is now losing about a quarter of its commissioned performance a year and must be cleaned at the October shutdown at the latest, and that coil 2 is slowing down.

The chart cannot say how much of the bend is the sensor. Unit 1's tile reads through the discharge sensor whose drift from window 20 Chapter 10 confirmed, and a drift that starts in the middle of a record bends any line drawn through it. Nor can the chart say whether a bend of that size, in 37 scattered windows, is more than the month-to-month wander that a constant rate already allows, or whether coil 2's bend is a slower rate or something in its path in year 2, where Chapter 10 saw its steepest months, windows 13 to 16. The question of this chapter is

whether each coil's rate of loss is itself changing, read as a curvature entry on the trend rows of the condition record's fit, with the probability that the loss accelerates and the evidence that the rate changes at all.

It is answered with the method of Chapter 5, on the stacked state of Chapter 10 with one entry per coil added to the trend rows of Chapter 4, and with the evidence of Chapter 7.

Jargon: curvature of a degradation path, accelerating, constant and slowing loss, knee of a rate, the curvature's wander floor, Lindley's effect, power of an evidence test, preposterior of a Bayes factor

The *curvature of a degradation path* is the change of its rate per year, the second derivative of the log-derate path; an entry of the state that carries it is the acceleration entry. A loss is *accelerating* when its rate grows more negative with time, *slowing* when it grows less negative, and *constant* when the curvature is zero. The *knee of a rate* is a window after which the rate changes and before which it is constant. The *curvature's wander floor* is the standard deviation of a curvature whose path is known exactly at every window of the record: the part of its band that the path's own wander sets and no reading removes. *Lindley's effect* is the dependence of a Bayes factor on the prior spread of the entry it frees: as that spread widens, the factor against the larger drawing grows without limit, while the posterior of the entry within the larger drawing settles. The *power of an evidence test* is the share of fresh records, drawn at a stated truth, in which its Bayes factor passes a stated level. The *preposterior of a Bayes factor* is the distribution of the factor that a record not yet read would give, simulated from the drawings' posteriors.

Key idea

Whether a rate changes is two questions. The sign of the curvature, given that the rate may change, is a posterior probability within the drawing that frees it. Whether the record needs a changing rate at all is a Bayes factor, which charges the curvature for its prior width. The record can be sure of the first and undecided on the second. Both are read on the coil's own conductance, after the physics has taken out the load, the season and the other coil, and the drift entry the sensor.

11.1.1 A rate that may change

Chapter 10 tied each window's log-derate to the next with a constant expected increment $\gamma_u \Delta t$, the trend rows (10.5). Here the rate itself may move. With t_{y} the time since commissioning in years, window t at $t_{\mathrm{y}} = t\,\Delta t$, the rate of coil u at time t_{y} is

$$\gamma_u(t_{\mathrm{y}}) = \gamma_u + \dot{\gamma}_u\,(t_{\mathrm{y}} - \bar{t}_{\mathrm{y}}), \qquad \bar{t}_{\mathrm{y}} = 1.5\,\mathrm{yr}\ \text{(window 18)}, \tag{11.1}$$

so that $\gamma_u = \gamma_u(\bar{t}_{\mathrm{y}})$ is now the rate at the record's center and $\dot{\gamma}_u$, in yr^{-2}, the change of the rate per year: the curvature of the log-derate path, whose second derivative it is. A negative $\dot{\gamma}_u$ is an accelerating loss, the rate growing more negative with time. Over the month from window t to window $t+1$ the path's expected change is the integral of the rate over the month,

$$\int_{t\Delta t}^{(t+1)\Delta t} \gamma_u(t_{\mathrm{y}})\,\mathrm{d}t_{\mathrm{y}} = \Delta t\,\left[\gamma_u + \dot{\gamma}_u\,x_t\right], \qquad x_t = (t + \tfrac{1}{2})\Delta t - \bar{t}_{\mathrm{y}}, \tag{11.2}$$

with x_t in years the curvature's lever on that increment: the distance of the month's midpoint from the center. Rates are reported as $100(\mathrm{e}^{\gamma_u(t_{\mathrm{y}})} - 1)\,\%/\mathrm{yr}$, the form of equation (10.1), negative for a loss; a curvature is reported as $100\dot{\gamma}_u$, in percentage points of rate per year.

Four drawings of the plant make up the candidate family at the place of the question, Chapter 7's form. M_0 is Chapter 10's base fit, with both curvatures pinned at zero: the constant-rate drawing. M_1 frees coil 1's curvature, M_2 coil 2's, and M_{12} both. Every drawing carries the drift entry of unit 1's discharge sensor.

The curvature enters the trend rows only, through each increment's lever. A path's rate over the record is read from its monthly increments: γ_u is their mean and $\dot{\gamma}_u$ their slope against time. The curvature is therefore a property of the path the readings fix, as the rate was, and it can be no better known than the path's monthly increments allow.

Principle 11.1. *An acceleration is a change of the parameter's rate, read on the parameter. A bend in a performance figure is a bend in everything the figure depends on, and a sensor whose drift starts in the middle of a record bends every line drawn through its readings.*

11.1.2 The two-unit hall through its record (step 0)

The graph is Chapter 10's, Figure 10.1: Chapter 5's window graph at each of the 37 windows, with its five physical rows (5.2) at that window's known inputs $T_s(t)$ and $Q_o(t)$; the 72 trend rows between consecutive windows; the two trend entries γ_1 and γ_2, linked to every trend row of their coil; the recirculation fraction β_1, shared by all windows; and the drift node $\dot{b}_{d,1}$ on unit 1's discharge reading from window 20. This chapter adds one node per coil, the curvature $\dot{\gamma}_u$, linked to every trend row of its coil beside γ_u, with a lever x_t on each link that grows with the row's distance from window 18. Nothing else changes: no reading row sees $\dot{\gamma}_u$. The modeling choices of step 0 are Chapter 5's paragraph, cited through Chapter 10: steady state within each window, the airflow pinned with no discharge control, the season's quantities as known inputs, and one conductance per coil. The choices the curvature adds are named at step 2.

Figure 11.1 draws the trend layer of coil 1. The middle row is the chain of coil 1's window entries $g_{1,0}, \ldots, g_{1,36}$, with runs of windows drawn as dashed boxes and the trend rows as black squares between neighbors; each entry is also an entry of its window's graph, which the figure does not repeat. Above the chain, γ_1 feeds every trend row along the lower bus, and $\dot{\gamma}_1$ along the upper bus, with its lever written on four links. Below, the drift node is linked to unit 1's discharge readings from window 20.

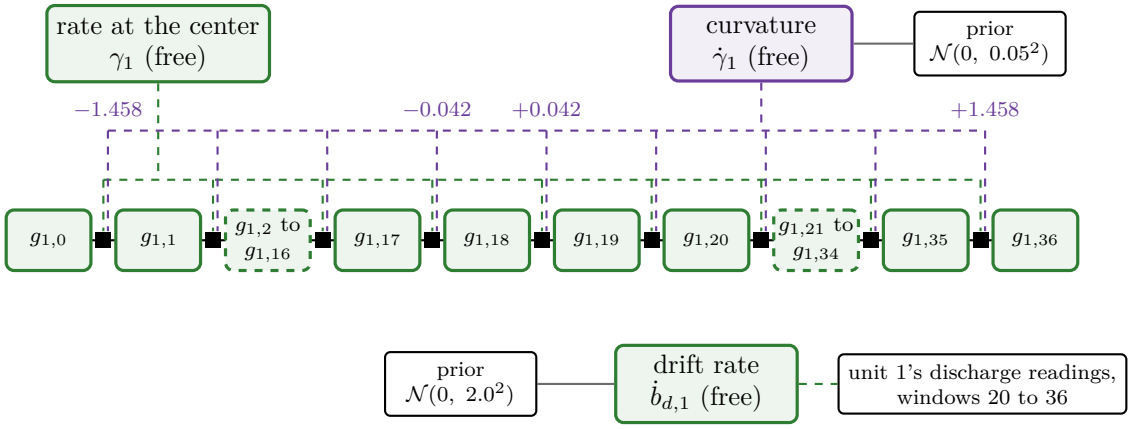


Figure 11.1: The trend layer of coil 1 in the drawing that frees its curvature. Free (green): coil 1's log-derates $g_{1,0}, \ldots, g_{1,36}$ (runs of windows as dashed boxes), each also an entry of its window's graph (Figure 10.1), with that window's load, coil 2's log-derate and the shared β_1; the rate at the center γ_1; the drift rate $\dot{b}_{d,1}$; and (purple) the curvature $\dot{\gamma}_1$. Black squares: the trend rows (10.5) between consecutive windows. Dashed green and purple buses: γ_1 and $\dot{\gamma}_1$ feed every trend row of the coil, the curvature with the lever x_t, written on the links 0 to 1, 17 to 18, 18 to 19 and 35 to 36. The drift node acts on unit 1's discharge readings of windows 20 to 36 only. Pinned, not drawn: each window's chilled-water supply and other heat (known inputs), the ratings and the capacity rates. Observed: each window's five readings, through its window graph. Priors (solid gray): the curvature's and the drift rate's. Coil 2's layer is the same.

The record. The chapter reads the whole record as Chapter 10 does, Table 10.1: window 0, the commissioning test, and windows 1 to 36, January of year 1 to December of year 3,

five readings each. Table 11.1 gives the dashboard's view of it: each row is a quantity for one coil, and the columns are the commissioning test and the three Decembers. The tiles are Chapter 5's raw-performance index $h_{u,t}^{\mathrm{raw}} = A_{u,0}/A_{u,t}$, the commissioning approach over the window's. Coil 1's year-3 drop is the largest of its three years and follows the smallest, and its tile ends below coil 2's; coil 2's drops shrink from year to year.

Table 11.1: The dashboard's tiles at commissioning and each December, the drop of each year (year 1 from commissioning), and the slope of the least-squares line through each year's twelve tiles.

Quantity [unit]	Window 0 (commissioning)	Window 12 (Dec Y1)	Window 24 (Dec Y2)	Window 36 (Dec Y3)
Tile, coil 1, $h_{1,t}^{\mathrm{raw}}$ [–]	1.0000	0.8013	0.7752	0.4873
Tile, coil 2, $h_{2,t}^{\mathrm{raw}}$ [–]	1.0000	0.7879	0.6830	0.5947
Drop over the year, coil 1 [–]	–	−0.1987	−0.0261	−0.2879
Drop over the year, coil 2 [–]	–	−0.2121	−0.1048	−0.0884
Line through the year's tiles, coil 1 [1/yr]	–	−0.1575	−0.0679	−0.2647
Line through the year's tiles, coil 2 [1/yr]	–	−0.0950	−0.1858	−0.0963

The two-unit hall's maintenance log holds, so far, the commissioning test of window 0; Chapter 5's condition assessment and Chapter 6's effort entry at window 18; and Chapter 7's assessment at window 25, which held coil 1's cleaning order and asked for a side-by-side check of unit 1's discharge sensor, a check that has not been made. Chapter 10's entry at window 36 read the rates on the conductance as constants and noted that the path departs from a straight line, in coil 1 through year 3 and in coil 2 around windows 13 to 16, without testing either. This chapter tests the first of those departures and reports the second as the record shows it.

11.1.3 The asset manager's questions (step 1)

The asset manager asks five questions of the record, and step 1 writes each as a probability over the stacked state, with the unknown on the left of the bar and the 185 readings and the priors on the right. Table 11.2 lists them, with the columns of Chapter 5's table: the question in the asset manager's words, the drill question first; the general form of its category from Chapter 2; its mathematical form for the hall; its category; the computation of Section 2.3.6 that answers it; and the subsection where it is answered. Both coils are answered. The first question is the trend form (2.21) one order up, its first line asked of the curvature in place of the rate. Whether either coil's path changed at a single window is Chapter 12's question; how much of the tiles' swing is the season, Chapter 13's; and when the capacity limit is crossed under either drawing, Chapter 14's, in Volume II. Which measurement would decide the acceleration asks nothing new of the record; it is the next measurement, Section 11.6.

11.1.4 The evidence as rows (step 2)

Step 2 attaches the record to the graph as rows of each drawing. Chapter 10's rows are kept as they are: each window's five readings, the return air, the two discharge temperatures and rack 1's inlet at 0.3 K and the load meter at 3.0 kW, with unit 1's discharge row carrying the drift (10.2) from window 20; the rating prior on window 0's coils; the survey prior on β_1; the drift prior; and the trend rows. What changes is the trend rows' expected increment, and the drawings that free a curvature add its prior row. Table 11.3 lists the four drawings: each row is a drawing, the columns the entries it frees beyond the base fit, its numbers of entries and rows, and its degrees of freedom.

Table 11.2: The questions asked of the two-unit hall's condition record at window 36 (Dec Y3). $\mathbf{y}$ is the 185 readings and the prior rows; $\dot{\gamma}_u$ is coil u's curvature and $\gamma_u(t_y)$ its rate at time t_y, equation (11.1); M_0, M_1, M_2 and M_{12} are the four drawings of Section 11.1.1.

Asset manager's question	General form	Mathematical form	Category	Computation	Answered in
Is the loss accelerating?	$\Pr(f(\mathbf{z}) \geq f^* \mid \mathbf{y})$, (2.12), on the curvature; (2.21), first line, one order up	$\Pr(\dot{\gamma}_u < 0 \mid \mathbf{y}, M_{12})$, $u = 1, 2$	prediction under a threshold	trend	11.3.1
Does the record need a changing rate at all, or does the constant rate do?	B_{AB}, (2.14); $\Pr(M_j \mid \mathbf{y})$, (2.47)	$B(M_u : M_0)$; $\Pr(M \mid \mathbf{y})$ over $\{M_0, M_1, M_2, M_{12}\}$; accelerating, constant and slowing as the family's split	hypothesis comparison	evidence	11.3.2
How fast is it losing now, against the record's average?	moments and credible interval, (2.9)–(2.11), over the family	$\gamma_u(3) = \gamma_u + \dot{\gamma}_u\,(3 - \bar{t}_y)$ in %/yr: median and 5–95 % band, averaged over the four drawings	estimation	trend	11.3.3
Is the dashboard's bend in year 3 the coil?	moments of a derived quantity, (2.10)	$\dot{s}_u^{\text{raw}}(\mathbf{z})$, the bend of the quadratic through the tiles $h_{u,t}^{\text{raw}}(\mathbf{z})$, split into load and season, coils, sensor and reading errors	estimation	trend	11.3.4
With the coil allowed to accelerate, does unit 1's sensor still need its drift?	B_{AB}, (2.14)	$B(\text{drift} : \text{none} \mid M_1)$; Pr over the drift alone, the acceleration alone, both	hypothesis comparison	evidence	11.3.5

Table 11.3: The four drawings of the condition record. Each frees the curvature of the coils named, with its prior row; all carry the drift entry.

Drawing	Freed beyond the base fit	Entries [–]	Rows [–]	Degrees of freedom [–]
M_0, constant rates (Chapter 10's base fit)	none	115	261	146
M_1, coil 1's rate may change	$\dot{\gamma}_1$	116	262	146
M_2, coil 2's rate may change	$\dot{\gamma}_2$	116	262	146
M_{12}, both rates may change	$\dot{\gamma}_1$, $\dot{\gamma}_2$	117	263	146

The stacked state. The state of M_{12} is Chapter 10's stacked state (10.3) with the two curvatures appended,

$$\mathbf{z} = (Q_0,\ g_{1,0},\ g_{2,0},\ \ldots,\ Q_{36},\ g_{1,36},\ g_{2,36},\ \beta_1,\ \gamma_1,\ \gamma_2,\ \dot{b}_{d,1},\ \dot{\gamma}_1,\ \dot{\gamma}_2)^{\mathsf{T}}, \tag{11.3}$$

117 entries: the loads in kW, the drift rate in K/yr, the rates in yr^{-1}, the curvatures in yr^{-2}, and the rest dimensionless. M_1 and M_2 drop one curvature and M_0 both.

The rows. Each trend row of Chapter 10, equation (10.5), takes the increment of equation (11.2) as its expectation:

$$r_{q,u,t}(\mathbf{z}) = \frac{g_{u,t+1} - g_{u,t} - \Delta t\,[\gamma_u + \dot{\gamma}_u\, x_t]}{\sigma_q}, \qquad t = 0, \ldots, 35, \quad u = 1, 2, \tag{11.4}$$

with $\sigma_q = 0.005$ per window, Chapter 10's, and the curvature's term present for the coils whose curvature the drawing frees. Each freed curvature adds the prior row $\dot{\gamma}_u/\sigma_{\dot{\gamma}}$, $\sigma_{\dot{\gamma}} = 0.05\,\mathrm{yr}^{-2}$. M_{12} has 263 rows on 117 entries, 146 degrees of freedom, as M_0 has: each curvature adds one entry and one row. The levers run from $x_0 = -1.4583\,\mathrm{yr}$ for the increment from window 0 to window 1 to $x_{35} = +1.4583\,\mathrm{yr}$ for the increment into window 36, and over the 36 increments $\sum_t x_t^2 = 26.9792\,\mathrm{yr}^2$.

Which row acts on which entry. The Jacobian of M_{12} is 263×117: Chapter 10's 261×115 Jacobian with its block pattern (10.6), bordered by the two curvature columns and their two prior rows. With the rows grouped as the readings, the rating, survey and drift priors, the 72 trend rows and the two curvature priors, and the columns as Chapter 10's 115 and then $(\dot{\gamma}_1, \dot{\gamma}_2)$,

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_{\text{read}} & \mathbf{0} \\ \mathbf{J}_{\text{prior}} & \mathbf{0} \\ \mathbf{J}_{\text{trend}} & \mathbf{F} \\ \mathbf{0} & \sigma_{\dot{\gamma}}^{-1}\,\mathbf{I}_2 \end{bmatrix}, \qquad \mathbf{J}_{\text{trend}} = \begin{bmatrix} \mathbf{D}_0 & \cdots & \mathbf{D}_{36} & \mathbf{0} & \mathbf{E} & \mathbf{0} \end{bmatrix}, \qquad \mathbf{F} = -\frac{\Delta t}{\sigma_q} \begin{bmatrix} \mathbf{x}_{\text{lev}} & \mathbf{0} \\ \mathbf{0} & \mathbf{x}_{\text{lev}} \end{bmatrix},$$

$$\mathbf{g}_{\dot{\gamma}_u} = \mathbf{e}_{\dot{\gamma}_u}, \qquad \mathbf{g}_{\gamma_u(3)} = \mathbf{e}_{\gamma_u} + 1.5\,\mathbf{e}_{\dot{\gamma}_u}, \qquad \mathbf{g}_{\hat{s}_u^{\text{raw}}} \text{ dense over the 37 windows' blocks.} \tag{11.5}$$

$\mathbf{J}_{\text{read}}$ holds the reading blocks $\mathbf{R}_t$ of equation (10.6), each window's five rows on its own three entries, with rack 1's rows in the column of β_1 and unit 1's discharge rows of windows 20 to 36 in the column of $\dot{b}_{d,1}$; $\mathbf{J}_{\text{prior}}$ holds the rating priors' block $\mathbf{P}_0$ and the survey and drift priors. In the trend rows, $\mathbf{D}_t$ is window t's block, $\pm 1/\sigma_q = \pm 200$ in the columns of $g_{u,t}$, and $\mathbf{E}$ the block in (γ_1, γ_2), $-\Delta t/\sigma_q$ in each coil's 36 rows, both as in Chapter 10. The new block $\mathbf{F}$ is $\mathbf{E}$ with each row weighted by its lever: $\mathbf{x}_{\text{lev}} = (x_0, \ldots, x_{35})^{\mathsf{T}}$. The second line gives the questions' gradients: the curvature, the rate today, and the dashboard's bend, the last a dense row over the 37 windows' entries and the drift rate. Of the 30 771 entries of $\mathbf{J}$, 829 can be nonzero, 2.69 %. The column of $\dot{\gamma}_u$ holds 37: coil u's 36 trend rows, each $-\Delta t\, x_t/\sigma_q$, from $+24.3056$ at the increment out of window 0 through ± 0.6944 at the center to -24.3056 at the increment into window 36, and its prior row, $1/\sigma_{\dot{\gamma}} = 20.0000$.

Two facts are read here, before any number of the record. First, the curvature is the slope of the path's monthly rates. Write $\dot{g}_{u,t} = (g_{u,t+1} - g_{u,t})/\Delta t$ for the path's monthly rate. The gradient $\mathbf{J}^{\mathsf{T}}\mathbf{r}$ vanishes at the MAP state in every column; in the columns of γ_u and $\dot{\gamma}_u$, with the lines of the pattern (11.5) and the levers symmetric about the center, $\sum_t x_t = 0$,

$$\begin{aligned} 0 &= -\frac{\Delta t}{\sigma_q} \sum_t r_{q,u,t} && \Longrightarrow \quad \hat{\gamma}_u = \frac{1}{36} \sum_t \hat{\dot{g}}_{u,t}, \\ 0 &= -\frac{\Delta t}{\sigma_q} \sum_t x_t\, r_{q,u,t} + \frac{\hat{\dot{\gamma}}_u}{\sigma_{\dot{\gamma}}^2} && \Longrightarrow \quad \hat{\dot{\gamma}}_u = \frac{\sum_t x_t\, \hat{\dot{g}}_{u,t}}{\sum_t x_t^2 + \sigma_q^2/(\Delta t^2\, \sigma_{\dot{\gamma}}^2)}, \end{aligned} \tag{11.6}$$

with $\hat{g}_{u,t}$ the fitted path's monthly rate. The rate at the center is the mean of the fitted path's monthly rates, Chapter 10's telescoped sum, and the curvature is their least-squares slope against the lever, shrunk by the prior by the factor $26.9792/(26.9792 + 1.4400) = 0.94933$, with $\sigma_q^2/(\Delta t^2\,\sigma_{\dot\gamma}^2) = 1.4400\,\mathrm{yr}^2$ the prior's term. Second, the curvature has a wander floor. With the path known exactly at every window, the only rows on $\dot\gamma_u$ are the trend rows, whose column is $-\Delta t\,\mathbf{x}_{\mathrm{lev}}/\sigma_q$, and its information is $\Delta t^2\,\mathbf{x}_{\mathrm{lev}}^{\mathsf{T}}\mathbf{x}_{\mathrm{lev}}/\sigma_q^2$; without the prior row,

$$\mathrm{sd}(\dot\gamma_u) \geq \frac{\sigma_q}{\Delta t\,\sqrt{\sum_t x_t^2}} = 0.01155\,\mathrm{yr}^{-2}, \tag{11.7}$$

which the prior row brings to 0.01126. No measurement on the existing windows reads the curvature more finely than this, and what a longer record adds is the lever arm $\sum_t x_t^2$: each new increment adds the square of its distance from the center, so a later window adds more than an earlier one (Exercise 11.2).

Modeling choices at the rows. Five choices make these rows. The first is that the rate changes linearly over the whole record, one curvature per coil, a quadratic path with wander. This buys one entry per coil, a sign and an evidence, with no knee to choose. It fails if the rate is constant and then changes at a knee: a knee late in the record is read as a smaller curvature spread over all of it. Section 11.4.3 sets knees at the starts of year 2 and year 3 beside it; ranking windows by evidence is Chapter 12's method, and the knee is not swept here. The second is the center at window 18, so that γ_u is the rate at the record's center and nearly uncorrelated with the curvature. With γ_u free and flat, the centering moves neither the evidence nor the curvature's posterior (Exercise 11.3); it moves only what γ_u means. The third is the curvature's prior, $\mathcal{N}(0, 0.05^2)$ in yr^{-2}: at one standard deviation the rate may change within a year by about its own size, Chapter 10's 0.04 to $0.06\,\mathrm{yr}^{-1}$. It is proper because the Bayes factor needs it, Chapter 7's Occam factor, and its spread is varied from 0.01 to 0.5 in step 4. The fourth is Chapter 10's base fit unchanged: $\sigma_q = 0.005$, the drift linear from window 20 with Chapter 7's prior, and no step on either path. The base fit is the constant-rate drawing M_0, and the choices Chapter 10 named for it hold here. The fifth is even prior odds over the four drawings, as in Chapter 7's family.

11.2 The posterior of the state (step 3)

All five questions of Table 11.2 are read from the posteriors of the four drawings under the rows of step 2: the answers within a drawing from that of M_{12}, and the evidence from all four.

11.2.1 The rows and the normal equations

The derivation form is Chapter 5's, on the stacked state of Chapter 10: the reading rows of each window on the reduced state at that window's inputs, Section 5.2.1; the trend rows (11.4); and the prior rows, stacked into one residual vector $\mathbf{r}(\mathbf{z})$ of 263 normalized rows, with the objective $\mathcal{L}(\mathbf{z}) = \frac{1}{2}\,\mathbf{r}^{\mathsf{T}}\mathbf{r}$. Each Gauss–Newton iteration solves the normal equations (5.9), $(\mathbf{J}^{\mathsf{T}}\mathbf{J})\,\Delta\mathbf{z} = -\mathbf{J}^{\mathsf{T}}\mathbf{r}$, and at the minimum the Laplace covariance is $\mathbf{\Sigma} = (\mathbf{J}^{\mathsf{T}}\mathbf{J})^{-1}$, equation (5.10). With every row normalized the information matrix is $\mathbf{\Lambda} = \mathbf{J}^{\mathsf{T}}\mathbf{J}$; with unnormalized rows and their weights it is $\mathbf{J}^{\mathsf{T}}\mathbf{W}\mathbf{J}$. The trend rows, the curvature's terms among them, are linear in the state, so the curvature adds nothing to the second-order term $\mathbf{C}$ of the objective, equation (4.44); the nonlinear rows are the readings', as in Chapter 10.

The Jacobian at the MAP state. Eight lines of $\mathbf{J}$, in the columns $(g_{1,24}, g_{1,25}, \gamma_1, \dot{\gamma}_1, \dot{\gamma}_2, \dot{b}_{d,1})$, show the new column against the old:

$$\begin{bmatrix} 0 & 0 & -16.6667 & 24.3056 & 0 & 0 \\ 0 & 0 & -16.6667 & 0.6944 & 0 & 0 \\ 0 & 0 & -16.6667 & -0.6944 & 0 & 0 \\ -200 & 200 & -16.6667 & -9.0278 & 0 & 0 \\ 0 & 0 & -16.6667 & -24.3056 & 0 & 0 \\ 0 & 0 & 0 & 0 & -24.3056 & 0 \\ 0 & 0 & 0 & 20.0000 & 0 & 0 \\ 0 & 0 & 0 & 0 & 20.0000 & 0 \end{bmatrix} \begin{array}{r} \text{trend row, coil 1, window 0 to 1} \\ \text{trend row, coil 1, window 17 to 18} \\ \text{trend row, coil 1, window 18 to 19} \\ \text{trend row, coil 1, window 24 to 25} \\ \text{trend row, coil 1, window 35 to 36} \\ \text{trend row, coil 2, window 35 to 36} \\ \text{curvature prior, coil 1} \\ \text{curvature prior, coil 2} \end{array} \tag{11.8}$$

the rate columns in yr, the curvature columns in yr^2 and the drift column in yr/K. Every trend row of coil 1 sees γ_1 alike, $-\Delta t/\sigma_q = -16.6667$; its slope on $\dot{\gamma}_1$ changes sign at the center and grows with the distance from it, the lever x_t times $-\Delta t/\sigma_q$. The rows on either side of the center nearly cancel each other's pull on the curvature, and the first and last months read it most strongly. The drift column is zero in every trend row: the drift and the curvature meet only through the readings.

11.2.2 The fit of the accelerating drawing

The iterations. Table 11.4 runs the normal equations of M_{12} from the state the record suggests: the meter's loads, the coils at their ratings, the survey's fraction, and no trend, drift or curvature. Each row is an iteration, with the objective and the norm of its gradient. The objective falls from 1250.311469 to 81.628688 in the first step and settles at 76.165116 by the fourth; the gradient reaches the floor that the Jacobian's round-off sets, about 1.4×10^{-7}, by the ninth.

Table 11.4: Gauss–Newton iterations of the drawing M_{12} on the 117-entry stacked state, from the meter's loads, the rated coils, the survey's fraction and no trend, drift or curvature.

Iteration l [–]	Objective $\mathcal{L}(\mathbf{z}^{(l)})$ [–]	$\lVert\mathbf{J}^{\mathsf{T}}\mathbf{r}\rVert$ [–]
0	1250.311469	3.649×10^{3}
1	81.628688	2.430×10^{2}
2	76.165917	2.075
3	76.165117	4.269×10^{-2}
4	76.165116	1.324×10^{-3}
5	76.165116	2.875×10^{-5}
6	76.165116	1.476×10^{-6}
7	76.165116	1.897×10^{-7}
8	76.165116	1.736×10^{-7}
9	76.165116	1.368×10^{-7}
10	76.165116	1.413×10^{-7}

The residuals. At the MAP state $2\mathcal{L}(\hat{\mathbf{z}}) = 152.3302$ on 146 degrees of freedom, and $\Pr(\chi^2_{146} > 152.3) = 0.343$, against Chapter 10's 154.9755 and 0.290 for M_0. The readings carry 147.5858 of it and the trend rows 4.0626 (150.1030 and 4.7022 in M_0). By coil, the trend rows' squared residuals fall from 2.0869 to 1.6134 for coil 1 and from 2.6153 to 2.4492 for coil 2: the curvature takes up coil 1's year-3 months, which the constant rate left in the trend residuals, and little of coil 2's.

The MAP state. Table 11.5 gives the entries the answers use, with their posterior spreads, in M_{12} and, beside them, in Chapter 10's base fit M_0. Each row is an entry. The curvature of coil 1 is negative and that of coil 2 positive, each with a spread larger than half its value. Freeing them moves coil 1's commissioned and window-36 entries down, its rate at the center from -0.04240 to $-0.04707\,\mathrm{yr}^{-1}$ and the drift rate from 0.6837 to 0.5642 K/yr; coil 2's window-36 entry moves up. Every spread of a log-derate widens, the window-36 entries most.

Table 11.5: The MAP state of the drawing M_{12} and of the base fit M_0, with the posterior spreads, for the entries the answers use.

Entry [unit]	M_{12}, $\hat{z}_i$	$\sqrt{\Sigma_{ii}}$	M_0, $\hat{z}_i$	$\sqrt{\Sigma_{ii}}$
$g_{1,0}$ [–]	−0.02206	0.01851	−0.00594	0.01543
$g_{2,0}$ [–]	−0.01589	0.01860	−0.02448	0.01558
$g_{1,36}$ [–]	−0.16328	0.03073	−0.13314	0.02386
$g_{2,36}$ [–]	−0.19602	0.02090	−0.21208	0.01690
β_1 [–]	0.21060	0.00903	0.21272	0.00893
γ_1 [yr^{-1}]	−0.04707	0.01500	−0.04240	0.01463
γ_2 [yr^{-1}]	−0.06004	0.01327	−0.06253	0.01323
$\dot{\gamma}_1$ [yr^{-2}]	−0.03341	0.02118	–	–
$\dot{\gamma}_2$ [yr^{-2}]	+0.01623	0.02009	–	–
$\dot{b}_{d,1}$ [K/yr]	0.5642	0.1838	0.6837	0.1681

The covariance. The Laplace covariance of these ten entries, in the order of Table 11.5, scaled by 10^4 and in the entries' units,

$$10^4\,\boldsymbol{\Sigma} = \begin{bmatrix} 3.4272 & -1.3171 & 0.80369 & -0.34436 & 0.16074 & -0.87452 & 0.32424 & 2.1267 & -0.83905 & 1.8332 \\ -1.3171 & 3.4614 & -0.073664 & 0.49602 & 0.14788 & 0.41447 & -0.98845 & -0.74815 & 2.0627 & 1.1180 \\ 0.80369 & -0.073664 & 9.4417 & -4.1533 & 0.94740 & 2.8793 & -1.3599 & 3.9984 & -1.7482 & 47.037 \\ -0.34436 & 0.49602 & -4.1533 & 4.3701 & -0.25772 & -1.2697 & 1.2914 & -1.7448 & 2.3426 & -20.814 \\ 0.16074 & 0.14788 & 0.94740 & -0.25772 & 0.81479 & 0.26222 & -0.13520 & 0.27928 & -0.12609 & 6.0018 \\ -0.87452 & 0.41447 & 2.8793 & -1.2697 & 0.26222 & 2.2513 & -0.56138 & 0.62387 & -0.30306 & 15.068 \\ 0.32424 & -0.98845 & -1.3599 & 1.2914 & -0.13520 & -0.56138 & 1.7599 & -0.33222 & 0.093289 & -7.3107 \\ 2.1267 & -0.74815 & 3.9984 & -1.7448 & 0.27928 & 0.62387 & -0.33222 & 4.4849 & -1.3538 & 15.755 \\ -0.83905 & 2.0627 & -1.7482 & 2.3426 & -0.12609 & -0.30306 & 0.093289 & -1.3538 & 4.0350 & -7.0206 \\ 1.8332 & 1.1180 & 47.037 & -20.814 & 6.0018 & 15.068 & -7.3107 & 15.755 & -7.0206 & 337.75 \end{bmatrix},$$

is read by its correlations, and the rows fix each sign.

- $\dot{\gamma}_1$ against $\dot{b}_{d,1}$, +0.4048: a faster drift explains more of unit 1's late rise in its discharge reading, so coil 1 needs less acceleration. This is Chapter 7's ridge between the coil and its sensor, now on the curvature; the drift rate against $g_{1,36}$ is +0.8329.
- $\dot{\gamma}_1$ against $g_{1,0}$, +0.5425, and against $g_{1,36}$, +0.6144: with the middle windows held, a higher commissioned value needs a steeper early fall and a higher value today a gentler late one, and both mean less acceleration.
- $\dot{\gamma}_1$ against $\dot{\gamma}_2$, −0.3183, and against $g_{2,36}$, −0.3941: the return air sees the total conductance, so a coil 1 that accelerates leaves less of the late loss to coil 2.
- $\dot{\gamma}_1$ against γ_1, +0.1963: small, by the centering of step 2.

The curvature as the path's reading. Under M_{12} coil 1's fitted path changes by −0.01825, −0.03661 and −0.08636 in years 1, 2 and 3, against −0.03873, −0.03172 and −0.05675 under M_0. By equation (11.6), the least-squares slope of the fitted path's monthly rates against the lever is $-0.03519\,\mathrm{yr}^{-2}$, and times the shrink factor 0.94933 it is −0.03341, the MAP curvature; the mean monthly rate is $-0.04707\,\mathrm{yr}^{-1}$, the MAP rate at the center. The increments that carry most of $\hat{\dot{\gamma}}_1$ are those into windows 36, 34 and 35 (−0.00544, −0.00504 and −0.00461). Coil 2's curvature is positive, carried by its early increments, out of windows 0, 3 and 6 (+0.00379, +0.00294 and +0.00267), and by the increments of windows 13 to 16, which together carry

+0.00304, 18.8 % of it. Whether those months are a slower rate or one event in coil 2's path is Chapter 12's question.

11.2.3 The evidence of the drawings and the full posterior

One choice enters here, as in Chapter 7: the evidence of each drawing is its Laplace evidence (7.7), the Gaussian integral of the objective about the drawing's own MAP state. Between two nested drawings it is equation (7.8), the gain in fit less the Occam factor, here with $\sigma^2 = \sigma_{\dot\gamma}^2$ for one freed curvature and $\sigma_{\dot\gamma}^4$ for two. The units cancel: $\sigma_{\dot\gamma}^2$ is in yr^{-4} and the determinant ratio in yr^4, because each curvature column of $\mathbf{J}$ is in yr^2. The posterior probability of each drawing is equation (7.9) over $\{M_0, M_1, M_2, M_{12}\}$ at even prior odds, and the Savage–Dickey ratio (7.21) gives each nested Bayes factor from the larger drawing's fit alone.

A quantity asked over the family has the mixture posterior (7.10), a drawing that pins the curvature contributing a point mass at zero, with the mean and the variance of equation (7.11). For coil u the family splits the probability into three: the loss accelerates, is constant, or slows,

$$\begin{aligned} \Pr(\text{accelerating}_u \mid \mathbf{y}) &= \sum_{M \ni \dot\gamma_u} \Pr(M \mid \mathbf{y})\, \Pr(\dot\gamma_u < 0 \mid \mathbf{y}, M), \\ \Pr(\text{constant}_u \mid \mathbf{y}) &= \sum_{M \not\ni \dot\gamma_u} \Pr(M \mid \mathbf{y}), \\ \Pr(\text{slowing}_u \mid \mathbf{y}) &= 1 - \Pr(\text{accelerating}_u \mid \mathbf{y}) - \Pr(\text{constant}_u \mid \mathbf{y}), \end{aligned} \tag{11.9}$$

the first sum over the drawings that free coil u's curvature and the second over those that pin it.

The full posterior of each drawing is computed by the reweighted draws of Section 5.2.4, with the drawing's Laplace normal as the proposal and the weights of equation (5.13): $N_\mathrm{d} = 100\,000$ draws per drawing from seeded generators (seeds 10 for M_0, Chapter 10's, 12 for M_1, 13 for M_2 and 11 for M_{12}), with Monte Carlo standard errors from 20 batches. The same draws estimate each evidence by equation (7.12). For M_{12} the effective sample size is 98 529, 98.5 % of N_d, and the largest weight 7.54×10^{-5}, in 117 dimensions; the other drawings' effective sample sizes are 98.7 to 98.8 %.

11.3 The answers (step 4)

Every answer is read from the fits of Section 11.2.2 and the evidences of Section 11.2.3, from the drawing each answer names, and that value is the one reported. A derived quantity f is carried by its gradient $\mathbf{g}$ at $\hat{\mathbf{z}}$, with variance $\mathbf{g}^\mathsf{T}\boldsymbol{\Sigma}\mathbf{g}$. The full posterior by reweighted draws stands beside each value, a median beside a median.

11.3.1 Whether the loss accelerates

Asset manager's question. Is the loss accelerating? It is asked of each coil, within the drawing that lets both rates change.

General form. A prediction under a threshold is a tail of a derived quantity under the posterior, equation (2.12),

$$\Pr(f(\mathbf{z}) \geq f^* \mid \mathbf{y}) = \int \mathbf{1}\big[f(\mathbf{z}) \geq f^*\big]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z},$$

and asked of a trend entry it is the first line of the trend form (2.21), here one order up: the sign of the change of the rate, not of the rate.

Specialized form. The question is equation (2.12) with $f(\mathbf{z}) = -\dot{\gamma}_u$, the event $f > 0$, an accelerating loss, and $p(\mathbf{z} \mid \mathbf{y})$ the Laplace normal of M_{12}, Section 11.2.2. One choice enters here: the probability is read within M_{12}, the drawing that asks the question of both coils. This is what makes the answer a tail of one entry of one posterior, and it is conditional on the rate being allowed to change: whether it should be is the next question. The curvature is an entry, with gradient $\mathbf{g}_{\dot{\gamma}_u} = \mathbf{e}_{\dot{\gamma}_u}$ from the pattern (11.5), so its variance is $\Sigma_{\dot{\gamma}_u\dot{\gamma}_u}$ and

$$\Pr(\dot{\gamma}_u < 0 \mid \mathbf{y}, M_{12}) = \Phi\left(\frac{-\hat{\dot{\gamma}}_u}{\sqrt{\Sigma_{\dot{\gamma}_u\dot{\gamma}_u}}}\right). \tag{11.10}$$

Building blocks. The MAP state and the covariance of Section 11.2.2; the gradient $\mathbf{e}_{\dot{\gamma}_u}$; the normal distribution function Φ.

The sign of each curvature. For coil 1, $\hat{\dot{\gamma}}_1 = -0.03341\,\mathrm{yr}^{-2}$ with standard deviation 0.02118, a 5–95 % band $[-0.06825,\ +0.00142]$, or -3.341 percentage points of rate per year with band $[-6.825,\ +0.142]$; by equation (11.10),

$$\Pr(\dot{\gamma}_1 < 0 \mid \mathbf{y}, M_{12}) = \Phi(1.5777) = 0.943.$$

The full posterior gives a median of -0.03296, a band of $[-0.06800,\ +0.00129]$ and $\Pr = 0.9437$. For coil 2, $\hat{\dot{\gamma}}_2 = +0.01623$ with standard deviation 0.02009, band $[-0.01681,\ +0.04927]$, and $\Pr(\dot{\gamma}_2 < 0 \mid \mathbf{y}, M_{12}) = \Phi(-0.8079) = 0.210$: a slowing rate with probability 0.7904. The full posterior gives $+0.01577$, $[-0.01711,\ +0.04873]$ and 0.2135. With one coil's curvature freed alone, M_1 reads coil 1's as -0.02806 (standard deviation 0.02002, $\Pr = 0.919$) and M_2 coil 2's as $+0.00626$ (0.01912, $\Pr = 0.372$). Freeing coil 2's curvature beside coil 1's moves coil 1's by -0.00535, through the correlation -0.3183 of the two curvatures.

The two posteriors are drawn in Figure 11.2. The horizontal axis is the curvature as $100\dot{\gamma}_u$, in percentage points of rate per year, the vertical axis its posterior density; the shaded areas lie below zero, the event of the question, and the solid gray line marks a constant rate. The dotted lines are the drawing's pseudo-true values, check markers for Section 11.4.6. The dashboard's bends, a different quantity in the commissioned tile per yr^2, are not drawn on this axis; the note in the upper right gives them, and Section 11.3.4 takes them apart. Coil 1's density lies mostly below zero, coil 2's mostly above it, and both are wide against the distance between them.

11.3.2 Whether the record needs a changing rate

Asset manager's question. Does the record need a changing rate at all, or does the constant rate do? The answer is a Bayes factor for each drawing against the constant rate, the probabilities of the four drawings, and for each coil the split of its loss into accelerating, constant and slowing.

General form. A hypothesis comparison is the ratio of the evidences of two complete drawings fitted to the same record, equation (2.14),

$$B_{AB} = \frac{p(\mathbf{y} \mid M_A)}{p(\mathbf{y} \mid M_B)},$$

and over a family of drawings each has the posterior probability of equation (2.47),

$$\Pr(M_j \mid \mathbf{y}) = \frac{p(\mathbf{y} \mid M_j)\ \Pr(M_j)}{\sum_k p(\mathbf{y} \mid M_k)\ \Pr(M_k)}.$$

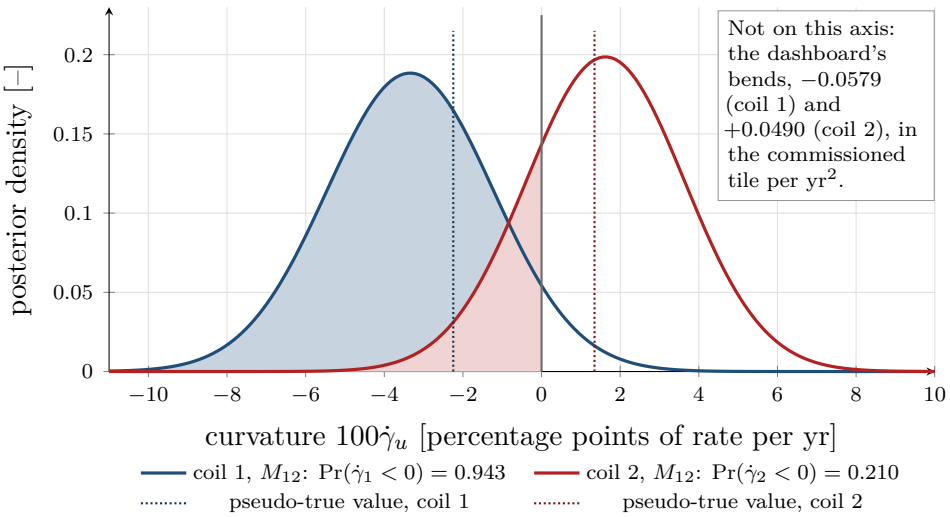


Figure 11.2: The posterior densities of the two curvatures in the drawing M_{12}, normals with means −3.341 and +1.623 and standard deviations 2.118 and 2.009, in percentage points of rate per year. Shaded: the probability of an accelerating loss. Solid gray: a constant rate. Dotted: the pseudo-true values −2.245 and +1.347, a check marker (Section 11.4.6). Boxed note: the dashboard's bends, a different quantity, given as numbers (Section 11.3.4).

Specialized form. The question is equation (2.14) with $M_A = M_u$ or M_{12} and $M_B = M_0$, and equation (2.47) over the four drawings with $\Pr(M_j) = \frac{1}{4}$, which is equation (7.9). Each evidence is the Laplace evidence (7.7) of the drawing's own fit, and since M_0 is nested in each of the others, equation (7.8) gives each log Bayes factor as the gain in fit, $\mathcal{L}_0(\hat{\mathbf{z}}_0) - \mathcal{L}_M(\hat{\mathbf{z}}_M)$, less the Occam factor, $\frac{1}{2}\ln(\sigma_{\dot{\gamma}}^2 \det \mathbf{\Lambda}_M / \det \mathbf{\Lambda}_0)$ for one freed curvature. The family's split for each coil is equation (11.9), with the signs of equation (11.10) read in M_u and M_{12}. Beside it, the Savage–Dickey ratio (7.21) on the Laplace marginal of M_1, $\ln B(M_1 : M_0) = \ln \mathcal{N}(0;\ 0,\ \sigma_{\dot{\gamma}}^2) - \ln \mathcal{N}(0;\ \hat{\dot{\gamma}}_1,\ \Sigma_{\dot{\gamma}_1\dot{\gamma}_1})$, needs only the larger drawing's fit. One choice enters here: the prior spread $\sigma_{\dot{\gamma}}$. It enters the Occam factor directly and the gain through the prior row, and the two answers respond to it differently: as the spread widens the Bayes factor falls without limit, while the sign settles, Lindley's effect. This is what lets the evidence charge the curvature for the range of rates it allows, and it means the verdict is only as firm as its range over plausible spreads, which the end of this subsection gives.

Building blocks. Each drawing's objective at its MAP state and $\ln\det \mathbf{\Lambda}_M$ (Section 11.2.2); the prior spread $\sigma_{\dot{\gamma}}$; the prior probabilities; the signs within each drawing (Section 11.3.1).

The evidences. Table 11.6 evaluates equations (7.7) to (7.9). Each row is a drawing; the columns give twice its objective, $\ln\det \mathbf{\Lambda}_M$, its log evidence, its log Bayes factor against M_0 with that factor's two parts, the gain in fit and the Occam factor, and its posterior probability.

The record does not decide whether coil 1's rate changes:

$$B(M_1 : M_0) = 1.056.$$

Coil 1's curvature buys 0.9942 in fit and costs 0.9392 in Occam factor. The Savage–Dickey ratio gives $\ln B(M_1 : M_0) = 2.0768 - 2.0101 = +0.0667$, and importance sampling $\ln B(M_1 : M_0) = +0.0583$ (standard error 0.0005) beside the Laplace +0.0549, a factor of 1.060 beside 1.056. Coil 2's curvature buys almost nothing, 0.0539, and the record leans against it, $B(M_2 : M_0) = 0.403$; freeing both gives $B(M_{12} : M_0) = 0.585$. Over the family, by equation (11.9),

Table 11.6: The evidence of each drawing by equation (7.7), up to the constant of the flat priors common to all four, and the bookkeeping of equation (7.8) against M_0. The determinant $\det \mathbf{\Lambda}_M$ is evaluated with the entries in the units of Table 11.5 and the loads in kW, so its logarithm depends on those units; they cancel in every ratio of two drawings, and the other columns are dimensionless.

Drawing	$2\mathcal{L}(\hat{\mathbf{z}})$ [–]	$\ln\det\mathbf{\Lambda}_M$ [–]	$\ln p(\mathbf{y} \mid M)$ [–]	$\ln B$ vs M_0 [–]	gain [–]	Occam [–]	$\Pr(M \mid \mathbf{y})$ [–]
M_0, constant	154.9755	763.8513	−68.9335	0	–	–	0.3284
M_1, coil 1	152.9872	771.7213	−68.8785	+0.0549	0.9942	0.9392	0.3470
M_2, coil 2	154.8676	771.7687	−69.8425	−0.9090	0.0539	0.9629	0.1323
M_{12}, both	152.3302	779.5504	−69.4689	−0.5354	1.3226	1.8581	0.1923

coil 1's curvature is in the drawing with probability 0.5392, and coil 1's loss is

accelerating 0.500, constant 0.461, slowing 0.039;

coil 2's is accelerating with probability 0.090, constant 0.675 and slowing 0.235. The step's finding is the gap between the two answers: within M_{12} the record is fairly sure of the sign of coil 1's curvature, 0.943, and between the drawings it is even, 1.056. The first is conditional on a changing rate; the second asks whether to allow one, and three years of record cannot pay for the entry.

Lindley's effect. Table 11.7 repeats the evidence and the sign at other prior spreads of the curvature; each column is a spread, and the rows give the two log Bayes factors and the sign probability of coil 1 in M_{12}. Figure 11.3 plots them against the spread on a logarithmic axis: the upper panel the two log Bayes factors, with the band $|\ln B| \le 1$ shaded and the reported spread and the most favorable one marked, the lower panel the sign probability. $\ln B(M_1 : M_0)$ peaks near $0.024\,\mathrm{yr}^{-2}$ and falls steadily beyond it; $\ln B(M_{12} : M_0)$ is largest at the smallest spread and falls from there; and the sign probability rises and flattens.

Table 11.7: The evidence and the sign at six prior spreads of the curvature. The chapter's spread is $0.05\,\mathrm{yr}^{-2}$.

$\sigma_{\dot\gamma}$ [yr^{-2}]	0.01	0.02	0.05	0.1	0.2	0.5
$\ln B(M_1 : M_0)$ [–]	+0.1068	+0.2247	+0.0549	−0.4425	−1.0795	−1.9795
$\ln B(M_{12} : M_0)$ [–]	+0.0234	+0.0024	−0.5354	−1.5427	−2.8153	−4.6143
$\Pr(\dot\gamma_1 < 0 \mid \mathbf{y}, M_{12})$ [–]	0.7440	0.8639	0.9427	0.9570	0.9606	0.9617

No prior spread lets the record decide in favor of a changing rate. The most favorable spread for M_1, over a grid from 0.002 to $1.0\,\mathrm{yr}^{-2}$, is 0.0240, and there $\ln B(M_1 : M_0) = +0.2380$, a factor of 1.269. A wide prior charges the curvature for rates the record never shows and turns the verdict against it, −1.0795 at 0.2 and −1.9795 at 0.5, while from 0.05 up the sign probability settles between 0.9427 and 0.9617. Only the evidence depends on the prior without limit.

11.3.3 The rate today

Asset manager's question. How fast is it losing now, against the record's average? The answer is each coil's rate at window 36 in %/yr, with its band, averaged over the four drawings, and the rate in each drawing beside it.

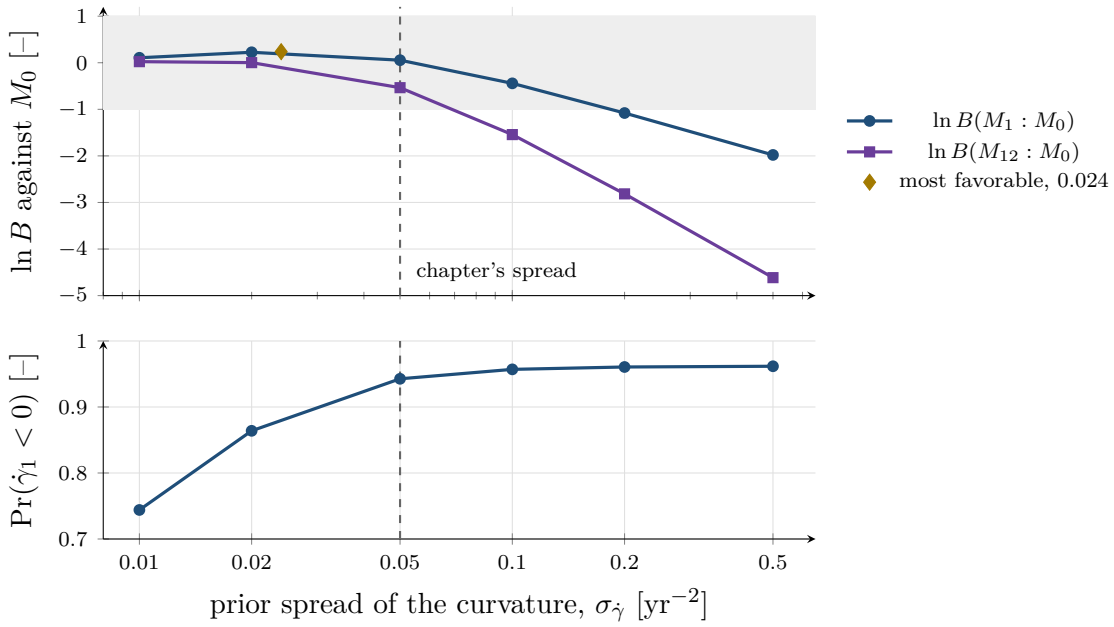


Figure 11.3: Lindley's effect on the curvature's evidence. Upper panel: the log Bayes factors of M_1 and M_{12} against the constant-rate drawing at the six prior spreads of Table 11.7, with the band $|\ln B| \leq 1$ shaded and the most favorable spread for M_1 (diamond). Lower panel: the sign probability of coil 1's curvature in M_{12}. Dashed: the chapter's spread, $0.05\,\mathrm{yr}^{-2}$.

General form. An estimation question asks for the posterior distribution of a quantity, equation (2.9),

$$\Pr(f(\mathbf{z}) \leq f^* \mid \mathbf{y}) = \int \mathbf{1}[f(\mathbf{z}) \leq f^*]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z},$$

summarized by its median and the 5–95 % credible interval of equation (2.11); asked over a family of drawings, $p(\mathbf{z} \mid \mathbf{y})$ is the family's mixture.

Specialized form. The question is equations (2.9) and (2.11) with $f = \gamma_u(t_\mathrm{y})$ of equation (11.1) at $t_\mathrm{y} = 3\,\mathrm{yr}$, and the posterior the mixture (7.10) of the four drawings with the probabilities of Table 11.6. In a drawing that frees coil u's curvature the rate is linear in the state, with the gradient of the pattern (11.5) and its variance by the delta method (4.53),

$$\begin{aligned} \mathbf{g}_{\gamma_u(t_\mathrm{y})} &= \mathbf{e}_{\gamma_u} + (t_\mathrm{y} - \bar{t}_\mathrm{y})\, \mathbf{e}_{\dot{\gamma}_u}, \\ \mathrm{Var}(\gamma_u(t_\mathrm{y}) \mid \mathbf{y}, M) = \mathbf{g}_{\gamma_u(t_\mathrm{y})}^\mathsf{T} \boldsymbol{\Sigma}\, \mathbf{g}_{\gamma_u(t_\mathrm{y})} &= \Sigma_{\gamma_u\gamma_u} + (t_\mathrm{y} - \bar{t}_\mathrm{y})^2\, \Sigma_{\dot{\gamma}_u\dot{\gamma}_u} + 2(t_\mathrm{y} - \bar{t}_\mathrm{y})\, \Sigma_{\gamma_u\dot{\gamma}_u}, \end{aligned} \tag{11.11}$$

which the delta method gives with no linearization error, the rate being linear in the state, so that the drawing's Laplace normal carries it without further approximation; in a drawing that pins the curvature, $\gamma_u(t_\mathrm{y}) = \gamma_u$ at every t_y. The mixture's quantile f_q solves $\sum_j \Pr(M_j \mid \mathbf{y})\, \Phi((f_q - \bar{f}_j)/\varsigma_j) = q$, with $\bar{f}_j$ and ς_j the rate's mean and standard deviation in drawing j, and its mean and variance are equation (7.11). The quantiles pass to %/yr through the increasing map of equation (10.1). One choice enters here: the reported value is the family's average, not the rate of one drawing. This is what carries the record's indecision over the form into the answer, and it rests on the even prior odds of step 2; the rates of the drawings stand beside it.

Building blocks. The fits of the four drawings, their probabilities (Table 11.6), the variance (11.11), the mixture's quantiles, and the map to %/yr.

The rate of each coil. Table 11.8 gives the numbers. Each row is a drawing, or the family's average, for one coil; the columns give the rate at commissioning, at the center and today, and the band today, all in %/yr.

Table 11.8: Each coil's rate at commissioning ($t_{\mathrm{y}} = 0$), at the record's center ($t_{\mathrm{y}} = 1.5\,\mathrm{yr}$) and today ($t_{\mathrm{y}} = 3\,\mathrm{yr}$), as $100(\mathrm{e}^{\gamma_u(t_{\mathrm{y}})} - 1)$, in each drawing and averaged over the family; bands are 5–95 %.

Coil	Drawing	At commissioning [%/yr]	At the center [%/yr]	Today [%/yr]	Band today [%/yr]
1	M_0	−4.151	−4.151	−4.151	−6.430 to −1.817
1	M_1	−0.374	−4.480	−8.417	−13.647 to −2.871
1	M_2	−4.173	−4.173	−4.173	−6.454 to −1.836
1	M_{12}	+0.305	−4.598	−9.262	−14.718 to −3.457
1	averaged	–	–	**−5.697**	**−13.058 to −2.122**
2	M_0	−6.062	−6.062	−6.062	−8.084 to −3.995
2	M_1	−5.867	−5.867	−5.867	−7.902 to −3.788
2	M_2	−6.938	−6.060	−5.173	−9.977 to −0.112
2	M_{12}	−8.092	−5.828	−3.507	−8.657 to +1.933
2	averaged	–	–	**−5.691**	**−8.318 to −0.831**

Coil 1 is losing 5.70 %/yr today averaged over the family, with band −13.06 to −2.12 %/yr; the full posterior gives a median of −5.665 and a band of [−12.996, −2.106], with Monte Carlo standard errors of $0.00006\,\mathrm{yr}^{-1}$ on the median and 0.00015 and 0.00006 on the band ends. If the rate may change it is −9.262 %/yr, 2.23 times the constant rate's −4.151. The variance of the average is 1.4168×10^{-3}, of which 8.1533×10^{-4} is within the drawings and 6.0151×10^{-4} between them, by equation (7.11): a share of 0.425 of the variance is the record's indecision over the form. In M_{12}, by equation (11.11) with $t_{\mathrm{y}} - \bar{t}_{\mathrm{y}} = 1.5\,\mathrm{yr}$,

$$\begin{aligned} \hat{\gamma}_1(3) &= -0.04707 + 1.5\,(-0.03341) = -0.09719\,\mathrm{yr}^{-1}, \\ \mathrm{Var}(\gamma_1(3) \mid \mathbf{y}, M_{12}) &= 2.2513 \times 10^{-4} + 2.25\,(4.4849 \times 10^{-4}) + 3\,(6.2387 \times 10^{-5}) \\ &= 1.4214 \times 10^{-3}\,\mathrm{yr}^{-2}, \end{aligned}$$

and the cross term, positive through the correlation +0.1963 of the rate and the curvature, adds to the variance. Through the lever 1.5 yr the curvature's term carries 0.710 of it, the rate's 0.158 and the cross term 0.132. Today's rate exceeds the center's in M_{12} exactly when the curvature is negative, with probability 0.9427. Coil 2 is losing 5.69 %/yr today averaged, band −8.32 to −0.83; the full posterior gives −5.719 and [−8.358, −0.922].

Coil 1's rate across the record is drawn in Figure 11.4. The horizontal axis is the time since commissioning, the vertical axis the rate in %/yr. The gray band and line are M_0's constant rate, the blue band and line M_{12}'s rate (11.1), whose band widens away from the center; the interval at today's window is the family's average, and the dotted line the stated truth, a check marker for Section 11.4.6. The two drawings cross near the center and part toward both ends of the record; the averaged interval today spans most of both drawings' bands.

The operational stake is one line: the forecast of the capacity limit, Chapter 14's question, reads today's rate, and the two drawings disagree on it by a factor of 2.23; the forecast must carry both, or their average.

11.3.4 The dashboard's bend as a quantity of the state

Asset manager's question. Is the dashboard's bend in year 3 the coil? The question asks what the bend of the dashboard's quadratic measures, and how much of it the coils, the load and the season, and the sensor each make.

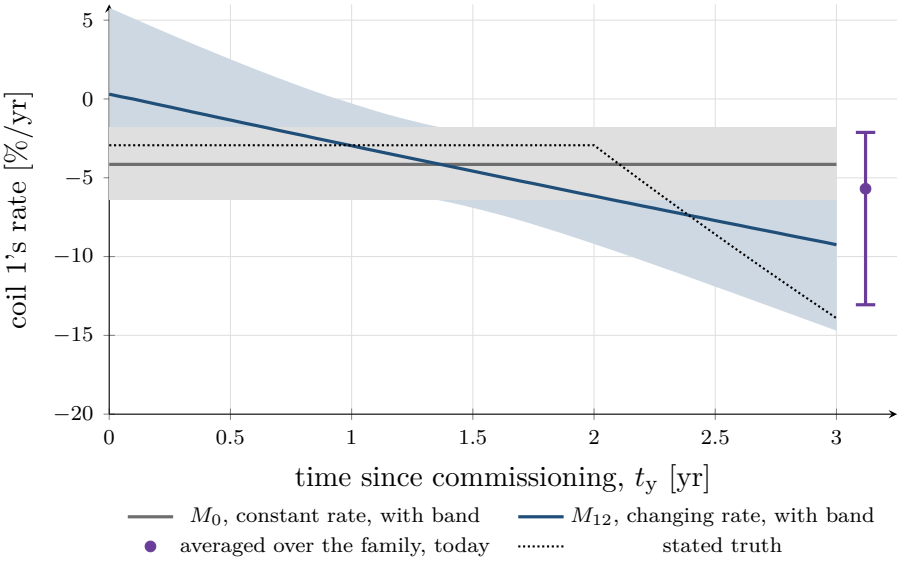


Figure 11.4: Coil 1's rate against time. Gray: the constant-rate drawing M_0, $-4.151\,\%/\mathrm{yr}$ with band -6.430 to -1.817. Blue: M_{12}'s rate $100(e^{\gamma_1(t_y)} - 1)$ with $\gamma_1(t_y) = -0.04707 - 0.03341(t_y - 1.5)$ and its 5–95 % band from equation (11.11) with $\Sigma_{\gamma_1\gamma_1} = 2.2513 \times 10^{-4}$, $\Sigma_{\gamma_1\dot{\gamma}_1} = 4.4849 \times 10^{-4}$ and $\Sigma_{\gamma_1\dot{\gamma}_1} = 6.2387 \times 10^{-5}$. Purple: the family's average today, median -5.697 and band -13.058 to -2.122, drawn just right of $t_y = 3$. Dotted: the stated truth, $-0.0300\,\mathrm{yr}^{-1}$ to $t_y = 2$ and falling by $0.12\,\mathrm{yr}^{-2}$ after it, a check marker (Section 11.4.6).

General form. An estimation question about a derived quantity is summarized by its moments, equation (2.10),

$$\mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}] = \int f(\mathbf{z})\, p(\mathbf{z} \mid \mathbf{y})\, d\mathbf{z}, \qquad \mathrm{Var}(f(\mathbf{z}) \mid \mathbf{y}) = \int \big(f(\mathbf{z}) - \mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}]\big)^2\, p(\mathbf{z} \mid \mathbf{y})\, d\mathbf{z}.$$

Specialized form. The question is equation (2.10) with $f(\mathbf{z}) = \dot{s}_u^{\mathrm{raw}}(\mathbf{z})$ and $p(\mathbf{z} \mid \mathbf{y})$ the Laplace normal of M_{12}. The tiles are Chapter 5's $h_{u,t}^{\mathrm{raw}}(\mathbf{z}) = A_{u,0}(\mathbf{z})/A_{u,t}(\mathbf{z})$, with the approach $A_{u,t}$ of equation (5.18) at window t's entries and inputs and unit 1's discharge carrying the drift (10.2). The dashboard fits a quadratic in $t_y - \bar{t}_y$, with $t_y = t\,\Delta t$ and window 18 at the center, to the 37 tiles by least squares, and reports its bend, twice the quadratic coefficient:

$$\dot{s}_u^{\mathrm{raw}}(\mathbf{z}) = 2\,\mathbf{e}_3^{\mathsf{T}}(\mathbf{X}^{\mathsf{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathsf{T}}\,\mathbf{h}_u^{\mathrm{raw}}(\mathbf{z}), \qquad \mathbf{X} = \begin{bmatrix} 1 & t_{y,0} - \bar{t}_y & (t_{y,0} - \bar{t}_y)^2 \\ \vdots & \vdots & \vdots \\ 1 & t_{y,36} - \bar{t}_y & (t_{y,36} - \bar{t}_y)^2 \end{bmatrix}, \tag{11.12}$$

with $\mathbf{h}_u^{\mathrm{raw}}$ the column of the 37 tiles and $\mathbf{e}_3$ the third unit vector. The row $2\,\mathbf{e}_3^{\mathsf{T}}(\mathbf{X}^{\mathsf{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathsf{T}}$ is fixed by the windows' times alone. The quadratic's slope today, at $t_y = 3\,\mathrm{yr}$, is $(\mathbf{e}_2 + 3\,\mathbf{e}_3)^{\mathsf{T}}(\mathbf{X}^{\mathsf{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathsf{T}}\mathbf{h}_u^{\mathrm{raw}}$. The bend is a fixed linear combination of the tiles, so its gradient is $2\,\mathbf{e}_3^{\mathsf{T}}(\mathbf{X}^{\mathsf{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathsf{T}}$ times the 37×117 Jacobian of the tiles, dense over every window's block and the drift rate, and its variance is $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$ by the delta method (4.53). One choice enters here: the split is sequential, that of Chapter 10's Section 10.3.3. The bend of unaged coils with a sound sensor is the part load and season make; the coils as fitted, still with a sound sensor, add the coils' part; the drift adds the sensor's; and the readings' own bend less that of the fitted state is the reading errors'. The order moves the terms by second-order amounts, and the split is used only to take the dashboard's number apart.

Building blocks. The MAP state and the covariance of M_{12}; the approach (5.18) with the drift; the fixed row of equation (11.12); the gradient of the bend.

The bend taken apart. Table 11.9 gives the split. Each row is a coil; the columns give the bend of the readings' tiles and its four parts, each with its share, in units of the commissioned tile per yr^2.

Table 11.9: The dashboard's bend in coil 1's and coil 2's tiles taken apart on the fitted state of M_{12}, in units of the commissioned tile per yr^2, with the share of the readings' bend in parentheses.

Coil	Readings' bend [1/yr^2]	Load and season [1/yr^2] (%)	Coils [1/yr^2] (%)	Sensor [1/yr^2] (%)	Reading errors [1/yr^2] (%)
1	−0.0579	+0.0025 (−4.4)	−0.0180 (31.1)	−0.0312 (53.9)	−0.0112 (19.4)
2	+0.0490	+0.0025 (5.2)	+0.0609 (124.3)	0 (0)	−0.0144 (−29.5)

More than half of coil 1's bend is the sensor, −0.0312 of the readings' −0.0579, 53.9 %. A drift that starts at window 20 and grows linearly is a hinge in unit 1's discharge reading, and a quadratic through a hinge bends. The coils' part is less than a third, 31.1 %, and the reading errors add a fifth. From the state, the bend is $-0.04662\,\mathrm{yr}^{-2}$ for coil 1, standard deviation 0.02735 and 5–95 % band [−0.09161, −0.00164]; the full posterior gives a median of −0.04670 and a band of [−0.09316, −0.00330]. For coil 2 it is +0.06339 (standard deviation 0.02144). The quadratic through the coil's own health index $h_{u,t}$ along the fitted path bends by −0.02808 for coil 1 and +0.02442 for coil 2. Coil 2's tile reads no drift: its coils' part exceeds its whole bend, and its reading errors take part of it back. The dashboard's slope today for coil 1 is −0.22517 from the state (standard deviation 0.04152), against the readings' −0.24441: load and season −0.01997, the coils −0.12645 and the sensor −0.07876, shares 0.089, 0.562 and 0.350. The planner's "a quarter a year" is a third sensor.

The tiles of coil 1 are set on one axis in Figure 11.5. The horizontal axis is the window, the vertical axis the tile in units of the commissioned one; the dots are the readings' tiles and the black line the dashboard's quadratic through them; the gray line is the tile of unaged coils with a sound sensor, the blue line the tile of the coils as fitted with a sound sensor, and the purple line the tile of the full fitted state, drift included. The gray line falls in the first month and then swings with the season below the commissioning value; the blue line falls steadily; the purple line leaves the blue at window 20 and falls below it by the drift, and the dashboard's quadratic follows the purple line and the scatter of the dots.

11.3.5 The drift with the acceleration allowed

Asset manager's question. With the coil allowed to accelerate, does unit 1's sensor still need its drift? Both the drift and coil 1's acceleration raise unit 1's discharge reading late in the record, and the question is whether the curvature takes the drift's place.

General form. A hypothesis comparison is the ratio of the evidences of two drawings, equation (2.14),

$$B_{AB} = \frac{p(\mathbf{y} \mid M_A)}{p(\mathbf{y} \mid M_B)}.$$

Specialized form. The question is equation (2.14) with M_A a drawing of the family and M_B the same drawing without the drift entry $\dot{b}_{d,1}$: a nested pair, so equation (7.8) gives the log factor with $\sigma = \sigma_{\dot{b}}$. Three explanations of the late rise of unit 1's reading are then set side by side at even prior odds by equation (7.9): the drift alone, M_0; coil 1's acceleration alone, M_1 without the drift; and both, M_1.

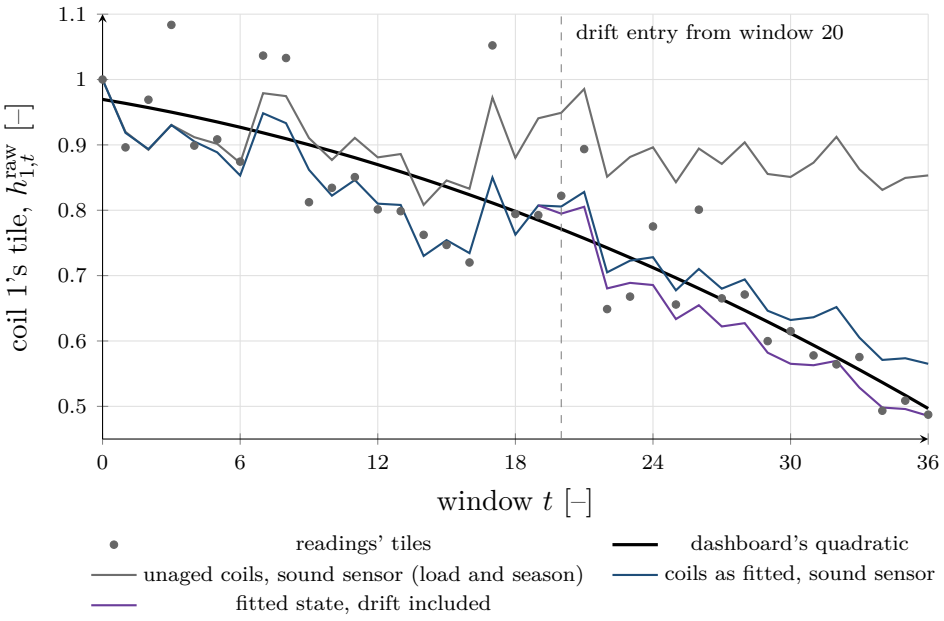


Figure 11.5: Coil 1's tiles, the commissioning approach over each window's, in the 37 windows. Dots: the readings' tiles. Black: the dashboard's least-squares quadratic through them, bend $-0.0579\,\text{yr}^{-2}$. Gray: the tiles of unaged coils with a sound sensor at each window's load and inputs. Blue: the tiles of the coils as fitted by M_{12}, with a sound sensor. Purple: the tiles of the full fitted state, with the drift on unit 1's discharge from window 20; before window 20 it coincides with the blue line.

Building blocks. The fits and evidences of Section 11.2.3 for the drawings with and without the drift entry; the prior spread $\sigma_{\dot{b}} = 2.0\,\text{K/yr}$.

The drift. The drift is still needed. Within M_1,

$$\ln B(\text{drift} : \text{none} \mid M_1) = +3.146, \qquad B = 23.24,$$

and within M_{12} the log factor is $+2.511$, $B = 12.31$, against Chapter 10's $+6.071$, $B = 433.3$, within M_0. The curvature takes $0.1195\,\text{K/yr}$ of the drift: the drift rate is $0.5642\,\text{K/yr}$ (standard deviation 0.1838, band $[0.2620,\ 0.8665]$) in M_{12}, against a full posterior median of 0.5739 and band of $[0.2747,\ 0.8764]$; 0.5926 (0.1803) in M_1; and 0.6837 (0.1681) in M_0. At window 36 the drift in M_{12} is $0.7994\,\text{K}$ (standard deviation 0.2604, band $[0.3711,\ 1.2276]$). Without the drift entry the record would read coil 1 as accelerating: $\ln B(M_1 : M_0) = +2.980$, a factor of 19.69, and in M_{12} the curvature -0.06000 (standard deviation 0.01958) with $\Pr(\dot{\gamma}_1 < 0) = 0.9989$. That drawing's own evidence, $\ln p(\mathbf{y} \mid M) = -72.0245$, stands against M_0's -68.9335 and M_{12}'s -69.4689, and over the three explanations at even odds,

$$\text{the drift alone } 0.476, \qquad \text{the acceleration alone } 0.022, \qquad \text{both } 0.503.$$

Along the ridge of the covariance, a drift faster by $0.1\,\text{K/yr}$ takes $+0.00466\,\text{yr}^{-2}$ from coil 1's curvature. The dashboard's error is in one line: the planner's acceleration is the drawing without the sensor's drift, which the record rejects.

11.4 The checks (step 5)

Each answer of Section 11.3 rests on approximations, priors and choices that the record lets be tested, and on one record of sensor errors, which fresh records at the stated truth put in context.

11.4.1 The full posterior against the Laplace answers

The reweighted draws of Section 11.2.3 have stood beside every answer of step 4. Their Monte Carlo standard errors, by batch means, are at most $0.00013\,\mathrm{yr}^{-2}$ on the medians and band ends of the curvatures, 0.0006 and 0.0013 on the two sign probabilities, at most 0.00040 on the log evidences, and $0.00006\,\mathrm{yr}^{-1}$ on the median of coil 1's averaged rate today. The medians of the curvatures lie 0.00045 and 0.00046 from the MAP values, 4.5 and 5.2 standard errors: a resolved gap, small against the posterior standard deviations of 0.021. The sign probabilities differ by 0.0010 for coil 1, 1.7 standard errors, not resolved, and by 0.0039 for coil 2, 3.0. The log evidences by importance sampling lie below the Laplace evidences by 0.0040, 0.0006, 0.0080 and 0.0081 for M_0, M_1, M_2 and M_{12}, so $\ln B(M_1 : M_0)$ moves by $+0.0034$, and the four drawings' probabilities agree to 0.0012. M_{12}'s rate today has a full posterior median of $-0.09612\,\mathrm{yr}^{-1}$ (standard error 0.00018), $-9.165\,\%/\mathrm{yr}$, against the Laplace -0.09719, and the family's averaged rate a median of -5.665 against $-5.697\,\%/\mathrm{yr}$. Coil 1's dashboard bend has a full posterior median of -0.04670 (standard error 0.00010) against -0.04662, and the drift rate $0.5739\,\mathrm{K/yr}$ (standard error 0.0007) against 0.5642, the largest gap against its posterior standard deviation, which the draws resolve: the drift's posterior leans toward faster drifts. The Laplace normal is the posterior to within a fiftieth of a standard deviation in the curvatures, a thirtieth in M_{12}'s rate today and a tenth in the drift rate; the draws give the full posterior, and the Laplace route is not called exact.

11.4.2 The limit of exact rows and the second-order term

The fit runs on the reduced state. The check fits M_{12}'s own full state, the 117 entries of equation (11.3) and each window's five temperatures, 302 entries in all, with the 185 physical rows (5.2) at a common finite tolerance beside the readings, the trend rows and the priors, and lets the tolerance shrink. Table 11.10 gives the result: each row is a tolerance, the columns the two curvatures and the drift rate with their standard deviations and the iterations taken, and the last row the reduced state. At 0.1 the spreads are slightly wider; at 0.01 and 0.001 the full state reproduces the reduced state to the fourth decimal.

Table 11.10: The limit of exact physical rows on the 302-entry full state of M_{12}: the curvatures and the drift rate with their standard deviations at three tolerances, against the reduced state.

Tolerance σ_p [kW or K]	$\dot\gamma_1$ [yr^{-2}]	sd [yr^{-2}]	$\dot\gamma_2$ [yr^{-2}]	sd [yr^{-2}]	$\dot b_{d,1}$ [K/yr]	sd [K/yr]	Iterations [–]
0.1	−0.03351	0.02126	+0.01609	0.02012	0.5651	0.1856	12
0.01	−0.03341	0.02118	+0.01623	0.02009	0.5643	0.1838	5
0.001	−0.03341	0.02118	+0.01623	0.02009	0.5642	0.1838	4
reduced state	−0.03341	0.02118	+0.01623	0.02009	0.5642	0.1838	–

The covariance is the Gauss–Newton form, without the second-order term $\mathbf{C} = \mathbf{H} - \mathbf{J}^{\mathsf{T}}\mathbf{J}$ of equation (4.44). The full Hessian at $\hat{\mathbf{z}}$, computed by differences, gives the Newton covariance, in which the standard deviations of $\dot\gamma_1$, $\dot\gamma_2$ and $\dot b_{d,1}$ are 0.02105, 0.02000 and 0.18335, ratios of 0.9941, 0.9955 and 0.9976 to Gauss–Newton's; those of γ_1 and γ_2 are in ratios of 0.9979 and 0.9991. The largest entry of $\mathbf{C}$ is 8.39×10^{-4} of the largest entry of $\mathbf{J}^{\mathsf{T}}\mathbf{J}$, and $\ln\det\mathbf{H} - \ln\det\mathbf{J}^{\mathsf{T}}\mathbf{J} = 0.0216$, which would enter the Laplace evidence (7.7) of M_{12} with the factor $-\frac{1}{2}$. The Gauss–Newton covariance is the one reported.

11.4.3 Sensitivity to the priors and the choices

Table 11.11 repeats the fit under other priors and choices; each variant refits M_0, M_1 and M_{12} at the same options. Each row is a variant; the columns give each coil's curvature in M_{12} with its standard deviation and sign probability, the two log Bayes factors against M_0, and the drift rate.

Table 11.11: The curvatures, their evidence and the drift rate under other priors and choices. $P_u = \Pr(\dot\gamma_u < 0 \mid \mathbf{y}, M_{12})$.

Variant	$\dot\gamma_1$ [yr^{-2}]	sd [yr^{-2}]	P_1 [–]	$\dot\gamma_2$ [yr^{-2}]	P_2 [–]	$\ln B$, M_1 [–]	$\ln B$, M_{12} [–]	$\dot b_{d,1}$ [K/yr]
the reported drawings	−0.03341	0.02118	0.9427	+0.01623	0.2096	+0.0549	−0.5354	0.5642
rating prior sd 0.05	−0.03228	0.02080	0.9397	+0.01672	0.1980	−0.0082	−0.5851	0.5661
rating prior sd 0.30	−0.03353	0.02122	0.9430	+0.01618	0.2108	+0.0613	−0.5301	0.5641
drift prior sd 0.7059 K/yr (1.0 K at window 36)	−0.03486	0.02110	0.9507	+0.01687	0.2001	+0.1518	−0.4124	0.5328
drift prior sd 4.0 K/yr	−0.03325	0.02119	0.9417	+0.01615	0.2107	+0.0443	−0.5489	0.5678
survey sd 0.04 on β_1	−0.03354	0.02117	0.9434	+0.01629	0.2087	+0.0634	−0.5245	0.5614
$\sigma_q = 0.0025$	−0.03244	0.01898	0.9563	+0.02074	0.1192	−0.1909	−0.5455	0.5619
$\sigma_q = 0.01$	−0.03299	0.02693	0.8898	+0.00698	0.3955	+0.0797	−0.5276	0.5613
the straight line, $\sigma_q \to 0$	−0.03186	0.01811	0.9607	+0.02269	0.0857	−0.3859	−0.5580	0.5602
drift onset window 16	−0.04284	0.02034	0.9824	+0.02035	0.1531	+0.8525	+0.4498	0.3850
drift onset window 24	−0.02519	0.02207	0.8732	+0.01286	0.2628	−0.4041	−1.1097	0.8316
drift onset window 28	−0.01978	0.02244	0.8110	+0.01111	0.2921	−0.6008	−1.3558	1.3585
no drift entry	−0.06000	0.01958	0.9989	+0.02735	0.0808	+2.9803	+3.0252	–
window 20's unit-2 discharge reading left out	−0.03100	0.02120	0.9282	+0.00908	0.3263	+0.0500	−0.7600	0.5510

The priors, coil by coil. The rating, survey and drift-size priors move $\ln B(M_1 : M_0)$ by at most 0.0969, $\ln B(M_{12} : M_0)$ by at most 0.1230, coil 1's curvature by at most 0.00145 and coil 2's by at most 0.00064. A rating prior three times tighter moves both curvatures up, coil 1's from −0.03341 to −0.03228 and coil 2's from +0.01623 to +0.01672: both commissioned entries lie below the rating, so the tighter prior raises both, and each curvature is correlated positively with its coil's commissioned entry (+0.5425 and +0.5519), so coil 1 accelerates less and coil 2 slows more. A tighter drift prior shrinks the drift to 0.5328 K/yr, and the ridge gives coil 1's curvature what the drift gives up, −0.03486; coil 2's moves up through its negative correlation with the drift, to +0.01687. The wander changes the curvature's spread, 0.01898 at $\sigma_q = 0.0025$ and 0.01811 on the straight line, because a smaller wander reads the curvature more directly from the readings, but not the verdict: $\ln B(M_1 : M_0)$ stays between −0.3859 and +0.0797. Coil 2's slowing leans on window 20's unit-2 discharge reading, Chapter 10's largest normalized residual: without it P_2 rises from 0.2096 to 0.3263.

The drift's onset. The drift's onset moves the verdict most, $\ln B(M_1 : M_0)$ from +0.8525 at window 16 to −0.6008 at window 28. An earlier drift is slower and takes less of the late rise, leaving more to coil 1's acceleration; a later one is steeper and takes more. This is Chapter 10's ridge between the drift and coil 1, now on the curvature. Over onsets 16 to 28 no log Bayes factor reaches +1, and at onset 28 $\ln B(M_{12} : M_0)$ falls to −1.3558.

The form of the change. Table 11.12 sets two knees beside the reported form. In a knee form the rate is γ_u to window t_{knee} and changes by $\dot\gamma_u$ per year after it, $\gamma_u(t_y) = \gamma_u + \dot\gamma_u\,(t_y - t_{\text{knee}}\Delta t)_+$, with the same prior spread on the change; each row is a form, the columns coil 1's curvature, its spread and sign, the two log Bayes factors and coil 1's rate today. No form reaches $|\ln B| = 1$.

The knee at the start of year 3 leads the whole-record curvature by 0.5371 in $\ln B(M_1 : M_0)$ and moves coil 1's rate today by +0.432 points; its curvature is steeper, −0.05461, because it acts over one year instead of three, and less certain.

Table 11.12: The form of the change of rate: over the whole record (the reported form) and constant to a knee t_{knee}, changing after it.

Form	$\dot\gamma_1$ [yr^{-2}]	sd [yr^{-2}]	P_1 [–]	$\ln B$, M_1 [–]	$\ln B$, M_{12} [–]	Rate today, coil 1 [%/yr]
over the whole record (reported)	−0.03341	0.02118	0.9427	+0.0549	−0.5354	−9.262
constant to window 12, changing after	−0.04358	0.02518	0.9582	+0.2912	+0.3196	−10.251
constant to window 24, changing after	−0.05461	0.04161	0.9053	+0.5921	+0.5469	−8.830

11.4.4 The acceleration as of each window

The record cut at an earlier window t_{as}, with nothing later than t_{as}, gives the answers as they stood then. Table 11.13 gives them at the end of year 2 and each quarter of year 3; each row is an as-of window, the columns each coil's curvature in M_{12} with its spread and sign probability, and the two log Bayes factors. Figure 11.6 plots the sign probabilities in the upper panel and the log Bayes factors in the lower, against t_{as}, with the band $|\ln B| \le 1$ shaded.

Table 11.13: The acceleration as of each window t_{as}, from the record of windows 0 to t_{as} only.

t_{as} [window]	$\dot\gamma_1$ [yr^{-2}]	sd [yr^{-2}]	P_1 [–]	$\dot\gamma_2$ [yr^{-2}]	sd [yr^{-2}]	P_2 [–]	$\ln B$, M_1 [–]	$\ln B$, M_{12} [–]
24 (Dec Y2)	+0.01006	0.03772	0.3948	−0.04535	0.03593	0.8966	−0.3075	+0.1647
27 (Mar Y3)	+0.00832	0.03370	0.4025	−0.01724	0.03146	0.7082	−0.4302	−0.7400
30 (Jun Y3)	−0.00106	0.02911	0.5146	−0.00382	0.02725	0.5557	−0.5905	−1.1867
33 (Sep Y3)	−0.01285	0.02497	0.6967	+0.01215	0.02352	0.3028	−0.6899	−1.3148
36 (Dec Y3)	−0.03341	0.02118	0.9427	+0.01623	0.02009	0.2096	+0.0549	−0.5354

Coil 1's sign probability climbs through year 3, from 0.3948 at window 24 to 0.9427 today, while $\ln B(M_1 : M_0)$ stays between −0.6899 and +0.0549: the sign gathers and the verdict does not. Coil 2 runs the other way. At window 24 it reads as accelerating, with probability 0.8966, and by window 36 as slowing. Its steep months of windows 13 to 16 sit late in a 24-window record and early in a 36-window one, so a slope of the monthly rates against time reads them as a late steepening in the first and an early one in the second. This is coil 2's path as the record shows it, and Chapter 12 examines it. Each as-of fit uses nothing later than its window, and none is the time at which a verdict first holds, so none needs a spread over fresh records.

11.4.5 Fresh records and the power of the test

The record is one draw of the sensor errors. The check repeats the three fits on 200 fresh records at the stated truth (seeds 1000 to 1199), each with fresh sensor errors and the same inputs, and sets the result beside the drawing's pseudo-true values, its fit to the truth's noise-free readings. Table 11.14 gives the distribution over the fresh records: each row is a quantity, the columns its median or mean, its spread, the shares beyond stated levels, and where this record falls.

At this truth the record has no power to decide the acceleration at three years. In 0.920 of the fresh records $\ln B(M_1 : M_0)$ is at or below +1, even or against, and only 0.080 pass +1;

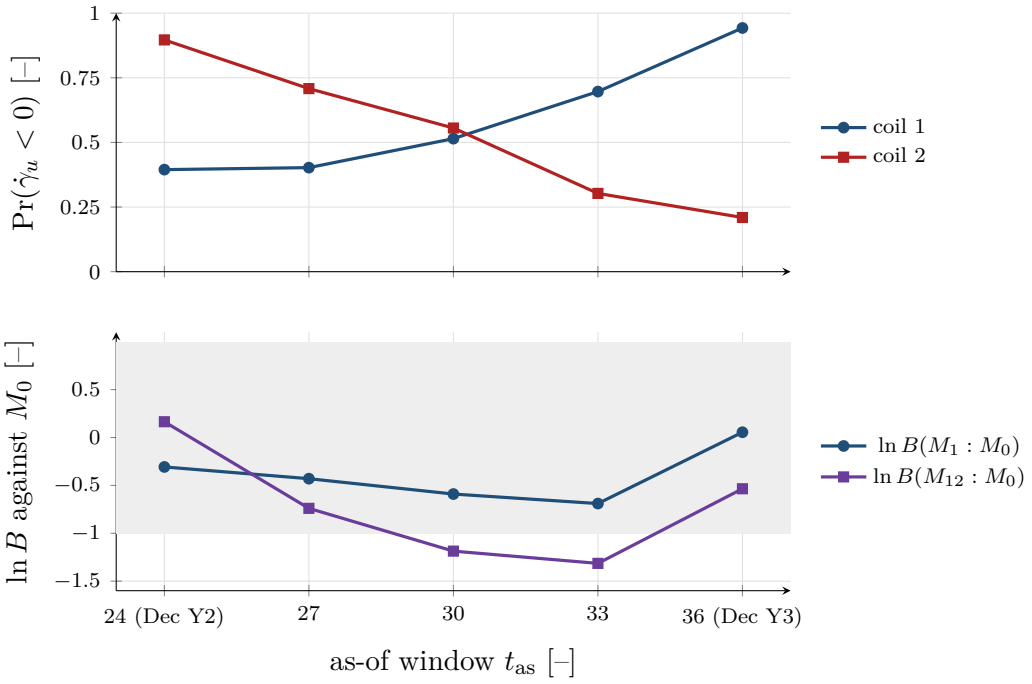


Figure 11.6: The acceleration as of each window t_{as} (Table 11.13). Upper panel: the sign probabilities of the two curvatures in M_{12}. Lower panel: the log Bayes factors of M_1 and M_{12} against the constant rate, with the band $|\ln B| \leq 1$ shaded.

Table 11.14: The evidence and the curvatures over 200 fresh records at the stated truth. Bands are 5–95 % over the records; the last column is this record's quantile among them.

Quantity	Center and spread	Shares over the 200 records [–]	This record's quantile [–]
$\ln B(M_1 : M_0)$ [–]	median −0.5108; band −0.9073 to +1.4736	above +1: 0.080; above $\ln 19$: 0.010; below −1: 0.000	0.755
$\ln B(M_{12} : M_0)$ [–]	median −0.9031; band −1.7411 to +1.8203	above +1: 0.070; above $\ln 19$: 0.010; below −1: 0.465	0.640
$\hat{\gamma}_1$ in M_{12} [yr^{-2}]	mean −0.02362; sd 0.01577 (posterior sd, mean 0.02132)	sign probability: median 0.8524, above 0.95 in 0.225; band covers the pseudo-true value in 0.980	0.255
$\hat{\gamma}_2$ in M_{12} [yr^{-2}]	mean +0.01529; sd 0.01495 (posterior sd, mean 0.02008)	sign probability: median 0.2048, above 0.95 in 0.000; band covers the pseudo-true value in 0.970	0.490

this record's +0.0549 sits at the 0.755 quantile, one of the luckier ones. The sign is easier: its probability has median 0.8524 and passes 0.95 in 0.225 of the records. The fitted curvature scatters with standard deviation 0.01577 against a mean posterior spread of 0.02132, and the band covers the pseudo-true value in 0.980 of the records for coil 1 and 0.970 for coil 2: the band is conservative, as Chapter 10's was, by the wander. This is the record's limit, not the drawing's error.

11.4.6 The answers against the stated truth

The last check sets the answers against the stated truth of the record (Section 11.6.6). Table 11.15 collects the routes: each row is a route or a reference value, the columns coil 1's curvature with its band and sign probability, coil 2's curvature and sign probability, $\ln B(M_1 : M_0)$, and coil 1's rate today averaged over the family.

Table 11.15: The routes to the curvatures, their evidence and coil 1's rate today, with the pseudo-true values and the stated truth. Bands are 5–95 %; $P_u = \Pr(\dot\gamma_u < 0 \mid \mathbf{y}, M_{12})$.

Route	$\dot\gamma_1$ [yr^{-2}]	band [yr^{-2}]	P_1 [–]	$\dot\gamma_2$ [yr^{-2}]	P_2 [–]	$\ln B$ $M_1 : M_0$ [–]	Rate today, coil 1 [%/yr]
Laplace on the stacked state (reported)	−0.03341	−0.06825 to +0.00142	0.9427	+0.01623	0.2096	+0.0549	−5.697
Full posterior, reweighted draws	−0.03296	−0.06800 to +0.00129	0.9437	+0.01577	0.2135	+0.0583	−5.665
Savage–Dickey on the Laplace marginal	–	–	–	–	–	+0.0667	–
Full state, $\sigma_p = 0.1$	−0.03351	sd 0.02126	–	+0.01609	–	–	–
Newton covariance	–	sd 0.02105	–	sd 0.02000	–	–	–
Drawing's pseudo-true value	−0.02245	–	–	+0.01347	–	−0.5264	–
The true path's own curvature; the true rate today	−0.03110	–	–	+0.00908	–	–	−13.929

Coil 1's true path. Coil 1 accelerates from the start of year 3. Its true rate is $-0.0300\,\mathrm{yr}^{-1}$ to window 24 and -0.1500 at window 36, −13.929 %/yr: a curvature of $-0.1200\,\mathrm{yr}^{-2}$ in year 3 and none before. The least-squares slope of the true path's monthly rates over the whole record is −0.03110; the drawing's pseudo-true value is −0.02245, smaller because at the truth's noise-free readings the drift entry and the return air share the late rise, with a pseudo-true drift of 0.6136 K/yr against the true 0.6000. The posterior lies −0.52 standard deviations from the pseudo-true value. Coil 2's stated truth has no curvature: the slope of its true path's monthly rates is +0.00908, its pseudo-true value +0.01347, and the posterior lies +0.14 standard deviations from it.

The rate today. Coil 1's true rate today is $-0.1500\,\mathrm{yr}^{-1}$. M_{12}'s −0.09719 lies +1.40 standard deviations from it, its pseudo-true value being −0.07570; M_0's −0.04240 lies +7.36 standard deviations from it; and the truth sits at the 0.0298 quantile of the family's averaged posterior. The constant-rate drawing is the one that misses.

The health index at window 36. Under M_{12} coil 1's $\ln h_{1,36}$ is −0.14122 (standard deviation 0.03356, $h = 0.8683$), +0.26 standard deviations from the true −0.1500 ($h = 0.8607$), against +0.71 under M_0; coil 2's is −0.18013 (0.02615, $h = 0.8352$), −1.92 standard deviations from the true −0.1300 ($h = 0.8781$), against −2.22 under M_0. The curvature gives back to coil 1's

window-36 entry part of what the constant rate handed to the drift and to coil 2: $g_{1,36}$ is -0.16328 under M_{12} against -0.13314 under M_0 and the true -0.1700. Coil 2's remaining gap is not a matter of curvature, and Chapter 12 takes up coil 2's path. The total conductance at window 36 is 11.6995 kW/K (standard deviation 0.1394), -0.80 standard deviations from the true 11.8113.

The dashboard and the drift. At the true state the dashboard's bend in coil 1's tiles is $-0.03766\,\mathrm{yr}^{-2}$, of which the sensor makes -0.03408 and the coils -0.01093: in the truth too, the bend is mostly the sensor. Coil 2's is $+0.05010$, its coils $+0.04276$. The true drift rate is 0.6000 K/yr, and M_{12}'s 0.5642 lies -0.19 standard deviations from it.

11.5 The report (step 6)

The dashboard's year-3 plunge in coil 1 is more than half the drifting sensor: 53.9 % of its bend. Read on the conductance, coil 1's rate is probably rising, with probability 0.943 if it may change at all, but the record does not decide whether it changes: the Bayes factor against the constant rate is 1.056, even, and no prior spread on a grid from 0.002 to $1.0\,\mathrm{yr}^{-2}$ takes $\ln B$ above $+0.2380$. Averaged over that indecision, coil 1 is losing 5.7 %/yr today (2.1 to 13.1); the constant rate says 4.2 and the accelerating drawing 9.3. Coil 2 is not accelerating, 0.090 over the family; its record leans to a slowing rate, carried by its early months and by windows 13 to 16. Without the drift entry the record would read coil 1's acceleration as nearly certain, 0.9989 with a Bayes factor of 19.69; with the acceleration allowed, the drift is still required, by a factor of 23.24. Three years of record cannot decide an acceleration of this size, and the next measurement is the record itself.

11.5.1 The answers and their bands

Table 11.16 gives the answers in the order of Table 11.2. Each row is a question; the second column gives the reported value, the third the full posterior's value beside it, and the fourth the band: the 5–95 % band of a quantity, and for a Bayes factor its range over the prior spreads and the drift's onsets of Section 11.4.3.

11.5.2 The entry in the maintenance log

The trend assessment of window 36 is the next entry of the two-unit hall's maintenance log, after Chapter 10's entry at the same window. It uses nothing later than window 36, and in the values of Table 11.16 it reads:

> *Window 36 (Dec Y3), whether the coils' rates of loss are changing, from the condition record (windows 0 to 36).*
> Coil 1: accelerating 0.50, constant 0.46, slowing 0.04. The rate is more likely rising than falling (0.94 if it changes), but the record does not decide whether it changes (Bayes factor 1.06). Rate today -5.7 %/yr (-13.1 to -2.1), averaged; -4.2 if constant, -9.3 if accelerating.
> Coil 2: no sign of acceleration (0.09); the record leans to a slowing rate, not tested against a single event in its path (Chapter 12).
> Unit 1's discharge sensor: the drift is still required with the acceleration allowed (Bayes factor 23); 0.56 K/yr (0.26 to 0.87), 0.80 K today (0.37 to 1.23).
> Dashboard bend in year 3 (coil 1's tile): more than half the sensor; not a measure of the coil's acceleration.
> Action: no cleaning decision from the acceleration alone; the forecast of the capacity limit must carry both drawings (Chapter 14), and the October shutdown's work order waits on it.
> Next measurement: continue the record to the shutdown and refit at window 45 (Sep Y4); a reference condition test now if one window is to be spent.

Table 11.16: The report on the two-unit hall's rates of loss at window 36 (Dec Y3). Reported values from the Laplace normal of each drawing; bands are 5–95 %, or for a Bayes factor the range over the variants named.

Asset manager's question	Reported value	Full posterior	Band
Is the loss accelerating?	$\Pr(\dot\gamma_u < 0 \mid \mathbf{y}, M_{12})$: coil 1 0.943, coil 2 0.210 [–]	0.9437; 0.2135 [–]	$\dot\gamma_1$: −0.06825 to +0.00142; $\dot\gamma_2$: −0.01681 to +0.04927 [yr^{-2}]
Does the record need a changing rate at all, or does the constant rate do?	$B(M_1 : M_0) = 1.056$ [–]; coil 1 accelerating 0.500, constant 0.461, slowing 0.039; coil 2 0.090, 0.675, 0.235 [–]	$B(M_1 : M_0) = 1.060$ [–]	$\ln B(M_1 : M_0)$: −1.9795 to +0.2247 over $\sigma_{\dot\gamma}$ from 0.01 to 0.5; −0.6008 to +0.8525 over onsets 16 to 28 [–]
How fast is it losing now, against the record's average?	averaged: coil 1 −5.697, coil 2 −5.691 [%/yr]; coil 1 in M_0 −4.151, in M_{12} −9.262 [%/yr]	−5.665; −5.719 [%/yr]	coil 1: −13.058 to −2.122; coil 2: −8.318 to −0.831 [%/yr]
Is the dashboard's bend in year 3 the coil?	coil 1's bend from the state −0.04662 [1/yr^2]; of the readings' −0.0579, sensor 53.9 %, coils 31.1 %, reading errors 19.4 %, load and season −4.4 %	−0.04670 [1/yr^2]	−0.09161 to −0.00164 [1/yr^2]
With the coil allowed to accelerate, does unit 1's sensor still need its drift?	yes: $B(\text{drift} : \text{none} \mid M_1) = 23.24$ [–]; $\dot b_{d,1} = 0.5642$ [K/yr]	$\dot b_{d,1}$: 0.5739 [K/yr]	0.2620 to 0.8665 [K/yr]; drift at window 36 0.3711 to 1.2276 [K]

Whether either coil is to be cleaned at the shutdown is not settled here; it needs the crossing date of Chapter 14 and the decision level of Chapter 19, both in Volume II.

Common mistake

Reading the sign probability of a curvature as evidence that the rate changes. The sign is read within a drawing that already lets the rate change; whether the record needs that drawing is the Bayes factor, which charges the curvature for its prior width. A record can give a sign probability of 0.94 and an even Bayes factor at once, as this one does.

11.6 The next measurement (step 7)

The report leaves coil 1's acceleration undecided, with a curvature known to 0.02118 yr^{-2}. A further measurement is one row or more of the residual vector, now or in windows not yet read, and what it removes from that spread is known before its value is.

11.6.1 The preposterior of the curvature

The preposterior is equation (4.65) on M_{12}'s stacked state extended to window 48, as in Chapter 10's step 7: the forecast windows carry their loads at the load's forecast growth and their path on the fitted curve, linked by their own trend rows (11.4), 153 entries in all. With no new rows the extension leaves the curvatures' spreads where they are, 0.02118 and 0.02009. A candidate a with Jacobian lines $\mathbf{J}_a$ in the columns of the extended state gives

$$\boldsymbol{\Sigma}_a = \left(\boldsymbol{\Sigma}^{-1} + \mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a\right)^{-1}, \qquad \mathrm{sd}_a(\dot\gamma_u) = \sqrt{\mathbf{e}_{\dot\gamma_u}^{\mathsf{T}}\,\boldsymbol{\Sigma}_a\,\mathbf{e}_{\dot\gamma_u}}, \tag{11.13}$$

and the width it takes off is $1 - \mathrm{sd}_a/\mathrm{sd}$ of $\dot\gamma_u$. The candidates are of four kinds. A side-by-side check of unit 1's discharge sensor now, ten paired readings against a reference of 0.1 K, reads the drift at window 36. A reference condition test now reads the return air, both discharges

and rack 1's inlet at 0.1 K. The record continued adds each later window's five readings, to the October shutdown (windows 37 to 45, to September of year 4) or through a fourth year (to window 48); with a fourth year the permanent instruments come in, water-side heat meters on both coils or a second discharge sensor on unit 1, the permanent instrument at the place of the question. The last two candidates are not measurements: the path known exactly at every window of the record, the curvature's floor, and at every window to 48.

11.6.2 The candidates

Table 11.17 ranks the candidates. Each row is a measurement; the columns give, for each coil, the curvature's spread after it by equation (11.13) and the share of the present spread it removes, and the drift rate's spread after it. Figure 11.7 has two panels. The upper one draws the widths removed as bars, two per candidate, ranked by coil 1's, with the 36-window floor as dashed lines. The lower one, for the next subsection, draws the predicted distributions of $\ln B(M_1 : M_0)$ at windows 45 and 48 as their medians and 5–95 % ranges, for records drawn from each drawing and from the stated truth, with the band $|\ln B| \le 1$ shaded.

Table 11.17: Candidate measurements for the curvatures, from the posterior of M_{12} extended to window 48; the width removed is taken from the present $\mathrm{sd}(\dot\gamma_1) = 0.02118$ and $\mathrm{sd}(\dot\gamma_2) = 0.02009\,\mathrm{yr}^{-2}$.

Candidate	$\mathrm{sd}_a(\dot\gamma_1)$ [yr^{-2}]	coil 1 [%]	$\mathrm{sd}_a(\dot\gamma_2)$ [yr^{-2}]	coil 2 [%]	$\mathrm{sd}_a(\dot b_{d,1})$ [K/yr]
Side-by-side check of unit 1's discharge sensor now, ten pairs (0.1 K)	0.01961	7.4	0.01977	1.6	0.0659
Reference condition test now (0.1 K on return air, both discharges, rack 1)	0.01750	17.4	0.01782	11.3	0.1168
The record continued to the shutdown, windows 37 to 45 (to Sep Y4)	0.01356	36.0	0.01316	34.5	0.1034
A fourth year of the record, windows 37 to 48	0.01186	44.0	0.01156	42.4	0.0886
The fourth year with water-side heat meters on both coils, permanent	0.01102	48.0	0.01106	44.9	0.0609
The fourth year with a second discharge sensor on unit 1, permanent	0.01104	47.9	0.01142	43.2	0.0555
The fourth year with the side-by-side check now	0.01162	45.2	0.01151	42.7	0.0552
The path known exactly at windows 0 to 36 (the curvature's floor)	0.01126	46.9	0.01126	44.0	0.0874
The path known exactly at windows 0 to 48	0.00742	65.0	0.00742	63.1	0.0874

The side-by-side check, which settles the sensor, takes only 7.4 % off coil 1's curvature, although it narrows the drift rate from 0.1838 to 0.0659 K/yr. The ridge with the drift, +0.4048, is not what holds the acceleration back. The lever arm is: a trend row's slope on $\dot\gamma_1$ grows with its distance from the center, −24.3056 for the increment into window 36 and −40.9722 for the increment into window 48 ($x_t = +1.4583$ and $+2.4583$ yr), so each later window reads the curvature more strongly, and the floor of equation (11.7) itself falls from 0.01155 to $0.00688\,\mathrm{yr}^{-2}$ with a fourth year. A fourth year alone, 0.01186, nearly reaches the 36-window floor, 0.01126, and a reference condition test now, 17.4 %, does more than the check. Of the permanent instruments, heat meters on both coils take 48.0 % off coil 1's spread with a fourth year and a second discharge sensor on unit 1 47.9 %, against the fourth year's 44.0 % alone.

11.6.3 The preposterior of the Bayes factor

A narrower curvature is not yet a verdict. Whether waiting decides the question is the distribution of the Bayes factor that the continued record would give, a preposterior of the

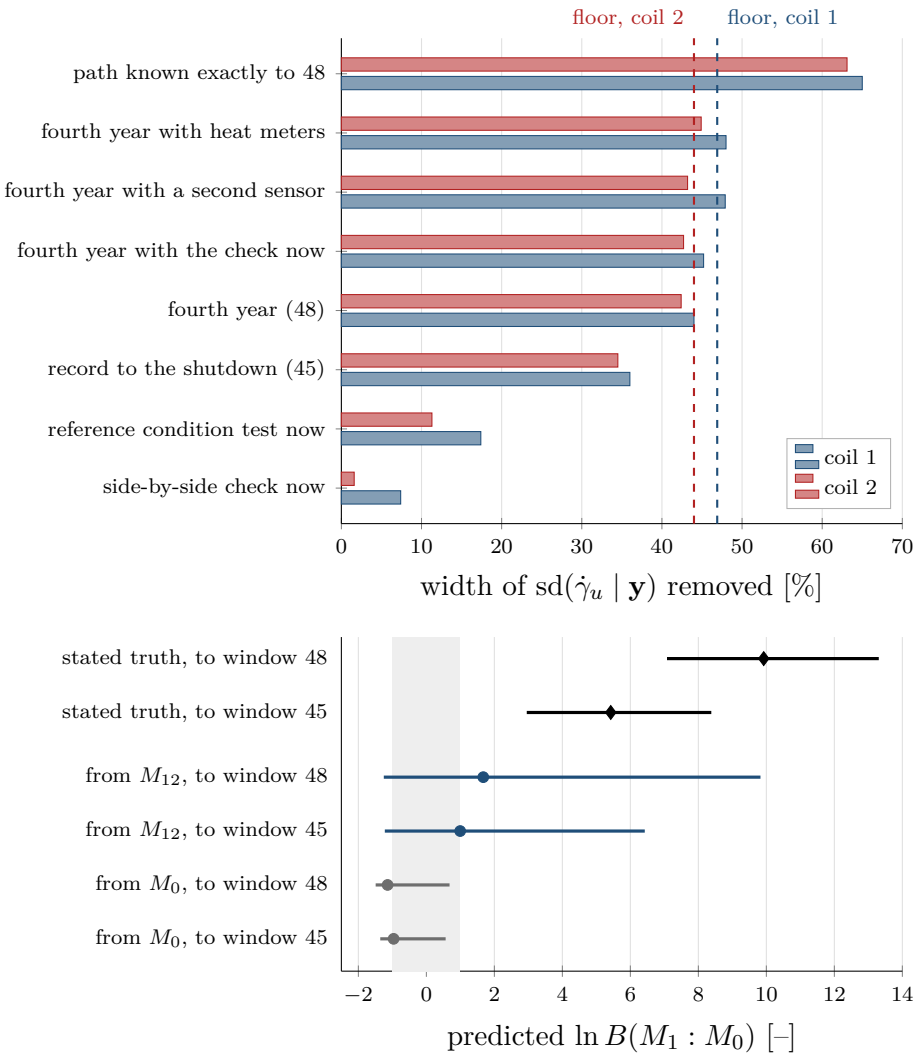


Figure 11.7: The next measurement for the curvatures. Upper panel: the width removed from each coil's sd($\dot{\gamma}_u$) by the candidates of Table 11.17, by the preposterior (11.13); dashed, the floor of the 36-window record. Lower panel: the predicted $\ln B(M_1 : M_0)$ of the record continued to window 45 or 48 (Table 11.18), median (dot) and 5–95 % range (line), for records drawn from M_0 (gray), from M_{12} (blue) and from the stated truth (black, a check marker); shaded, $|\ln B| \leq 1$.

evidence:

$$\Pr(|\ln B(\tilde{\mathbf{y}})| > 1 \mid \mathbf{y}) = \sum_{M} \Pr(M \mid \mathbf{y}) \int \Pr(|\ln B(\tilde{\mathbf{y}})| > 1 \mid \mathbf{z}, M)\, p(\mathbf{z} \mid \mathbf{y}, M)\, \mathrm{d}\mathbf{z}, \qquad (11.14)$$

with $\tilde{\mathbf{y}}$ the record continued and $\ln B(\tilde{\mathbf{y}})$ the log Bayes factor of M_1 against M_0 refitted on it. It is computed by simulation, 100 records per drawing: a state drawn from the drawing's Laplace normal, M_{12}'s (seeds 5000 to 5099) or M_0's (seeds 6000 to 6099), carried forward on the extended state; the later windows' readings predicted from it with sensor errors as the record's; and M_0 and M_1 refitted on windows 0 to 45 and 0 to 48. The simulation runs two arms in place of the four drawings of the sum: records drawn from M_{12} stand for the drawings that free coil 1's curvature, M_1 and M_{12}, and records drawn from M_0 for those that pin it, M_0 and M_2. The arms are weighted by those drawings' summed probabilities, 0.5392 and 0.4608. Table 11.18 gives the result: each row is an arm and a last window, the columns the median and range of the predicted $\ln B$ and the shares beyond the levels.

Table 11.18: The preposterior of $\ln B(M_1 : M_0)$ for the record continued, from 100 simulated records per arm; ranges are 5–95 %. The weighted rows combine the two drawings' arms with weights 0.5392 and 0.4608.

Records drawn from	Read to	Median $\ln B$ [–]	Range [–]	$\ln B > +1$ [–]	$\ln B > \ln 19$ [–]	$\ln B < -1$ [–]
M_{12}'s posterior	window 45	+0.9948	−1.2218 to +6.4244	0.50	0.25	0.10
M_{12}'s posterior	window 48	+1.6742	−1.2518 to +9.8273	0.61	0.35	0.09
M_0's posterior	window 45	−0.9638	−1.3558 to +0.5670	0.03	0.01	0.45
M_0's posterior	window 48	−1.1404	−1.4912 to +0.6822	0.02	0.00	0.61
weighted by the drawings	window 45	–	–	0.283	–	0.261
weighted by the drawings	window 48	–	–	0.338	–	0.330
the stated truth, continued	window 45	+5.4238	+2.9533 to +8.3809	–	0.95	–
the stated truth, continued	window 48	+9.9228	+7.0780 to +13.3070	–	1.00	–

Weighted by the drawings, the record continued to the shutdown decides the question one way or the other, $|\ln B| > 1$, in 0.545 of the predicted records, and a fourth year in 0.668; undecided remain 0.455 and 0.332. The continuation decides a large acceleration and leaves a small one open: among the records drawn from M_{12}, the half with the steeper drawn curvature (median $-0.04318\,\mathrm{yr}^{-2}$) reaches a median $\ln B$ of +2.6254 by window 45, and the other half (median −0.01884) +0.0465. At the stated truth continued, 100 fresh continuations of this record (seeds 7000 to 7099), $\ln B$ reaches a median of +5.4238 by window 45, above $\ln 19$ in 0.95 of them, and +9.9228 by window 48: the truth's acceleration grows with the square of time, and nine more months find it.

11.6.4 The recommendation

Keep the record running to the shutdown and refit at window 45, in September of year 4, before the shutdown's work order is written: that is the measurement that decides the acceleration, in 0.545 of the predicted records, at no cost beyond the record already kept. If one window is to be spent now, a reference condition test takes 17.4 % off coil 1's curvature, against the side-by-side check's 7.4 %; the side-by-side check stays on the list for Chapter 10's reason, because it closes Chapter 7's entry on the sensor. Of the permanent instruments, a second discharge sensor on unit 1 or heat meters on both coils add about four points to a fourth year, 47.9 and 48.0 against 44.0 %; their value for the crossing date is Chapter 23's question, in Volume II.

11.6.5 What the analysis showed

The dashboard's curve said that coil 1's fouling had taken off in year 3, a quarter of its performance a year, and that coil 2 was slowing. Read on the conductance, the bend is more than half the sensor, whose drift from window 20 puts a hinge in unit 1's tile; the physics takes out the load, the season and the other coil, the drift entry takes out the sensor, and the curvature, one entry per coil on the trend rows, carries what is left. What is left is a curvature whose sign is likely and whose existence the record does not decide. The probability adds two answers the chart cannot give, and they differ here: the sign within the drawing that allows a change, and the evidence for allowing it. It adds a third, the rate today averaged over the form, whose band spans both drawings and which the crossing forecast must carry. The record's three years cannot pay for the curvature, because a curvature is read through the lever of each month's distance from the center, and the months that read it most are the ones not yet recorded. Read jointly with Chapter 12's question of a change at a single window, as Chapter 14 in Volume II reads it when it forecasts the crossing, coil 1's curvature has probability 0.634. The analysis did not test a knee swept over windows, nor what carries coil 2's months 13 to 16.

11.6.6 Limits of the model

A constructed record. The condition record is generated from a stated truth, with the stated sensor errors, and step 5 compares the answers with that truth (Section 11.4.6). It is not field data, and the agreement is a check of the method on its own model. The truth's acceleration starts at a knee, and the reported drawing spreads it over the record.

Low power at three years. For an acceleration of this size the record's evidence is weak: 0.080 of fresh records reach $\ln B > 1$. The undecided verdict is the record's limit, not a failure of the drawing.

The drift's onset and the priors. The verdict depends on the drift's onset more than on any prior, through Chapter 10's ridge between the drift and coil 1; the onset and σ_q are Chapter 10's, not re-chosen here. The four drawings carry even prior odds, and the curvature's prior spread is chosen and varied.

Chapter 10's limits. The readings are independent across windows, the drift is linear in time, the commissioning test lies on the same path as the first month, and each window is in steady state with known inputs, as in Chapter 5.

The next chapter. Chapter 12 asks whether coil 2's steep months are one event, read as a step entry freed at each candidate window, and through it whether coil 1's verdict moves once coil 2's path is drawn as the record requires. The crossing forecast of Chapter 14 carries both of this chapter's drawings.

11.7 What to retain

Definition 11.1 (Accelerating degradation)**.** A degradation rate that may change is $\gamma_u(t_{\mathrm{y}}) = \gamma_u + \dot{\gamma}_u(t_{\mathrm{y}} - \bar{t}_{\mathrm{y}})$, with the curvature $\dot{\gamma}_u$ an entry of the state seen only by the trend rows, each through its lever $x_t = (t + \frac{1}{2})\Delta t - \bar{t}_{\mathrm{y}}$. Its MAP value is the least-squares slope of the fitted path's monthly rates against the lever, shrunk by its prior, and its spread has a floor $\sigma_q/(\Delta t\sqrt{\sum_t x_t^2})$ that only a longer record lowers. Whether the loss accelerates is two answers: the sign probability $\Pr(\dot{\gamma}_u < 0 \mid \mathbf{y}, M)$ within a drawing that frees the curvature, and the Bayes factor of that drawing against the constant rate, which falls without limit as the curvature's

prior widens. Over the family the two combine into the probabilities of an accelerating, constant and slowing loss, and the rate today is averaged over the drawings, with its variance split within and between them.

Three conclusions carry beyond the hall. An acceleration is read on the parameter, and a bend in a performance figure is a bend in everything it depends on, a drifting sensor first. The sign of a curvature and the need for one are different questions, and a record can be sure of the first and undecided on the second; the report gives both, and the rate today averaged over the drawings. And a curvature is read through the lever arm of the record, so that the most informative measurement is often the record continued, and the preposterior of the Bayes factor says how often waiting decides.

Table 11.19 lists the terms of the chapter with their expressions, units and the subsection that defines each.

Table 11.19: The terms of accelerating degradation and where each is defined. Units for the two-unit hall.

Term	Expression	Units	Meaning	Defined in
Rate that may change	$\gamma_u(t_y) = \gamma_u + \dot{\gamma}_u(t_y - \bar{t}_y)$, eq. (11.1)	yr^{-1}	the rate at time t_y	11.1.1
Curvature on the trend rows	eq. (11.4), lever x_t of eq. (11.2)	yr^{-2}	the rate's change per year	11.1.4
Curvature as the path's slope	eq. (11.6)	yr^{-2}	slope of the monthly rates, shrunk by the prior	11.1.4
Curvature's wander floor	eq. (11.7)	yr^{-2}	what no reading on the record removes	11.1.4
Questions and answers				
Sign probability	eq. (11.10)	–	accelerating, if the rate may change	11.3.1
Bayes factor against the constant rate	eqs. (7.7), (7.8)	–	whether the record needs the curvature	11.3.2
Lindley's effect	Table 11.7	–	the factor falls as the prior widens	11.3.2
Accelerating, constant, slowing	eq. (11.9)	–	the family's split	11.2.3
Rate today over the family	eq. (11.11), mixture (7.10)	%/yr	within and between the drawings	11.3.3
Dashboard's bend	eq. (11.12)	$1/\mathrm{yr}^2$	a derived quantity; a drift bends it	11.3.4
Drift and curvature	Section 11.3.5	K/yr, yr^{-2}	two explanations of one late rise	11.3.5
Power of an evidence test	Table 11.14	–	the share of fresh records that decide	11.4.5
Preposterior of the curvature	eq. (11.13), Table 11.17	yr^{-2}	later months read the curvature more	11.6.2
Preposterior of a Bayes factor	eq. (11.14)	–	how often waiting decides	11.6.3

11.8 Exercises

Exercise 11.1. *Formulation.* Write the question "is coil 1 losing more than 5 % a year today?" in the form of Table 11.2 and evaluate it under M_{12} and under M_0. (A prediction under a threshold, equation (2.12) with $f = -\gamma_1(3)$ and $f^* = -\ln 0.95$: under M_{12}, $\Pr(\gamma_1(3) < \ln 0.95 \mid \mathbf{y}) = \Phi\big((\ln 0.95 + 0.09719)/0.03770\big) = 0.8883$; under M_0, $\Phi\big((\ln 0.95 + 0.04240)/0.01463\big) = 0.2716$.)

Exercise 11.2. *Derivation.* With γ_u and $\dot{\gamma}_u$ free, show from the trend rows (11.4) that $\hat{\dot{\gamma}}_u$ is the shrunk least-squares slope of the path's monthly rates against x_t, and that a path known

exactly gives $\mathrm{sd}(\dot\gamma_u) = \sigma_q/(\Delta t\sqrt{\sum_t x_t^2})$ without the prior row. (The columns of γ_u and $\dot\gamma_u$ are $-\Delta t\,\mathbf{1}/\sigma_q$ and $-\Delta t\,\mathbf{x}_{\mathrm{lev}}/\sigma_q$, with $\mathbf{1}^{\mathsf{T}}\mathbf{x}_{\mathrm{lev}} = 0$; $\mathbf{J}^{\mathsf{T}}\mathbf{r} = 0$ in them gives equation (11.6). With the path fixed, the information on $\dot\gamma_u$ is $\Delta t^2\,\mathbf{x}_{\mathrm{lev}}^{\mathsf{T}}\mathbf{x}_{\mathrm{lev}}/\sigma_q^2$. $\sum_t x_t^2 = 26.9792\,\mathrm{yr}^2$, shrink factor 0.94933, floor $0.01155\,\mathrm{yr}^{-2}$.)

Exercise 11.3. *Derivation.* Show that the Savage–Dickey ratio (7.21) for M_1 is $\ln\mathcal{N}(0;\,0,\,\sigma_{\dot\gamma}^2) - \ln\mathcal{N}(0;\,\hat{\dot\gamma}_1,\,\varsigma_{\dot\gamma}^2)$ on the Laplace marginal, that it does not depend on the centering $\bar t_{\mathrm{y}}$, and evaluate it. (Moving the center to $\bar t'_{\mathrm{y}}$ replaces γ_1 by $\gamma'_1 = \gamma_1 + \dot\gamma_1(\bar t'_{\mathrm{y}} - \bar t_{\mathrm{y}})$ and leaves $\dot\gamma_1$ unchanged, a change of variables with unit determinant; γ_1 has no prior row, so the objective and the marginal of $\dot\gamma_1$ are the same. With $\hat{\dot\gamma}_1 = -0.02806$ and $\varsigma_{\dot\gamma} = 0.02002$: $2.0768 - 2.0101 = +0.0667$, against the Laplace evidence's $+0.0549$.)

Exercise 11.4. *Application.* Refit with $\sigma_{\dot\gamma} = 0.2\,\mathrm{yr}^{-2}$ and say why the Bayes factor falls while the sign probability rises. (The wider prior spreads its mass over rates the record does not show, so its density near the fitted curvature falls and the Occam factor of equation (7.8) grows by about the log of the ratio of the spreads, while within M_{12} the prior shrinks the curvature less toward zero: $\ln B(M_1 : M_0) = -1.0795$; $\hat{\dot\gamma}_1 = -0.04137$, $\Pr = 0.9606$.)

Exercise 11.5. *Application.* Refit with the rate constant to window 24 and changing after it, and compare the curvature and today's rate with the reported form. (The lever becomes $(t_{\mathrm{y}} - 2)_+$ integrated over each month: $\dot\gamma_1 = -0.05461$, standard deviation 0.04161, $\Pr = 0.9053$; $\ln B(M_1 : M_0) = +0.5921$; rate today $-8.830\,\%/\mathrm{yr}$.)

Chapter 12

Change-Point Detection

12.1 A step on the dashboard and a pump change in the work orders

The hall is three years past its handover, and today is December of year 3, window 36. Beside the condition record the site keeps its work-order history, and in it stands a chilled-water pump change in March of year 2, window 15. The asset manager wants to know whether that work did anything to the coils. The condition-monitoring dashboard has a step finder for this: it fits each coil's 37 health-index tiles, the raw-performance tiles of Chapter 10, with a straight line plus a step at each window, and flags the window whose step fits best. For coil 2 it flags a drop of 0.177 of the commissioned tile into window 14, February of year 2, with $F_{\mathrm{st}} = 17.56$ on $(1, 34)$ degrees of freedom and "$P = 0.0002$"; at March itself coil 2's step is -0.143. For coil 1 it flags a rise of 0.1656 into window 17, May of year 2, with $F_{\mathrm{st}} = 12.72$ and "$P = 0.0011$". The conclusion the chart invites is that coil 2 lost 18 % in February, a month before the pump change, so the pump work did not do it, and that coil 1 recovered 17 % in May, though nobody cleaned it.

The chart cannot say whether a step in a tile is a step in the coil. The tile mixes the load, the season, both coils and the sensor (Chapter 10), and window 14 carries the record's highest load on the meter, 139.82 kW, and window 17 its lowest, 112.81 kW. Nor can it say what its "P" means when the window was picked as the best of 34, or whether a step it finds is on the coil or on the sensor that reads it. The question of this chapter is

whether coil conductance stepped in March of year 2, read as the evidence for a step entry in the base fit at that window against none; what stepped, when, and by how much; and what the step does to Chapter 10's rates.

It is answered with the method of Chapter 5, on Chapter 10's stacked state, with Chapter 7's evidence over a family of drawings.

Jargon: change point, candidate window, candidate family of steps, named window, location prior, scan, look-elsewhere charge, false alarm of a scan, closed-form scan, influence of one window, time to decide

A *change point* (a step) is a window at which a parameter's path or a sensor's offset shifts by a fixed amount and stays shifted; a *candidate window* is a window at which a step is allowed to enter. The *candidate family of steps* holds one drawing for each part that could have stepped and each candidate window, kinds times windows. A *named window* is a candidate stated before the record is read, here from a work order. The *location prior* is the prior probability over the candidate windows. A *scan* scores every candidate of the family against the drawing without a step, and its *winner* is the candidate with the largest evidence. The *look-elsewhere charge* is the multiple-comparisons cost of a scan: what separates the evidence of its best candidate from the evidence for a step somewhere. A *false alarm of a scan* is a candidate whose evidence comes from the record's errors rather than from an event. The *closed-form scan* scores every candidate from the base fit alone, without refitting. The *influence of one window* is the change of a candidate's evidence when that window's readings are left out. The *time to decide* (detection delay) is the first window at which a verdict on a step holds.

Key idea

A change point is one more entry of the same fit: a step on one trend row of a coil's path, or an offset on one sensor from a window on. Each candidate, what stepped and when, is a drawing, and the evidence ranks them. Because a step on a trend row is a bias on one row, the base fit alone scores every candidate in closed form, by Chapter 7's formula for a bias on one reading. The evidence at a window named in advance answers "did it happen then"; "did anything happen" pays for every window it looked at.

12.1.1 A step as an entry of the path

Chapter 7 freed a bias on one reading of one window and compared the drawing that carries it with the drawing that does not. Chapter 10 tied each window's log-derate to the next by the trend rows (10.5), with the spread σ_q of one month's wander, and reported its largest trend residuals at coil 2's windows 13 to 16, untested. A step of size δ in coil u's path into window t_0 is a bias on the one trend row that crosses it,

$$r_{q,u,t_0-1}(\mathbf{z}) = \frac{g_{u,t_0} - g_{u,t_0-1} - \gamma_u \, \Delta t - \delta}{\sigma_q}, \tag{12.1}$$

which leaves every other row of the record as it was: the path moves by δ across that month and keeps its rate on either side. A sensor can step as well, an offset that appears in one window and stays in its readings from then on. Freed at each window, and on each part that could have stepped, either coil, both coils, or one of the five readings, the step is the diagnosis of Chapter 2, equation (2.15), over a family of candidates generated from the graph.

> **Principle 12.1.** *A window named before the record is read is one hypothesis; a window found by scanning is one of many, and the evidence for "something, somewhere" is the average over all of the candidates, not the best of them.*

12.1.2 The two-unit hall through its record (step 0)

The plant graph is Chapter 10's, cited and not re-derived (Figure 10.1): 37 copies of Chapter 5's window, each with its free load Q_t and the two log-derates $g_{1,t}$, $g_{2,t}$ under the five physical

rows (5.2) at the window's known inputs $T_s(t)$ and $Q_o(t)$; one trend row per coil between consecutive windows, linked to the trend entries γ_1 and γ_2; the shared recirculation fraction β_1; and the drift node $\dot{b}_{d,1}$ on unit 1's discharge reading from window 20, equation (10.2). The record is the whole of it, window 0 (commissioning) and windows 1 to 36, January of year 1 to December of year 3, five readings each. The modeling choices of step 0 are Chapter 5's paragraph, carried through Chapter 10: steady state within each window, the airflow pinned with no discharge control, the season's quantities known inputs, and one conductance per coil. The choices a step adds are named at step 2.

A candidate adds one node to this graph. A coil step attaches to one trend square of its coil, the square into the candidate window, and a step on both coils attaches one node to both squares into the same window. A reading step attaches to one sensor's reading edges from the candidate window to window 36. Each candidate is a separate drawing, and no drawing carries two steps. Figure 12.1 draws the graph reduced to what the candidates touch: the two coils' paths as strips of nodes with the trend squares between them, five of the 37 windows, and the step node of coil 2 on its square into window 15 (solid), with the competing step on unit 2's discharge readings from window 15 dashed.

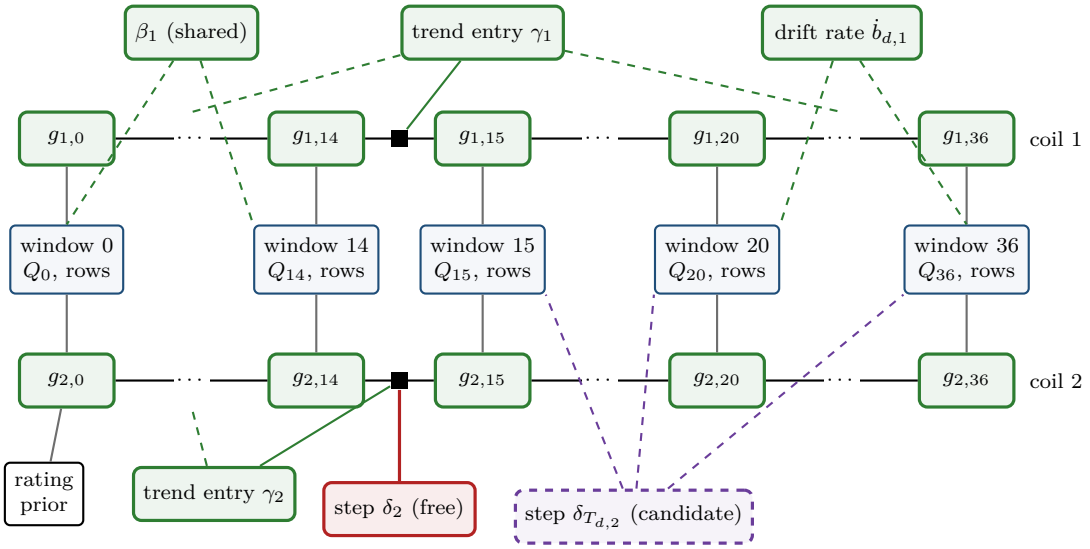


Figure 12.1: The record's graph reduced to what the step candidates touch, five of the 37 windows. Free (green): each window's log-derates $g_{u,t}$ and load Q_t (inside each window's block, with Chapter 5's physical rows and five readings), the trend entries γ_1, γ_2, the shared fraction β_1 and the drift rate $\dot{b}_{d,1}$ on unit 1's discharge readings from window 20. Black squares: the trend rows between consecutive windows, the runs between the drawn windows marked by dots; the trend entries link to every square of their coil (dashed toward the runs). Red: the step δ_2 on coil 2's square into window 15, the drawing $M_{2,15}$. Dashed purple: the competing drawing $M_{T_{d,2},15}$, a step on unit 2's discharge readings from window 15 to 36. Known inputs of each window, pinned and not drawn: $T_s(t)$ and $Q_o(t)$, the ratings and the capacity rates. Observed: the five readings of every window. Priors: the rating prior on window 0's coils (drawn for coil 2), the survey prior on β_1, the drift prior, and the step's prior on δ_2.

12.1.3 The asset manager's questions (step 1)

The asset manager asks six questions of the record, and step 1 writes each as a probability over the drawings of the family or over the state of one of them, with the unknown on the left of the bar and the 185 readings and the priors on the right. Table 12.1 lists them. The first column is the question in the asset manager's words, the first of them the drill question; the second

the general form of its category from Chapter 2; the third its mathematical form for the hall, with $B_{j,s}$ the Bayes factor of a step of kind j into window s against no step, $\bar{B}_s$ its mean over the eight kinds at window s and $\bar{B}$ its mean over all 288 candidates; the fourth its category; the fifth the computation of Section 2.3.6 that answers it; and the last the subsection where it is answered. Whether coil 1's rate changes is Chapter 11's question, how much of a summer tile is the season Chapter 13's, and when the margin closes Chapter 14's, in Volume II. Which measurement would decide the step asks nothing new of the record; it is step 7.

Table 12.1: The questions asked of the two-unit hall's condition record at window 36 about a step in March of year 2 (window 15). $\mathbf{y}$ is the 185 readings of windows 0–36 and the prior rows; M_{none} the base fit and $M_{j,s}$ the drawing with one step of kind j into window s; $h_{2,36}$ coil 2's health index today.

Asset manager's question	General form	Mathematical form	Category	Computation	Answered in
Did something happen in March?	(2.15) at the named window; (2.14)	$\Pr(t_0 = 15 \mid \mathbf{y}) = \bar{B}_{15}/(1+\bar{B}_{15})$, $\bar{B}_{15} = \frac{1}{8}\sum_j B_{j,15}$	diagnosis	evidence	12.3.1
What stepped: a coil, or a sensor reading it?	(2.15)	$\Pr(M_{j,15} \mid \mathbf{y}, t_0 = 15)$ over the eight kinds j	diagnosis	evidence	12.3.2
Was it March, or the month either side?	(2.15) over windows	$\Pr(t_0 = s \mid \mathbf{y}, \text{coil 2})$, $s \in \{14, 15, 16\}$	diagnosis	evidence	12.3.3
How large was the step, and how much of coil 2's loss is it?	(2.9), (2.11), (2.10)	$p(\delta_2 \mid \mathbf{y}, M_{2,15})$, as $100(\mathrm{e}^{\delta_2} - 1)\,\%$ of conductance; $\delta_2 / \ln h_{2,36}$	estimation	posterior	12.3.4
With the step allowed for, how fast is each coil losing conductance, and which is faster?	(2.21) (line 1), (2.12); (7.10)	$p(\gamma_u \mid \mathbf{y})$ mixed over the named-window family; $\Pr(\gamma_2 < \gamma_1 \mid \mathbf{y})$	estimation	trend	12.3.5
Did a step happen anywhere in the three years?	(2.15) over the whole record	$\Pr(\text{a step} \mid \mathbf{y})$ over windows 1–36 $= \bar{B}/(1+\bar{B})$; the ranking of the 288 candidates	diagnosis	evidence	12.3.6

12.1.4 The evidence as rows (step 2)

Step 2 attaches the record to the graph as Chapter 10's rows and adds, for each candidate, one entry and one prior row. The base drawing M_{none} is Chapter 10's base fit, the drawing Chapter 11 calls M_0 (written M_{none} here, the drawing with no step): the stacked state (10.3) of 115 entries, 261 rows (the 185 readings, the rating priors on window 0's coils, the survey prior on β_1, the 72 trend rows (10.5) at $\sigma_q = 0.005$ and the drift prior), and 146 degrees of freedom. A step drawing $M_{j,s}$ adds a step δ_j of kind j into window $s = 1, \ldots, 36$, for the eight kinds of Table 12.2. Each row of the table is a kind; the columns give where its step enters the rows, its prior, and its column $\mathbf{a}_{j,s}$ in the normalized rows, the derivative of the residual vector with respect to the step.

The state and the rows of a step drawing. The state of $M_{j,s}$ is the base state with the step appended, and its rows are the base rows with the step's column added and its prior row

Table 12.2: The candidate family: one step of kind j into window s, for eight kinds and 36 windows, 288 step drawings beside M_{none}. Sensor standard deviations as in Chapter 5: 0.3 K on each temperature, 3.0 kW on the meter.

Kind j	Where the step enters	Prior on δ_j	Column $\mathbf{a}_{j,s}$ (normalized rows)
coil 1 ($j = 1$), coil 2 ($j = 2$)	the coil's trend row into s, equation (12.1)	$\mathcal{N}(0,\ 0.15^2)$ [–]	one nonzero, $-1/\sigma_q = -200$
both coils ($j = 12$)	both trend rows into s, one δ	$\mathcal{N}(0,\ 0.15^2)$ [–]	two nonzeros, -200 each
return air (T_r), unit 1's discharge ($T_{d,1}$), unit 2's discharge ($T_{d,2}$), rack 1's inlet (T_1)	the sensor's reading of every window $t \geq s$	$\mathcal{N}(0,\ 1.0^2)$ [K]	$36 - s + 1$ nonzeros, $1/0.3 = 3.3333$ per K
load meter (Q)	the meter's reading of every window $t \geq s$	$\mathcal{N}(0,\ 10^2)$ [kW]	$36 - s + 1$ nonzeros, 0.3333 per kW

below them,

$$\mathbf{z}_{j,s} = \begin{bmatrix} \mathbf{z}_{\text{none}} \\ \delta_j \end{bmatrix}, \qquad \mathbf{r}_{j,s}(\mathbf{z}_{j,s}) = \begin{bmatrix} \mathbf{r}_{\text{none}}(\mathbf{z}_{\text{none}}) + \mathbf{a}_{j,s}\,\delta_j \\ \delta_j/\sigma_\delta \end{bmatrix}, \qquad \mathcal{L}_{j,s} = \tfrac{1}{2}\,\mathbf{r}_{j,s}^{\mathsf{T}}\mathbf{r}_{j,s}, \tag{12.2}$$

with $\mathbf{z}_{\text{none}}$ the stacked state (10.3), $\mathbf{r}_{\text{none}}$ its 261 rows and σ_δ the step's prior spread. A step drawing has 116 entries and 262 rows, so 146 degrees of freedom. The step enters every row it touches linearly: a coil step through the trend row (12.1), a reading step as an offset $+\delta_j$ on the predicted reading of every window from s on. The column is fixed by the graph and does not depend on the state.

Which row acts on which entry. The Jacobian of a step drawing is the base Jacobian, whose block-banded pattern is equation (10.6), with one column and one row added. For the drawing $M_{2,15}$, with k the index of coil 2's trend row into window 15, the new column has one nonzero in that trend row and one in the prior row, and the block of those two rows on the columns $(g_{2,14}, g_{2,15}, \gamma_2, \delta_2)$ is

$$\mathbf{J}_{2,15} = \begin{bmatrix} \mathbf{J}_{\text{none}} & \mathbf{a}_{2,15} \\ \mathbf{0}^{\mathsf{T}} & 1/\sigma_\delta \end{bmatrix}, \quad \mathbf{a}_{2,15} = -\frac{\mathbf{e}_k}{\sigma_q}, \quad \begin{bmatrix} * & * & * & * \\ 0 & 0 & 0 & * \end{bmatrix} \begin{matrix} \text{coil 2's trend row into 15} \\ \text{prior row on } \delta_2 \end{matrix},$$
$$\mathbf{g}_\delta^{\mathsf{T}} = \mathbf{e}_\delta^{\mathsf{T}}, \qquad \mathbf{g}_{\gamma_u}^{\mathsf{T}} = \mathbf{e}_{\gamma_u}^{\mathsf{T}}, \qquad \mathbf{g}_{D_\gamma}^{\mathsf{T}} = \mathbf{e}_{\gamma_1}^{\mathsf{T}} - \mathbf{e}_{\gamma_2}^{\mathsf{T}}, \qquad \mathbf{g}_{\text{split}}^{\mathsf{T}} = 36\,\Delta t\,\mathbf{e}_{\gamma_2}^{\mathsf{T}} + \mathbf{e}_\delta^{\mathsf{T}}, \tag{12.3}$$

with $\mathbf{e}_i$ the unit vector of entry i. The second line is not a row: it gives the gradients of the questions' quantities, the step, each rate, the difference $D_\gamma = \gamma_1 - \gamma_2$ of the two rates, and the sum $36\,\Delta t\,\gamma_2 + \delta_2$ of coil 2's two parts of loss, its rate over the record and its step. With γ_2 free the residuals of coil 2's 36 trend rows sum to zero at the MAP state, Chapter 10's telescoped sum, so that sum equals coil 2's loss since commissioning there, a MAP identity:

$$\widehat{\ln h}_{2,36} = \hat{g}_{2,36} - \hat{g}_{2,0} = 36\,\Delta t\,\hat{\gamma}_2 + \hat{\delta}_2. \tag{12.4}$$

Away from the MAP the two differ by the summed wander of the 36 rows, so the sum is not $\ln h_{2,36}$ as a quantity of the state, and its spread is not the index's. Two facts of the graph are read here, before any number. A coil step is seen only through one trend row: no reading row has a nonzero in its column, so the step is read from the path the readings fix, as the rate is (Chapter 10). And the step enters the residuals linearly, $\mathbf{r}_{j,s} = \mathbf{r}_{\text{none}} + \mathbf{a}_{j,s}\,\delta_j$, which is what lets the base fit alone score every candidate in closed form (Section 12.2.2).

Modeling choices at the rows. Six choices make the family. *At most one step in the record*: this buys a family of 289 drawings, the base and 288 steps, and it fails if two events happened; Section 12.4.3 frees a second step beside the winner. *A step is permanent*, a level shift of the path or of the sensor's offset: this buys one entry, and it fails for an event later reversed, which would read as two steps of opposite sign. *The step enters between two monthly windows*: this buys the record's own resolution, and it cannot place the event within the month. *The step's prior width*: $\sigma_\delta = 0.15$ on a coil, the rating prior's width, with either sign, because a cleaning or a valve reopened raises conductance; 1.0 K on a temperature sensor, Chapter 7's prior on an offset; 10 kW on the meter. The evidence depends on it, the dependence known as Lindley's effect (Exercise 12.5). *The prior over the family*: even prior odds of a step against none, equal weight on the eight kinds, and over the windows a location prior that is the named window for the first three questions and uniform over windows 1–36 for the last,

$$\Pr(M_{\text{none}}) = \tfrac{1}{2}, \qquad \Pr(M_{j,s}) = \tfrac{1}{2} \cdot \tfrac{1}{8} \cdot \Pi_s, \qquad \Pi_s = \begin{cases} \mathbf{1}[s = 15] & \text{named window,} \\ 1/36 & \text{scan.} \end{cases} \tag{12.5}$$

The base fit's choices are kept: $\sigma_q = 0.005$, the drift linear from window 20 with Chapter 7's prior, and a flat prior on each trend entry, as Chapter 10 decided; Section 12.4.3 varies each.

12.2 The posterior of the state (step 3)

Every answer of this chapter is read from the posteriors of the drawings of step 2: the evidence of each, and the state of the drawing the named window selects.

12.2.1 The rows and the normal equations of a step drawing

The derivation form is Chapter 5's, on Chapter 10's stacked state with one entry more. The state and the residual vector of a step drawing are equation (12.2); its Jacobian, 262×116, is the pattern (12.3), evaluated at each iterate because the reading rows are nonlinear in the log-derates; the MAP state solves the normal equations (5.9) by Gauss–Newton, and the Laplace covariance is $\mathbf{\Sigma}_{j,s} = (\mathbf{J}_{j,s}^{\mathsf{T}}\mathbf{J}_{j,s})^{-1}$ at the minimum, equation (5.10). With every row normalized the information is $\mathbf{J}^{\mathsf{T}}\mathbf{J}$; with unnormalized rows and their weights it would be $\mathbf{J}^{\mathsf{T}}\mathbf{W}\mathbf{J}$. The step's rows are linear in the state and add nothing to the curvature term $\mathbf{C}$ of equation (4.44). The evidence of each drawing is its Laplace evidence, equation (7.7), and the Bayes factor of a step against none is

$$B_{j,s} = \frac{p(\mathbf{y} \mid M_{j,s})}{p(\mathbf{y} \mid M_{\text{none}})}, \qquad \ln B_{j,s} = \left[\mathcal{L}_{\text{none}}(\hat{\mathbf{z}}_{\text{none}}) - \mathcal{L}_{j,s}(\hat{\mathbf{z}}_{j,s})\right] - \frac{1}{2} \ln \frac{\sigma_\delta^2 \det \mathbf{\Lambda}_{j,s}}{\det \mathbf{\Lambda}_{\text{none}}}, \tag{12.6}$$

the gain in fit less the Occam factor, equation (7.8) between the nested pair. Every one of the 288 candidates is refitted this way, by Gauss–Newton from the base MAP state with $\delta_j = 0$.

12.2.2 The closed form of a step

The chapter owes $\ln B_{j,s}$ for 288 candidates, and the base fit alone gives each of them in closed form. The log Bayes factor is the logarithm of the ratio of two integrals of the same rows,

$$\ln B_{j,s} = \ln \iint \mathrm{e}^{-\mathcal{L}_{j,s}(\mathbf{z},\delta)} \, \mathrm{d}\mathbf{z} \, \mathrm{d}\delta - \ln(\sqrt{2\pi}\,\sigma_\delta) - \ln \int \mathrm{e}^{-\mathcal{L}_{\text{none}}(\mathbf{z})} \, \mathrm{d}\mathbf{z},$$

the second term the step's prior row's normalizing constant, the only constant the two drawings do not share, and both integrals over the $n_{\mathrm{z}} = 115$ entries of the base state. The rows of the step drawing are the base rows plus the column times the step, equation (12.2),

$$\mathcal{L}_{j,s}(\mathbf{z}, \delta) = \tfrac{1}{2} \left\| \mathbf{r}_{\text{none}}(\mathbf{z}) + \mathbf{a}\,\delta \right\|^2 + \frac{\delta^2}{2\sigma_\delta^2},$$

with $\mathbf{a} = \mathbf{a}_{j,s}$. Linearize the base rows about the base MAP state, $\mathbf{r}_{\text{none}}(\mathbf{z}) \approx \hat{\mathbf{r}} + \mathbf{J}\,(\mathbf{z} - \hat{\mathbf{z}}_{\text{none}})$, with $\mathbf{J}$ the base Jacobian, $\boldsymbol{\Sigma}$ its covariance and $\hat{\mathbf{r}}$ the base residual vector, for which $\mathbf{J}^{\mathsf{T}}\hat{\mathbf{r}} = \mathbf{0}$. The integral over the state at fixed δ is Gaussian, and it leaves the part of $\hat{\mathbf{r}} + \mathbf{a}\,\delta$ that the columns of $\mathbf{J}$ cannot absorb, through the hat matrix $\mathbf{K} = \mathbf{J}\boldsymbol{\Sigma}\mathbf{J}^{\mathsf{T}}$ of equation (7.18),

$$\int \mathrm{e}^{-\frac{1}{2}\|\hat{\mathbf{r}}+\mathbf{a}\delta+\mathbf{J}(\mathbf{z}-\hat{\mathbf{z}}_{\text{none}})\|^2}\,\mathrm{d}\mathbf{z} = (2\pi)^{n_z/2}\det(\mathbf{J}^{\mathsf{T}}\mathbf{J})^{-1/2}\,\mathrm{e}^{-\frac{1}{2}(\hat{\mathbf{r}}+\mathbf{a}\delta)^{\mathsf{T}}(\mathbf{I}-\mathbf{K})(\hat{\mathbf{r}}+\mathbf{a}\delta)}.$$

Because $\mathbf{K}\hat{\mathbf{r}} = \mathbf{J}\boldsymbol{\Sigma}\mathbf{J}^{\mathsf{T}}\hat{\mathbf{r}} = \mathbf{0}$, the exponent is $-\frac{1}{2}[\hat{\mathbf{r}}^{\mathsf{T}}\hat{\mathbf{r}} + 2\delta\,\mathbf{a}^{\mathsf{T}}\hat{\mathbf{r}} + \delta^2\,\mathbf{a}^{\mathsf{T}}(\mathbf{I}-\mathbf{K})\mathbf{a}]$. The first term and the factor before the exponential are the base drawing's own integral, and they cancel in the ratio, which leaves one Gaussian integral over the step,

$$B_{j,s} = \int \exp\Big[-\delta\,\mathbf{a}^{\mathsf{T}}\hat{\mathbf{r}} - \tfrac{1}{2}\,\delta^2\,\mathbf{a}^{\mathsf{T}}(\mathbf{I}-\mathbf{K})\mathbf{a}\Big]\,\frac{\mathrm{e}^{-\delta^2/(2\sigma_\delta^2)}}{\sqrt{2\pi}\,\sigma_\delta}\,\mathrm{d}\delta.$$

The data part has the precision $\mathbf{a}^{\mathsf{T}}(\mathbf{I}-\mathbf{K})\mathbf{a}$ and the score $\mathbf{a}^{\mathsf{T}}\hat{\mathbf{r}}$, and the prior adds $1/\sigma_\delta^2$. Completing the square, with the ratio of the prior's variance to the variance of the step from the other rows,

$$\begin{gathered}\psi = \sigma_\delta^2\,\mathbf{a}^{\mathsf{T}}(\mathbf{I}-\mathbf{K})\mathbf{a}, \qquad \ln B_{j,s} = \frac{1}{2}\left[\frac{\sigma_\delta^2\,(\mathbf{a}^{\mathsf{T}}\hat{\mathbf{r}})^2}{1+\psi} - \ln(1+\psi)\right], \\ \hat{\delta} = -\frac{\sigma_\delta^2\,\mathbf{a}^{\mathsf{T}}\hat{\mathbf{r}}}{1+\psi}, \qquad \mathrm{sd}(\delta \mid \mathbf{y}) = \frac{\sigma_\delta}{\sqrt{1+\psi}}.\end{gathered} \tag{12.7}$$

For a bias on one reading, $\mathbf{a} = \mathbf{e}_k/\sigma_k$, this is Chapter 7's closed form (7.22). For a coil step, $\mathbf{a} = -\mathbf{e}_k/\sigma_q$ on the trend row k into window s, so $\mathbf{a}^{\mathsf{T}}(\mathbf{I}-\mathbf{K})\mathbf{a} = (1-K_{kk})/\sigma_q^2$ and $\psi = \sigma_\delta^2(1-K_{kk})/\sigma_q^2$, and the first term is $\psi/(1+\psi)$ times the square of the trend row's normalized residual (7.20): the step test of a coil is the trend row's own residual test. For a reading step the column spreads over every window from s on, and the test is that of the sum of those readings' departures.

Coil 2's trend row into window 15. In the base fit this row has the residual -0.6490, the leverage 0.9557 and so the redundancy number 0.0443, and the normalized residual -3.0819. Then $\mathbf{a}^{\mathsf{T}}\hat{\mathbf{r}} = 129.8064$ and $\mathbf{a}^{\mathsf{T}}(\mathbf{I}-\mathbf{K})\mathbf{a} = 1773.9738$, so $\psi = 0.15^2 \times 1773.9738 = 39.9144$, and by equation (12.7)

$$\ln B_{2,15} = \tfrac{1}{2}\,(9.2661 - 3.7115) = +2.7773, \qquad \hat{\delta}_2 = -0.07138, \qquad \mathrm{sd}(\delta_2 \mid \mathbf{y}) = 0.02345.$$

The refit of the drawing gives $\ln B_{2,15} = +2.7648$. The trend row's raw residual, -0.65, understates its departure: the wander lets the neighboring windows absorb most of one row's departure, so the row's redundancy number is small, and its normalized residual, -3.08, is the measure of what the other rows cannot explain. Over the 288 candidates the closed form departs from the refits by at most 0.1577, and by at most 0.0489 among the candidates with $\ln B > 0$; the gap is the reading rows' nonlinearity, which the refitted state follows and the linearization about the base MAP does not.

12.2.3 The fit of the step at the named window

The iterations. Table 12.3 runs the normal equations of $M_{2,15}$ from the base MAP state with $\delta_2 = 0$. Each row is an iteration, with the objective and the norm of the gradient $\mathbf{J}^{\mathsf{T}}\mathbf{r}$. The objective falls from 77.487758, the base fit's own minimum, to 72.859760 in one step, and the gradient is below 10^{-4} from the fourth iteration, falling to the Jacobian's round-off floor.

Table 12.3: Gauss–Newton iterations of the step drawing $M_{2,15}$, from the base MAP state with the step at zero.

Iteration l [–]	Objective $\mathcal{L}(\mathbf{z}^{(l)})$ [–]	$\lVert\mathbf{J}^{\mathsf{T}}\mathbf{r}\rVert$ [–]
0	77.487758	1.298×10^{2}
1	72.859760	2.871
2	72.858277	5.423×10^{-2}
3	72.858277	1.064×10^{-3}
4 on	72.858277	$< 10^{-4}$

The Jacobian of the step's rows. At the MAP state the line of coil 2's trend row into window 15 is $(-200,\ +200,\ -16.6667,\ -200)$ on $(g_{2,14}, g_{2,15}, \gamma_2, \delta_2)$, with $-16.6667 = -\Delta t/\sigma_q$, and the step's prior row is $1/0.15 = 6.6667$ in the column of δ_2. The step and the rate meet in that one row, and nowhere else.

The residuals. At the MAP state twice the objective is $2\mathcal{L} = 145.7166$ on 146 degrees of freedom, against the base fit's 154.9755, and $\Pr(\chi^2_{146} > 145.7) = 0.491$. The gain in fit, half the drop, is 4.6295, and the Occam factor of equation (12.6) is 1.8647, so the refit's $\ln B_{2,15} = 4.6295 - 1.8647 = +2.7648$.

The MAP state. Table 12.4 gives the entries the answers use, with the base fit's beside them. Each row is an entry; the columns give its MAP value and posterior spread under $M_{2,15}$ and under M_{none}. Coil 2's path now crosses window 15 from -0.06756 to -0.14164, against the base fit's -0.09891 to -0.10736, and its rate falls by half, from -0.06253 to $-0.03139\,\text{yr}^{-1}$.

Table 12.4: The MAP state of the step drawing $M_{2,15}$ and of the base fit M_{none}, with the posterior spreads, for the entries the answers use.

Entry [unit]	$M_{2,15}$, MAP [entry's]	$\sqrt{\Sigma_{ii}}$ [entry's]	M_{none}, MAP [entry's]	$\sqrt{\Sigma_{ii}}$ [entry's]
$g_{2,0}$ [–]	−0.03303	0.01586	−0.02448	0.01558
$g_{2,14}$ [–]	−0.06756	0.01407	−0.09891	0.00977
$g_{2,15}$ [–]	−0.14164	0.01506	−0.10736	0.00976
$g_{2,36}$ [–]	−0.19858	0.01736	−0.21208	0.01690
γ_1 [yr^{-1}]	−0.04398	0.01465	−0.04240	0.01463
γ_2 [yr^{-1}]	−0.03139	0.01673	−0.06253	0.01323
δ_2 [–]	−0.07139	0.02349	–	–
$\dot{b}_{d,1}$ [K/yr]	0.66823	0.16809	0.68371	0.16812

The covariance. The block of $\mathbf{\Sigma}_{2,15}$ on these eight entries, in the order of Table 12.4, scaled by 10^4, is

$$10^4\,\mathbf{\Sigma} = \begin{bmatrix} 2.5139 & 0.056452 & 0.62886 & -0.85615 & 0.59113 & -1.3514 & 0.68427 & 5.0836 \\ 0.056452 & 1.9807 & -0.32438 & 0.56691 & -0.040262 & 0.96443 & -2.3828 & -1.6699 \\ 0.62886 & -0.32438 & 2.2689 & -0.33140 & 0.036593 & -1.2173 & 2.6917 & -1.1466 \\ -0.85615 & 0.56691 & -0.33140 & 3.0142 & -0.99395 & 1.6346 & -1.0334 & -13.960 \\ 0.59113 & -0.040262 & 0.036593 & -0.99395 & 2.1457 & -0.56976 & 0.12420 & 12.769 \\ -1.3514 & 0.96443 & -1.2173 & 1.6346 & -0.56976 & 2.7995 & -2.4124 & -6.7083 \\ 0.68427 & -2.3828 & 2.6917 & -1.0334 & 0.12420 & -2.4124 & 5.5195 & 1.0811 \\ 5.0836 & -1.6699 & -1.1466 & -13.960 & 12.769 & -6.7083 & 1.0811 & 282.55 \end{bmatrix},$$

the entries of the last row and column in K/yr and those of γ_u in yr^{-1}. The correlations read by sign, and the rows fix each sign. The step is correlated with $g_{2,14}$ by -0.7207 and with $g_{2,15}$ by $+0.7606$: it is the difference across its row, so a higher window 15 or a lower window 14 is a smaller loss at the step. The step is correlated with the rate γ_2 by -0.6137: the step and the rate share coil 2's loss between windows 0 and 36, so a larger loss at the step, δ_2 more negative, leaves a slower loss to the rate. This is why freeing the step moves the rate. The rate is correlated with the ends of the path by -0.5094 ($g_{2,0}$) and $+0.5627$ ($g_{2,36}$), Chapter 10's signs, and the two rates by -0.2325, because the return air sees the sum of the two conductances. The drift rate is correlated with γ_1 by $+0.5186$, Chapter 7's ridge along unit 1's discharge reading, and with the step by only $+0.0274$: unit 1's sensor does not see coil 2's step.

The full posterior. Beside the Laplace normal stands the full posterior of each drawing by reweighted draws, the route of Section 5.2.4, with the drawing's Laplace normal as the proposal, and the same draws estimate each evidence by equation (7.12). Section 12.4.1 states their Monte Carlo errors.

12.3 The answers (step 4)

Every answer is read from the Laplace evidences of the refits of Section 12.2.1 and from the pair $(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ of the drawing $M_{2,15}$ of Section 12.2.3, and that value is the one reported. The full posterior by reweighted draws, the Savage–Dickey ratio and the closed form (12.7) stand beside it. A step on a coil is reported as $100(\mathrm{e}^{\delta} - 1)\,\%$ of conductance, and a rate as $100(\mathrm{e}^{\gamma} - 1)\,\%/\mathrm{yr}$, Chapter 10's form (10.1); both maps are monotone, so quantiles pass through them.

12.3.1 A step in March

Asset manager's question. Did something happen in March? The question names its window before the record is read, because the work-order history names the pump change of March of year 2, window 15.

General form. A diagnosis is a hypothesis comparison run over a family of candidates generated from the plant graph, equation (2.15),

$$\Pr(M_j \mid \mathbf{y}) \propto p(\mathbf{y} \mid M_j)\,\Pr(M_j),$$

each term set against a reference drawing by its Bayes factor, equation (2.14), $B_{AB} = p(\mathbf{y} \mid M_A)/p(\mathbf{y} \mid M_B)$.

Specialized form. The question is equation (2.15) over the family M_{none} and $M_{j,15}$, j over the eight kinds, with the prior probabilities (12.5) at the named window and each term the Bayes factor (2.14) with $M_A = M_{j,15}$ and $M_B = M_{\text{none}}$, the $B_{j,15}$ of equation (12.6). Dividing every term by $p(\mathbf{y} \mid M_{\text{none}})\,\Pr(M_{\text{none}})$,

$$\begin{aligned} \Pr(t_0 = 15 \mid \mathbf{y}) = \sum_j \Pr(M_{j,15} \mid \mathbf{y}) = \frac{\bar{B}_{15}}{1 + \bar{B}_{15}}, \qquad \bar{B}_{15} = \frac{1}{8}\sum_j B_{j,15}, \\ \Pr(M_{j,15} \mid \mathbf{y}) = \frac{B_{j,15}/8}{1 + \bar{B}_{15}}. \end{aligned} \tag{12.8}$$

One choice enters here: the location prior is a point mass at window 15, because the work order names March before the record is read. This is what lets the evidence at window 15 be read without paying for the other 35 windows, and it fails if the month was found in the record rather than named outside it; the scan of Section 12.3.6 is the reading for that case. A verdict at the probability 0.95 needs $\bar{B}_{15} \geq 19$.

Building blocks. The Laplace evidences of the nine drawings at window 15 and their log Bayes factors (12.6) (Section 12.2.1); for each step drawing its MAP step and its spread.

The evidence at window 15. Table 12.5 gives the eight kinds at the named window. Each row is a kind; the columns give its log Bayes factor, the Bayes factor, the step with its posterior spread, and the kind's share of window 15's evidence, $B_{j,15} / \sum_{j'} B_{j',15}$.

Table 12.5: The eight step drawings at the named window, window 15 (Mar Y2), by their refits. A coil step is in log-derate units; a sensor step in its reading's units.

Kind at window 15	$\ln B_{j,15}$ [–]	$B_{j,15}$ [–]	Step (sd) [–, K or kW]	Share [–]
coil 1	−1.8680	0.1544	+0.0037 (0.0229)	0.0077
coil 2	+2.7648	15.8758	−0.0714 (0.0235)	0.7873
both coils	−0.9438	0.3891	−0.0234 (0.0138)	0.0193
return-air sensor	−2.2528	0.1051	−0.0483 (0.0950) K	0.0052
unit 1's discharge sensor	−1.3432	0.2610	−0.1624 (0.1396) K	0.0129
unit 2's discharge sensor	+1.0881	2.9687	+0.3759 (0.1580) K	0.1472
load meter	−1.6435	0.1933	−0.5059 (1.8153) kW	0.0096
rack 1's inlet sensor	−1.5277	0.2170	+0.1249 (0.1186) K	0.0108

By equation (12.8) the mean Bayes factor at window 15 is $\bar{B}_{15} = 2.5206$ ($\ln \bar{B}_{15} = +0.9245$), and

$$\Pr(t_0 = 15 \mid \mathbf{y}) = \frac{2.5206}{3.5206} = 0.716, \qquad \Pr(M_{2,15} \mid \mathbf{y}) = 0.5637.$$

A step in March is probable, not decided: the verdict at 0.95 needs a mean Bayes factor of 19, and window 15's is 2.52. Coil 2 alone, the drawing $M_{2,15}$ against M_{none}, has $B = 15.88$ and would give $\Pr = 0.941$; the coils' kinds alone, coil 1, coil 2 and both, give a mean Bayes factor of 5.4731, $\Pr = 0.8455$, with coil 2's share 0.967; and "March" read as February to April of year 2, windows 14–16 with every kind, gives a mean Bayes factor of 1.7076 and $\Pr = 0.631$. The named window carries a Bayes factor of 16 on coil 2, and still the record does not decide that something happened, because it must also say what: seven other parts could have stepped in March, the prior gives each of them the same weight as coil 2, and their evidence is weak.

The log Bayes factors of five kinds are plotted against the window in Figure 12.2. In the upper panel the horizontal axis is the candidate window s and the vertical axis $\ln B_{j,s}$; the lines are coil 1, coil 2, both coils and the two discharge sensors; the dashed horizontal lines are zero and $\ln 19$; the gray band marks window 15, the pump change of the work order, and the dotted marker the stated truth's step (Section 12.4.5). The lower panel gives the base fit's normalized residual of coil 2's trend row into each window. Coil 2's line rises above zero at windows 14 to 16 and at windows 2 and 3, and nowhere else; the rows into windows 14 to 16 carry the base fit's most negative trend residuals. The both-coils line peaks at window 2, above every other candidate, and unit 1's sensor at window 24.

12.3.2 The part that stepped

Asset manager's question. What stepped: a coil, or a sensor reading it? The pump change could have moved a coil's water side, or the readings could carry a shift that only looks like one.

General form. The diagnosis (2.15), the ranking of the candidates by their posterior probabilities.

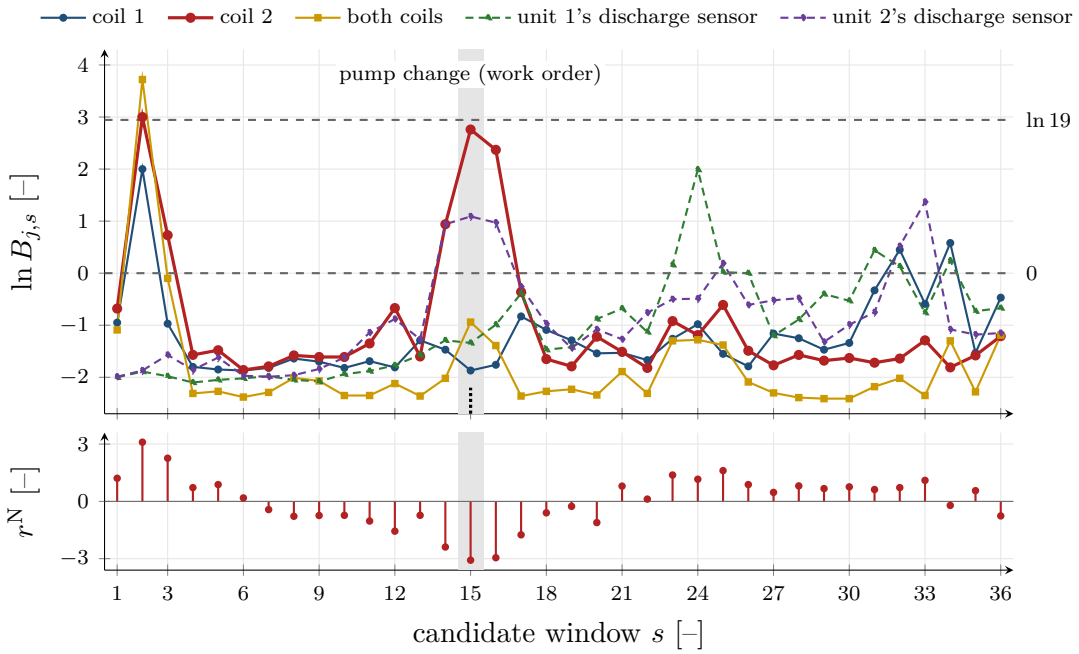


Figure 12.2: The scan. Upper panel: the log Bayes factor of a step of each of five kinds into window s against no step, by the refits, equation (12.6). Dashed: zero and ln 19, the level a lone drawing needs for a probability of 0.95. Gray band: window 15 (Mar Y2), the pump change of the work order; dotted tick at its foot: the stated truth's step on coil 2, a check marker (Section 12.4.5). Lower panel: the base fit's normalized residual (7.20) of coil 2's trend row into each window, the statistic of the closed form (12.7).

Specialized form. The question is equation (2.15) over the eight drawings $M_{j,15}$, given that a step entered at window 15, with equal prior weights on the kinds,

$$\Pr(M_{j,15} \mid \mathbf{y}, t_0 = 15) = \frac{B_{j,15}}{\sum_{j'} B_{j',15}}, \tag{12.9}$$

the share column of Table 12.5. Why one kind beats another is read row by row: the gain in fit of equation (12.6) is half the drop in the squared normalized residuals, and it splits over the readings and the trend rows.

Building blocks. The Bayes factors $B_{j,15}$ (Table 12.5); each drawing's residual vector at its MAP state against the base fit's, row by row.

The kind of the step. By equation (12.9), coil 2 holds 0.787 and unit 2's discharge sensor 0.147; both coils hold 0.019, unit 1's discharge sensor 0.013, and the rest less than 0.011 each. Coil 2's step gains 4.1455 on unit 2's discharge readings, 0.5148 on rack 1's inlet and 0.4668 on the trend rows, and loses 0.3266 on the return air; the step on unit 2's sensor gains 3.7956 on its own readings and loses 1.0713 on unit 1's discharge and 0.3895 on rack 1's inlet. The predicted readings say why. Averaged over windows 15 to 36, coil 2's step moves its own discharge by $+0.0446\,\mathrm{K}$, the return air by $+0.0111\,\mathrm{K}$ and rack 1's inlet by $+0.0203\,\mathrm{K}$, and unit 1's discharge by only $-0.0034\,\mathrm{K}$: a coil that loses conductance warms the return air and, through it, the rack's inlet, and the rack's readings favor that. The step on unit 2's sensor moves the return air by $-0.0483\,\mathrm{K}$, unit 1's discharge by $+0.0311\,\mathrm{K}$ and rack 1's inlet by $-0.0136\,\mathrm{K}$, as the refitted coils absorb the offset, and unit 1's discharge and rack 1's inlet do not support those moves. The readings that do not pass through unit 2's sensor decide between the two, Chapter 7's principle met again for a step.

Unit 1's sensor against coil 1. The family carries both kinds at every window, with the drift held at window 20. Unit 1's sensor candidates peak at window 24, with $\ln B = +1.9919$ and a step of -0.6543 K (sd 0.2513), and with that step the drift rate rises to 1.2817 K/yr from 0.6837. Together the step and the steeper drift draw a drift that starts later and climbs faster, the ridge of the drift's onset that Chapter 10 found (Table 10.9), met here as a step candidate. Coil 1's candidates at the same windows stay below zero, -0.98 at window 24, and window by window from 20 to 36 coil 1's $\ln B$ lies below unit 1's sensor's except at windows 27, 32, 33, 34 and 36, where both are below $+0.6$. The family puts the late change of unit 1's discharge reading on the sensor, not on coil 1.

Coil 1 in year 3. Coil 1's candidates rise above zero only at window 32, with $\ln B = +0.4508$ and a step of -0.0480, and at window 34, with $+0.5811$ and -0.0510; none decides, and coil 1's candidates of year 3 hold 0.0143 of the scan's probability. A bend spread over months is not one step; whether coil 1's rate changes is Chapter 11's question.

12.3.3 The month of the step

Asset manager's question. Was it March, or the month either side? The work order dates the pump change to a month; the dashboard put coil 2's drop a month earlier.

General form. The diagnosis (2.15) over windows: the candidates are one kind's steps at different windows.

Specialized form. The question is equation (2.15) over coil 2's drawings $M_{2,s}$, with a location prior uniform over the work order's neighborhood, $s \in \{14, 15, 16\}$,

$$\Pr(t_0 = s \mid \mathbf{y}, \text{coil 2}) = \frac{B_{2,s}}{\sum_{s'} B_{2,s'}}, \tag{12.10}$$

the sum over the same set; over windows 1–36 it is the location posterior of the scan of Section 12.3.6. The step's size over the three windows is the mixture (7.10) of the three drawings' normals with these weights.

Building blocks. The Bayes factors $B_{2,14}$, $B_{2,15}$, $B_{2,16}$ and the three drawings' steps with their spreads (Section 12.2.1).

The month. By equation (12.10) window 14 (Feb Y2) holds 0.0876, window 15 (Mar Y2) 0.5447 and window 16 (Apr Y2) 0.3677: March or April, not February. The dashboard's February step is not where the conductance stepped. The step's size barely depends on which month it was: -0.0553 (sd 0.0234) at window 14, -0.0714 (sd 0.0235) at 15 and -0.0687 (sd 0.0237) at 16, and averaged over the three by their weights it has median -0.06906 and band $[-0.10833, -0.02947]$. The month is the record's resolution: one representative hour per month cannot place the event within it.

The location posterior over the whole record is drawn in Figure 12.3. The horizontal axis is the window s, the vertical axis $\Pr(t_0 = s \mid \mathbf{y}, \text{coil 2})$ under a uniform prior over windows 1–36, the view of Section 12.3.6; the inset gives the three windows of this question renormalized. Two clusters hold most of it: windows 1 to 3 hold 0.3803, with the mode at window 2 (0.3369), and windows 14 to 16 hold 0.4913, the rest 0.1284. Within the second cluster March is the largest bar.

The dashboard's step is set beside the parameter's in Figure 12.4. The horizontal axis is the window and the vertical axis the health index against commissioning; the dots are coil 2's 37 raw-performance tiles, the dashed line the dashboard's line with its step into window 14, and

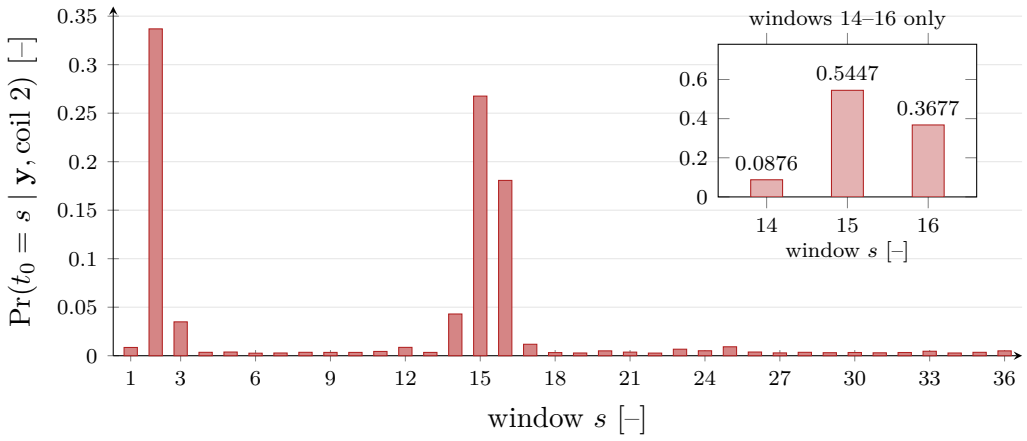


Figure 12.3: The location posterior of coil 2's step, equation (12.10), under a uniform prior over windows 1–36: $B_{2,s}$ normalized over the windows. Inset: the same over the work order's neighborhood, windows 14–16, renormalized (the question's view).

the solid line with its shaded band coil 2's health index $h_{2,t} = \mathrm{e}^{g_{2,t} - g_{2,0}}$ under $M_{2,15}$, its median and 5–95 % band from the step drawing's posterior. The tiles scatter far more widely than the band, and the tile of window 13 stands above every tile around it; the dashboard's line jumps between windows 13 and 14, while the index's step falls between windows 14 and 15, and the index has lost far less than the tiles say.

12.3.4 The size of the step and its share of the loss

Asset manager's question. How large was the step, and how much of coil 2's loss is it? The first part asks for the step on the coil's conductance; the second how coil 2's loss since commissioning divides between the event and the trend.

General form. An estimation question asks for the posterior distribution of a derived quantity, equation (2.9),

$$\Pr(f(\mathbf{z}) \le f^* \mid \mathbf{y}) = \int \mathbf{1}[f(\mathbf{z}) \le f^*]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z},$$

summarized by its moments, equation (2.10), and its credible interval, equation (2.11), at the levels 0.05 and 0.95.

Specialized form. The question is equations (2.9), (2.10) and (2.11) with $\mathbf{z}$ the state $\mathbf{z}_{2,15}$ of equation (12.2), $p(\mathbf{z} \mid \mathbf{y})$ its Laplace normal $\mathcal{N}(\hat{\mathbf{z}}_{2,15}, \boldsymbol{\Sigma}_{2,15})$, and $f(\mathbf{z}) = \delta_2$, the gradient $\mathbf{e}_\delta$ of the pattern (12.3); the step as a fraction of conductance is $100(\mathrm{e}^{\delta_2} - 1)$ %. For the share, the MAP identity (12.4) divides $\widehat{\ln h_{2,36}}$ into the rate's part and the step, and the step's share of the loss is $\hat{\delta}_2 / \widehat{\ln h_{2,36}}$; the parts' variances and cross term, and the spread of their sum, are $\mathbf{g}^\mathsf{T} \boldsymbol{\Sigma} \mathbf{g}$ on the gradients $36\,\Delta t\,\mathbf{e}_{\gamma_2}$, $\mathbf{e}_\delta$ and $\mathbf{g}_{\text{split}}$ of the pattern (12.3). What a reversal of the step's cause would restore is the conductance $G_{2,36}(\mathrm{e}^{-\delta_2} - 1)$ at window 36.

Building blocks. $\hat{\delta}_2$, $\hat{\gamma}_2$, $\hat{g}_{2,0}$, $\hat{g}_{2,36}$ and their block of $\boldsymbol{\Sigma}_{2,15}$ (Section 12.2.3); the normal quantile 1.6449 and the distribution function Φ.

The size of the step. The step is $\hat{\delta}_2 = -0.07139$, sd 0.02349, band $[-0.11003,\ -0.03274]$: a loss of 6.89 % of coil 2's conductance, with the band from 3.22 to 10.42 %, and $\Pr(\delta_2 < 0 \mid \mathbf{y}) = 0.9988$. The full posterior gives the median -0.07124 and the band $[-0.11010,\ -0.03255]$. The

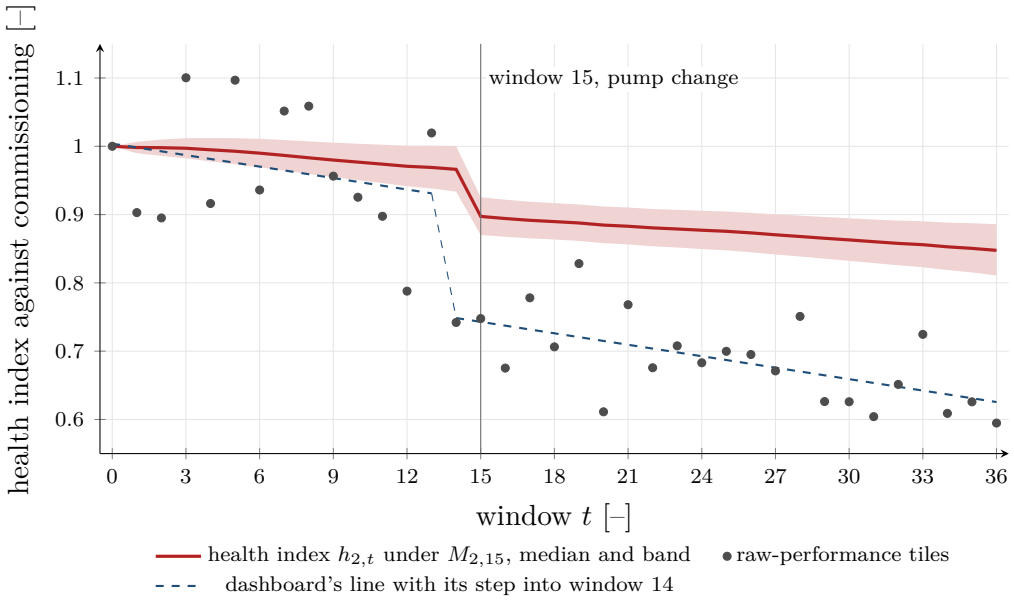


Figure 12.4: Coil 2's raw-performance tiles and the dashboard's step against the health index. Dots: the tiles $A_{2,0}/A_{2,t}$. Dashed: the dashboard's least-squares line with a step into window 14, intercept 1.0038, slope −0.0671 per yr and step −0.1770. Solid and shaded: $h_{2,t} = \mathrm{e}^{g_{2,t}-g_{2,0}}$ under the step drawing $M_{2,15}$, the median and the 5–95 % band, $\exp(\widehat{\ln h}_{2,t} \pm 1.6449\,\mathrm{sd})$. Gray line: window 15.

probability that the step is larger than 5 % is $\Pr(\delta_2 < \ln 0.95) = 0.8038$, and larger than 10 %, 0.074. The trend row's own spread is $\sigma_q = 0.005$: the path's wander from one month to the next cannot carry a step of this size.

The share of the loss. By the identity (12.4), $-0.16556 = -0.09417 + (-0.07139)$, so the step is 0.431 of coil 2's loss since commissioning and the fouling at the rate 0.5688. As quantities of the state the two parts have the variances 2.5195×10^{-3} (spread 0.05019) and 5.5195×10^{-4} (spread 0.02349), and their cross term, $2\,\mathrm{Cov} = -1.4474 \times 10^{-3}$, is negative, the step–rate correlation of Section 12.2.3. The cross term makes their sum, spread 0.04030, narrower than the rate's part, but not narrower than the step: the record knows the step better than it knows the sum. The sum is not the index as a quantity of the state, whose spread is 0.02691, because the index also carries the summed wander. In conductance, the step took 0.4247 kW/K of coil 2's 5.7392 kW/K at window 36, what a reversal of its cause would restore. The health indices today barely move: coil 2's is 0.8474, band [0.8107, 0.8858], against the base fit's 0.8289, and coil 1's 0.8764 against 0.8806. The readings fix today's conductance; the step changes how the loss is shared between the event and the trend.

The step is drawn in Figure 12.5. In the left panel the horizontal axis is the step in percent of coil 2's conductance and the vertical axis its posterior density; the solid curve is the Laplace normal of δ_2 carried through $100(\mathrm{e}^{\delta_2} - 1)$, the bars the reweighted draws' histogram, the thin curve the prior's density, the gray line zero and the dotted line the stated truth's step (Section 12.4.5). In the right panel the horizontal axis is the window and the vertical axis coil 2's log-derate $g_{2,t}$; the solid line and band are the path under $M_{2,15}$, its MAP and 5–95 % band, the dashed line the base fit's path. The posterior lies wholly on the side of a loss and far inside the prior; the stepped path is flatter than the base fit's on both sides of window 15 and drops across it.

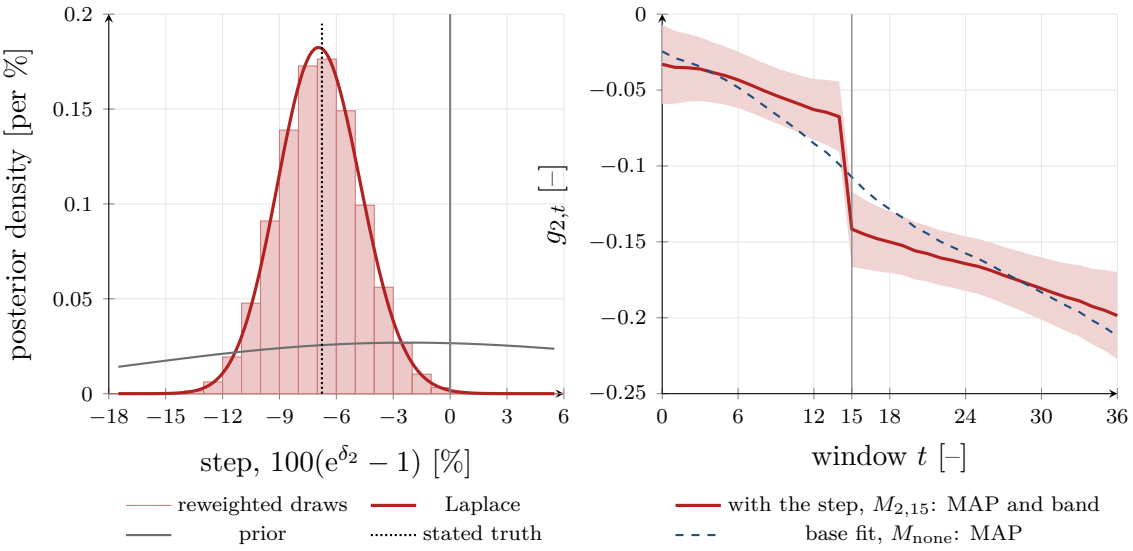


Figure 12.5: The step on coil 2 at window 15. Left: the posterior density of the step in percent of conductance, the Laplace normal $\delta_2 \sim \mathcal{N}(-0.07139,\ 0.02349^2)$ carried through $x = 100(\mathrm{e}^{\delta_2} - 1)$, density $\varphi((\ln(1+x/100) - \hat{\delta}_2)/s)/(s\,(100+x))$; the reweighted draws' histogram in 1 % bins; the prior's density, $\delta_2 \sim \mathcal{N}(0,\ 0.15^2)$, through the same map; gray, zero; dotted, the stated truth's step, a check marker (Section 12.4.5). Right: coil 2's log-derate path, the MAP of $M_{2,15}$ with its 5–95 % band (±1.6449 sd per window) and the base fit's MAP dashed; gray line, window 15.

12.3.5 The rates with the step allowed for

Asset manager's question. With the step allowed for, how fast is each coil losing conductance, and which is faster? Chapter 10 read coil 2 as the faster of the two with the path taken as a line with wander and no event.

General form. The rate is the first line of the trend form, equation (2.21), read as an estimation of the trend entry, and the comparison is a prediction under a threshold, equation (2.12),

$$\Pr(f(\mathbf{z}) \geq f^* \mid \mathbf{y}) = \int \mathbf{1}[f(\mathbf{z}) \geq f^*]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z},$$

here without an assurance level. Over a family that stays open, both are read from the mixture of model averaging, equation (7.10).

Specialized form. The question is equations (2.21) and (2.12) with $f(\mathbf{z}) = \gamma_u$ for the rate and $f(\mathbf{z}) = \gamma_1 - \gamma_2$, $f^* = 0$, for the comparison, under the mixture (7.10) over the named-window family: each drawing, M_{none} and the eight $M_{j,15}$, gives its Laplace normal of γ_u, and the weights are the posterior probabilities of equation (12.8), $1/(1 + \bar{B}_{15})$ for M_{none}. Given the step on coil 2, the comparison is the normal tail

$$\Pr(\gamma_2 < \gamma_1 \mid \mathbf{y}, M_{2,15}) = \Phi\left(\frac{\hat{\gamma}_1 - \hat{\gamma}_2}{\sqrt{\mathbf{g}_{D_\gamma}^{\mathsf{T}}\, \boldsymbol{\Sigma}_{2,15}\, \mathbf{g}_{D_\gamma}}}\right), \qquad \mathbf{g}_{D_\gamma}^{\mathsf{T}}\, \boldsymbol{\Sigma}\, \mathbf{g}_{D_\gamma} = \Sigma_{\gamma_1\gamma_1} + \Sigma_{\gamma_2\gamma_2} - 2\Sigma_{\gamma_1\gamma_2}, \tag{12.11}$$

on the gradient $\mathbf{g}_{D_\gamma}$ of the pattern (12.3). The mixture's quantiles are the mixture's own, read from the weighted sum of the nine normal distribution functions. One choice enters here: the reported rate is the mixture, not the rate given the step. This is what lets the rate carry the step's probability rather than assume it, and it fails only as the family fails; the rates given the step stand beside it.

Building blocks. Each drawing's $\hat{\gamma}_u$ and its spread; the weights of equation (12.8); the block of $\mathbf{\Sigma}_{2,15}$ on (γ_1, γ_2) (Section 12.2.3); Φ.

The rates. The weights of the named-window family are 0.2840 for M_{none}, 0.5637 for coil 2's step, 0.1054 for unit 2's sensor, 0.0138 for both coils, 0.0093 for unit 1's sensor, 0.0077 for rack 1's, 0.0069 for the meter, 0.0055 for coil 1 and 0.0037 for the return air. Under the mixture coil 1 loses 4.38 %/yr, band [−7.04, −1.94], and coil 2 4.14 %/yr, band [−7.36, −0.83], and

$$\Pr(\gamma_2 < \gamma_1 \mid \mathbf{y}) = 0.464.$$

Given the step on coil 2 at window 15, coil 1 loses 4.303 %/yr, band [−6.581, −1.969], and coil 2 3.090 %/yr, band [−5.721, −0.386]; with $\hat{\gamma}_1 - \hat{\gamma}_2 = -0.01259$, spread 0.02467 and correlation −0.2325, equation (12.11) gives $\Pr(\gamma_2 < \gamma_1 \mid \mathbf{y}, M_{2,15}) = 0.305$, and the full posterior 0.31179 (Monte Carlo standard error 0.00133), with coil 2's median rate at −3.120 %/yr. Chapter 10's base fit read −4.151 and −6.062 %/yr, with $\Pr(\gamma_2 < \gamma_1) = 0.8179$. Chapter 10's "coil 2 is losing faster" was the March step folded into a constant rate. Given the step, coil 2's rate halves, and with the step as probable as the record makes it, the record no longer says which coil is faster.

Coil 2's rate is also less certain with the step free, its spread rising from 0.01323 to 0.01673, because the step and the rate share the same windows. Its variance budget given the step is 53.0 % readings and 47.0 % wander, and coil 1's 43.8 and 56.2 %, the readings' part being the variance on the straight line of Chapter 10, $\sigma_q \to 0$. On that straight line with the step, the rates are −3.77 and −3.60 %/yr ($\gamma_1 = -0.03846$, $\gamma_2 = -0.03662$), the step −0.06430 (sd 0.02036), and $\Pr(\gamma_2 < \gamma_1) = 0.460$. The rates go to Table 12.12.

12.3.6 A step anywhere in the record

Asset manager's question. Did a step happen anywhere in the three years? The question asks of every window at once, March included, as it would be asked with no work order to name a month.

General form. The diagnosis (2.15) over the whole family.

Specialized form. The question is equation (2.15) over M_{none} and all 288 candidates, with the scan's uniform location prior of equation (12.5),

$$\bar{B} = \frac{1}{288} \sum_{j,s} B_{j,s}, \qquad \Pr(\text{a step somewhere} \mid \mathbf{y}) = \frac{\bar{B}}{1 + \bar{B}}, \qquad \Pr(M_{j,s} \mid \mathbf{y}) = \frac{B_{j,s}/288}{1 + \bar{B}}. \tag{12.12}$$

The look-elsewhere charge is $\ln B_{\max} - \ln \bar{B}$, the best candidate's log Bayes factor less the scan's: it is $\ln 288$ less the logarithm of the sum of all Bayes factors over the best one, at most $\ln 288$. What a candidate rests on is read by its influence: its $\ln B$ by the closed form (12.7) on the base fit refitted with one window's readings left out.

Building blocks. The 288 refits and their log Bayes factors (12.6); the closed form (12.7) beside each and on the base fit with each window's temperatures left out.

The ranking. Table 12.6 gives the twelve largest candidates. Each row is a candidate; the columns give its kind and window, its log Bayes factor by refit and by the closed form, the gain in fit and the Occam factor of equation (12.6), and its step with the spread, also in percent of conductance or in its reading's units. The record's largest candidate is not March: it is a rise of 6.06 % on both coils into window 2, February of year 1, with $\ln B = +3.7216$ ($B = 41.33$); then coil 2 rising 9.23 % into window 2 ($B = 19.99$), and coil 2 at window 15 third ($B = 15.88$).

Table 12.6: The twelve largest of the 288 candidates, by the refit's $\ln B_{j,s}$ against M_{none}; the closed form from the base fit alone beside it.

Rank [–]	Kind	Window [–]	$\ln B$, refit [–]	$\ln B$, closed [–]	Gain [–]	Occam [–]	Step (sd) [–, K]	Step [%, K]
1	both coils	2 (Feb Y1)	+3.7216	+3.7256	5.8988	2.1772	+0.0588 (0.0172)	+6.06 %
2	coil 2	2 (Feb Y1)	+2.9952	+2.9735	4.6194	1.6242	+0.0882 (0.0295)	+9.23 %
3	coil 2	15 (Mar Y2)	+2.7648	+2.7773	4.6295	1.8647	−0.0714 (0.0235)	−6.89 %
4	coil 2	16 (Apr Y2)	+2.3719	+2.4027	4.2271	1.8552	−0.0687 (0.0237)	−6.64 %
5	coil 1	2 (Feb Y1)	+2.0005	+1.9997	3.6588	1.6583	+0.0777 (0.0291)	+8.08 %
6	unit 1's discharge sensor	24 (Dec Y2)	+1.9919	+2.0094	3.3913	1.3995	−0.6543 (0.2513)	−0.654 K
7	unit 2's discharge sensor	33 (Sep Y3)	+1.3722	+1.4065	2.9755	1.6033	−0.4950 (0.2034)	−0.495 K
8	unit 2's discharge sensor	15 (Mar Y2)	+1.0881	+1.0451	2.8417	1.7536	+0.3759 (0.1580)	+0.376 K
9	unit 2's discharge sensor	16 (Apr Y2)	+0.9707	+0.9297	2.7211	1.7504	+0.3723 (0.1598)	+0.372 K
10	unit 2's discharge sensor	14 (Feb Y2)	+0.9416	+0.8928	2.6967	1.7550	+0.3638 (0.1570)	+0.364 K
11	coil 2	14 (Feb Y2)	+0.9374	+0.9328	2.8028	1.8654	−0.0553 (0.0234)	−5.38 %
12	coil 2	3 (Mar Y1)	+0.7279	+0.7055	2.4872	1.7593	+0.0570 (0.0257)	+5.87 %

A step somewhere. By equation (12.12) the mean Bayes factor over the 288 candidates is $\bar{B} = 0.6870$ ($\ln \bar{B} = -0.3754$), and

$$\Pr(\text{a step somewhere} \mid \mathbf{y}) = \frac{0.6870}{1.6870} = 0.407.$$

Scanned without a named window, the record does not say that anything happened. The look-elsewhere charge is $+3.7216 - (-0.3754) = 4.0970$, of the most $\ln 288 = 5.6630$ it could be. By period the scan's probability is 0.157 in windows 1–3, 0.084 in February to April of year 2, 0.087 in year 3 (windows 25–36) and 0.079 elsewhere. Coil 2 alone, scanned uniformly over its 36 windows, has its location split between window 2, 0.3369, and February to April of year 2, 0.491, and a mean Bayes factor of 1.648 against the 15.88 of the named window (Exercise 12.4).

What the window-2 candidate rests on. The candidate of February of year 1 rests on one window. With window 1's four temperature readings left out its $\ln B$ falls to -0.11, while coil 2's step at window 15 keeps $+2.30$; no single window's removal takes coil 2 at 15 below $+1.27$, the value with window 13 left out. The work-order history holds no cleaning or coil work in year 1: the candidate has no cause on record. It is reported here, and not entered in the fit.

12.4 The checks (step 5)

Each answer of Section 12.3 rests on the Laplace evidence, on the step's prior, on the base fit's choices and on one record's draws of the sensor errors, and this hall lets each be tested against the calculation it stands for.

12.4.1 The full posterior and the other routes to the evidence

The reweighted draws. The full posterior of the step drawing $M_{2,15}$ and of the base fit was computed by the route of Section 5.2.4, with $N_{\text{d}} = 100\,000$ draws per drawing from a seeded generator and 20 batches for the standard errors. The effective sample sizes are 98 804 and 98 841, 98.8 % of N_{d}, and the largest weights 2.60×10^{-5} and 2.51×10^{-5}. The Monte Carlo standard errors are 0.00011 on the step's median, 0.00012 and 0.00017 on its band ends, at

most 0.00006 on the rates' medians, 0.00133 on $\Pr(\gamma_2 < \gamma_1)$, and 0.00041 and 0.00025 on the two log evidences. Against the Laplace values the step's median lies 0.00015 away and its band ends 0.00007 and 0.00019, each within two standard errors. Coil 2's median rate is $-3.120\,\%/\mathrm{yr}$ against -3.090, and $\Pr(\gamma_2 < \gamma_1)$ is 0.31179 against 0.30485, a gap of 0.0069, 5.2 standard errors: resolved and real, the full posterior leaning slightly more toward coil 2 being the faster. By importance sampling, equation (7.12), the log evidences are -66.1771 and -68.9370 against the Laplace -66.1687 and -68.9335, so $\ln B_{2,15}$ is $+2.7599$ against $+2.7648$. These are the full posterior's values, beside the reported ones.

The Savage–Dickey ratio. The step drawing and the base are a nested pair, and equation (7.21) gives the Bayes factor from the step drawing's own fit. At the winner,

$$\ln B_{2,15} = \ln \frac{0.02349}{0.15} + \frac{0.07139^2}{2 \times 0.02349^2} = -1.8539 + 4.6164 = +2.7625,$$

against the Laplace $+2.7648$ (Exercise 12.3). Over the 288 candidates the largest gap between the two routes is 0.1570, and at the five largest candidates of Table 12.6 it is 0.0321, 0.1392, 0.0023, 0.0191 and 0.0683. The routes agree where the rows are nearly linear over the step's posterior, as in Chapter 7, and the closed form of Section 12.2.2 stays within 0.1577 of the refits.

12.4.2 The limit of exact rows and the curvature term

The fit runs on the reduced state. The check fits the winner's own full state, the 116 entries of $M_{2,15}$ and each window's five temperatures, 301 entries in all, with the 185 physical rows (5.2) at a common finite tolerance, and lets the tolerance shrink. Table 12.7 gives the result: each row is a tolerance, the columns the step and the two rates with their spreads, and the last row the reduced state. At 0.1 the spreads are slightly wider; at 0.01 and 0.001 the full state reproduces the reduced state to the fifth decimal.

Table 12.7: The limit of exact physical rows on the winner's full state of 301 entries: the step and the two rates with their posterior spreads at three tolerances, against the reduced state.

Tolerance σ_p [kW or K]	δ_2 [–]	sd [–]	γ_1 [yr^{-1}]	sd [yr^{-1}]	γ_2 [yr^{-1}]	sd [yr^{-1}]
0.1	−0.07131	0.02361	−0.04374	0.01471	−0.03142	0.01677
0.01	−0.07139	0.02349	−0.04398	0.01465	−0.03139	0.01673
0.001	−0.07139	0.02349	−0.04398	0.01465	−0.03139	0.01673
reduced state	−0.07139	0.02349	−0.04398	0.01465	−0.03139	0.01673

The covariance is the Gauss–Newton form, without the curvature term $\mathbf{C}$ of equation (4.44). The full Hessian at the winner's MAP state gives the Newton spreads 0.02348 for the step, 0.01457 for γ_1 and 0.01668 for γ_2, ratios of 0.9996, 0.9946 and 0.9968 to Gauss–Newton's. The term $\ln\det\mathbf{H} - \ln\det\mathbf{J}^{\mathsf{T}}\mathbf{J}$ is 0.022 for the winner and 0.012 for the base, which would move $\ln B_{2,15}$ by -0.005 if the Newton curvature were used in the evidence. Gauss–Newton is the one reported.

12.4.3 Sensitivity to the priors and the choices

Table 12.8 repeats the analysis under other priors and choices, each by its own base refit, the closed-form scan of all 288 candidates on that base, and the winner's refit. Each row is a variant; the columns give the winner's $\ln B_{2,15}$, its step with the spread, the named window's $\Pr(t_0 = 15 \mid \mathbf{y})$ and coil 2's share of it, the scan's largest candidate with its $\ln B$, the two rates given the step and $\Pr(\gamma_2 < \gamma_1)$ given the step. The probability columns come from the closed form, so the base row reads 0.7167 where question 1's refits read 0.7160.

Table 12.8: Sensitivity of the March step and of the scan to the priors and the choices. Rates given the step on coil 2 at window 15; the scan's top as kind and window ($T_{d,2}$: unit 2's discharge sensor). In the row with both coils' curvatures freed, the rates are the rates at the record's center, $\bar{t}_y = 1.5\,\text{yr}$, not constant rates, so they are not like with like with the other rows.

Variant [–]	$\ln B_{2,15}$ [–]	δ_2 [–]	sd [–]	Pr, step at 15 [–]	coil 2's share [–]	scan's top [–]	its $\ln B$ [–]	rate 1 [%/yr]	rate 2 [%/yr]	Pr, 2 faster [–]
the base fit (reported)	+2.765	−0.0714	0.0235	0.7167	0.7943	both@2	+3.73	−4.303	−3.090	0.3048
coil step prior sd 0.05	+3.018	−0.0597	0.0215	0.7630	0.8029	both@2	+4.22	−4.280	−3.583	0.3826
coil step prior sd 0.10	+3.018	−0.0693	0.0231	0.7585	0.8241	both@2	+4.03	−4.299	−3.180	0.3183
coil step prior sd 0.30	+2.167	−0.0727	0.0237	0.6143	0.6942	both@2	+3.10	−4.306	−3.034	0.2966
$\sigma_q = 0.0025$	+2.993	−0.0670	0.0213	0.7585	0.8054	both@2	+3.50	−3.950	−3.424	0.3935
$\sigma_q = 0.01$	+1.802	−0.0748	0.0284	0.5269	0.6824	both@2	+3.40	−4.856	−2.541	0.2537
the straight line, $\sigma_q \to 0$ (closed form linearized)	+3.016	−0.0643	0.0204	0.7634	0.8005	both@2	+3.26	−3.773	−3.596	0.4601
drift onset window 16	+2.972	−0.0730	0.0235	0.7752	0.7155	both@2	+3.85	−4.531	−2.936	0.2523
drift onset window 24	+2.589	−0.0699	0.0235	0.6931	0.7466	both@2	+3.70	−4.375	−3.071	0.2881
drift onset window 28	+2.527	−0.0693	0.0234	0.7148	0.6317	both@2	+3.73	−4.627	−2.917	0.2255
no drift entry	+3.086	−0.0732	0.0233	0.9950	0.0139	$T_{d,2}$@15	+7.35	−7.149	−1.633	0.0039
rating prior sd 0.05	+2.685	−0.0707	0.0235	0.7014	0.7893	both@2	+2.87	−4.266	−3.220	0.3270
both coils' curvatures freed ($\sigma_\dot{\gamma} = 0.05\,\text{yr}^{-2}$)	+2.928	−0.0749	0.0243	0.7865	0.6452	coil 2@2	+4.24	−4.714	−2.702	0.2035
window 20's unit-2 discharge reading left out	+1.529	−0.0613	0.0236	0.4517	0.7083	both@2	+3.80	−4.405	−3.435	0.3411
window 1's four temperature readings left out	+2.287	−0.0677	0.0235	0.6369	0.7121	coil 2@15	+2.30	−4.733	−3.601	0.3197
the coils' kinds only in the family	+2.765	−0.0714	0.0235	0.8471	0.9674	both@2	+3.73	−4.303	−3.090	0.3048

The step's size is stable, −0.0597 to −0.0749 across every row; its evidence is not, +1.529 to +3.086. The priors act on each coil through the step. A narrower step prior raises $\ln B_{2,15}$, because the Occam factor charges less for a narrow entry, and shrinks the step to −0.0597 at 0.05; the loss the step gives up returns to coil 2's rate, which steepens to −3.583 %/yr, while coil 1's moves only from −4.303 to −4.280 %/yr, through the −0.2325 correlation of the two rates. A wider prior lowers both the probability and coil 2's share. A larger wander, $\sigma_q = 0.01$, lets the path carry part of the step ($\ln B = +1.80$) and steepens coil 1 to −4.856 %/yr, as in Chapter 10; a smaller one does the reverse. The drift's onset moves coil 2's share between 0.6317 (onset 28) and 0.7943 (onset 20) through the sensor alternatives, while the step stays within −0.0693 to −0.0730.

Without the drift entry the family's top candidates become sensor steps, unit 2's at window 15 (+7.35), then unit 1's at window 31 (+7.29), unit 2's at 16 and 14, and coil 2's share at window 15 collapses to 0.0139: the sensor's change is found by the family even when the drift is not carried, on the sensor, and the base fit's drift is what lets the family read coil 2 at all.

The step beside a changing rate. Chapter 11 frees a curvature on each coil's trend rows, the rate's change per year $\dot{\gamma}_u$ on the lever $x_t = (t+\frac{1}{2})\Delta t - \bar{t}_y$ about the record's center $\bar{t}_y = 1.5\,\text{yr}$, with prior spread $\sigma_\dot{\gamma} = 0.05\,\text{yr}^{-2}$; its drawing with both curvatures is the row of Table 12.8 that frees them beside the step. The March step holds: $\ln B_{2,15} = +2.928$, a step of −0.0749 (sd 0.0243), and the named-window probability rises to 0.7865, while coil 2's share falls to 0.6452. Coil 1's curvature is −0.03341 (sd 0.02118) without the step and −0.03422 (sd 0.02119) with it, $\Pr(\dot{\gamma}_1 < 0) = 0.9469$: the step does not touch coil 1's bend. Coil 2's curvature is +0.01623 (sd 0.02009) without the step, $\Pr(\dot{\gamma}_2 < 0) = 0.2096$, a lean to a slowing rate, and +0.00061 (sd 0.02071) with it, $\Pr(\dot{\gamma}_2 < 0) = 0.4882$: the slowing that a changing rate reads on coil 2 is this step. The scan's top becomes coil 2 at window 2 (+4.24), then both coils there (+3.92), and coil 1's candidates of year 3 fall to at most −0.1161. Whether coil 1's rate changes is Chapter 11's question.

The influence of single windows. The March evidence leans on several windows at once: window 20's unit-2 discharge reading, Chapter 10's largest normalized residual, left out takes it to +1.53, and windows 13, 16, 17 and 20 left out each take it to between +1.27 and +1.79. No one window carries it, as one window carries the window-2 candidate.

A second step beside the winner. The closed form on the winner's fit ranks a second step on both coils at window 2 first, $\ln B = +2.829$ with a step of +0.0543. The two-step refit has $\ln B = +2.8258$ against the winner and +5.5906 against no step, with steps of −0.06426 (sd 0.02358) at window 15 and +0.05441 (sd 0.01727) at window 2, rates of −5.104 and −4.002 %/yr, and $\Pr(\gamma_2 < \gamma_1) = 0.3202$. It is not adopted: the second step is the one-window candidate of Section 12.3.6, with no work on record.

12.4.4 Fresh records and the time to decide

The verdicts of step 4 are one record's draws. The check refits 200 records of fresh sensor errors at the stated truth by the closed form and the winner's refit, and Table 12.9 summarizes them. Each row is a quantity; the columns give its median and 5–95 % range over the records, this record's value with its quantile among them, and the share of records in which the verdict holds.

Table 12.9: Fresh records at the stated truth: 200 records of fresh sensor errors, by the closed form and the winner's refit. This record's values are its closed-form values, like the records'. A verdict holds at $\Pr \geq 0.95$, for coil 2 alone at $B \geq 19$.

Quantity [–]	Median [–]	5 % [–]	95 % [–]	This record (quantile) [–]	Verdict holds [–]
$\ln B_{2,15}$, coil 2 at window 15	+2.118	−1.053	+6.932	+2.7773 (0.600)	0.365
$\Pr(t_0 = 15 \mid \mathbf{y})$, the named window	0.6674	0.1891	0.9923	0.7167 (0.570)	0.185
$\Pr(\text{a step somewhere} \mid \mathbf{y})$, the scan	0.3621	0.2100	0.8585	0.4079 (0.565)	0.020

This record sits near the middle of every row. The named-window probability reaches 0.75 in 0.405 of the records. The scan's winner is coil 2 at window 15 in 0.385 of the records, coil 2 at windows 14–16 in 0.460, a candidate in windows 1–3 in 0.015 and one in year 3 in 0.265. The truth holds no event in windows 1–3, and the largest $\ln B$ there has median −0.090 and 95th percentile +1.992 over the records: a candidate there as strong as this record's, +3.7256 by the closed form, appears in 0.015 of the records, and one anywhere away from the truth's events (windows 1–13 and 17–19) in 0.065, with $\ln B \geq \ln 19$ there in 0.110. The step at window 15 has mean −0.06519 and spread 0.02143 over the records, against the posterior's 0.02349; its band covers the true −0.07 in 0.925 of the records and the drawing's pseudo-true value in 0.935, and γ_2's band covers its pseudo-true value in 0.975. Given the step, coil 2 is read the faster of the two coils in 0.185 of the records. The truth's step of 6.76 % is at the edge of what monthly windows find: it is decided on coil 2 at its own window in 0.365 of the records, and the named-window question in 0.185.

The record as of each window. Table 12.10 cuts this record at each window t_{as} from 15 to 36, refits the base on windows 0 to t_{as} and scores every candidate up to t_{as} by the closed form; each fit uses nothing later than its window. Each row is an as-of window; the columns give coil 2's $\ln B_{2,15}$ with its step and spread, the named-window probability and coil 2's share of it, the probability of a step somewhere, and the scan's top candidate with its $\ln B$.

At $t_{\text{as}} = 15$ the step is read from one window, $\ln B = -0.68$ and a step of -0.04336 ± 0.03455; by $t_{\text{as}} = 17$, May of year 2, it is +2.18 with -0.07666 ± 0.02754, and from there $\ln B$ wanders between +2.04 and +2.78 to $t_{\text{as}} = 36$. The named-window probability peaks at 0.802 at $t_{\text{as}} = 26$

Table 12.10: The record as of each window t_{as}: the base fit cut at t_{as} and every candidate up to t_{as} scored by the closed form.

t_{as} [–]	Window [–]	$\ln B_{2,15}$ [–]	δ_2 [–]	sd [–]	Pr, step at 15 [–]	coil 2's share [–]	Pr, step somewhere [–]	scan's top [–]	its $\ln B$ [–]
15	Mar Y2	−0.6808	−0.04336	0.03455	0.3968	0.0962	0.5159	both@2	+3.966
16	Apr Y2	+0.9720	−0.06590	0.02877	0.4903	0.3434	0.6646	both@2	+4.855
17	May Y2	+2.1792	−0.07666	0.02754	0.6327	0.6413	0.7338	both@2	+5.085
18	Jun Y2	+2.1978	−0.07520	0.02686	0.6239	0.6786	0.7057	both@2	+4.914
20	Aug Y2	+2.1785	−0.07452	0.02666	0.7021	0.4685	0.7809	coil 2@2	+5.180
22	Oct Y2	+2.0362	−0.07311	0.02665	0.6862	0.4379	0.6886	coil 2@2	+4.967
24	Dec Y2	+2.0999	−0.07324	0.02644	0.6758	0.4896	0.5906	coil 2@2	+4.530
26	Feb Y3	+2.6849	−0.07752	0.02602	0.8020	0.4523	0.5184	coil 2@2	+3.724
28	Apr Y3	+2.6112	−0.07554	0.02552	0.7433	0.5878	0.4742	both@2	+3.915
30	Jun Y3	+2.5881	−0.07392	0.02497	0.7237	0.6348	0.4185	both@2	+3.640
32	Aug Y3	+2.2680	−0.07003	0.02452	0.6467	0.6597	0.3899	both@2	+3.542
34	Oct Y3	+2.6653	−0.07176	0.02392	0.7082	0.7404	0.4202	both@2	+3.782
36	Dec Y3	+2.7773	−0.07138	0.02345	0.7167	0.7943	0.4079	both@2	+3.726

and is 0.7167 at $t_{\text{as}} = 36$: at no as-of window does this record decide it, and coil 2 alone never reaches $\ln 19$ on the grid. The scan's top is a window-2 candidate at every t_{as}.

The time at which a verdict first holds is one record's draw, so its spread over fresh records is the measure of it. Over 100 fresh records, each cut at $t_{\text{as}} = 16$, 18, 20, 24, 28, 32 and 36, the share whose verdict holds at t_{as} is 0.14 for the named window and 0.20 for coil 2 alone at $t_{\text{as}} = 16$, then 0.15 and 0.35 at 18, 0.17 and 0.36 at 20, 0.16 and 0.39 at 24, 0.18 and 0.39 at 28, 0.22 and 0.42 at 32, and 0.22 and 0.43 at 36. A verdict holds at some $t_{\text{as}} \leq 36$ in 0.32 of the records for the named window and 0.51 for coil 2 alone, and among those records the first such t_{as} has median 18 (June of year 2) and 5–95 % range [16, 34]. What a record does not show soon after the event, three more years rarely add: the step's evidence is set by the windows next to it.

Both views are drawn in Figure 12.6. In the left panel the horizontal axis is the as-of window t_{as} and the vertical axis this record's $\ln B_{2,15}$, with $\ln 19$ dashed. In the right panel the same horizontal axis carries probabilities and shares: this record's named-window probability (solid), and the share of fresh records whose verdict holds at t_{as} for the named window and for coil 2 alone (dashed), with 0.95 dotted. This record's lines rise in the months after March and then run flat below their verdict levels. The fresh records' share for coil 2 alone rises between $t_{\text{as}} = 16$ and 18 and changes little after, and the named window's stays between 0.14 and 0.22.

12.4.5 The answers against the stated truth

The last check sets the answers against the stated truth of the record (Section 12.6.5). Table 12.11 collects the routes: each row is a route or a reference, and the columns give the step with its spread and band, coil 2's trend entry with its spread, the probability that coil 2 is the faster and $\ln B_{2,15}$. The like-with-like reference for a drawing is its pseudo-true value, the drawing fitted to the truth's noise-free readings, because coil 1's truth is not a straight line.

Coil 2's true step is −0.0700 into window 15, −6.76 % of its conductance, and the posterior sits −0.06 standard deviations from it; the drawing's pseudo-true step is −0.06296. Coil 2's true rate is $-0.0200\,\text{yr}^{-1}$, −1.980 %/yr, and the posterior sits −0.68 standard deviations from it and −0.34 from the pseudo-true −0.02566. Coil 1's pseudo-true rate is −0.04010 (how its truth bends is Chapter 11's), and the posterior sits −0.26 standard deviations from it. The true $\ln h_{2,36}$ is −0.1300, a step of −0.07 and fouling of $3 \times (-0.02) = -0.06$, against the posterior's −0.16556, −1.32 standard deviations: the drawing's straight coil 1 hands part of coil 1's year-3 loss to coil 2 through the return air, Chapter 10's finding, which Chapter 11 takes up. The

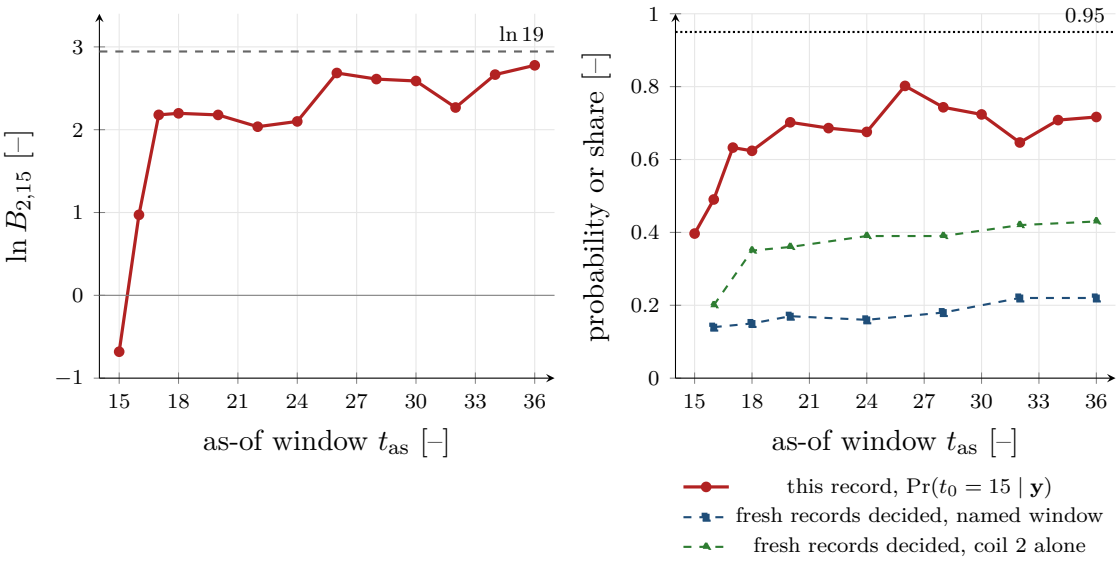


Figure 12.6: The step as of each window t_{as}, each fit using nothing later than t_{as} (Table 12.10). Left: this record's $\ln B_{2,15}$ by the closed form, with $\ln 19$ dashed. Right: this record's named-window probability (solid), and the share of 100 fresh records whose verdict holds at t_{as}, for the named window ($\Pr \geq 0.95$) and for coil 2 alone ($B \geq 19$) (dashed); dotted, 0.95.

Table 12.11: The routes to the step and to coil 2's rate given the step, the drawing's pseudo-true value and the stated truth. Bands are 5–95 %. For the truth the last probability column is the event itself: coil 1 is the faster. The Newton curvature would move $\ln B_{2,15}$ by -0.005.

Route [–]	δ_2 [–]	sd δ_2 [–]	band of δ_2 [–]	γ_2 [yr^{-1}]	sd γ_2 [yr^{-1}]	$\Pr(\gamma_2 < \gamma_1)$ [–]	$\ln B_{2,15}$ [–]
Laplace on the winner's state (reported)	−0.07139	0.02349	[−0.11003, −0.03274]	−0.03139	0.01673	0.3048	+2.7648
Full posterior, reweighted draws (medians)	−0.07124	–	[−0.11010, −0.03255]	−0.03169	–	0.31179	+2.7599
Savage–Dickey, the winner's own fit	–	–	–	–	–	–	+2.7625
Closed form, base fit alone	−0.07138	0.02345	–	–	–	–	+2.7773
$\sigma_p = 0.1$ on the full state	−0.07131	0.02361	–	−0.03142	0.01677	–	–
Newton covariance	–	0.02348	–	–	0.01668	–	–
Drawing's pseudo-true value	−0.06296	–	–	−0.02566	–	–	+1.7355
Stated truth	−0.0700	–	–	−0.0200	–	0	–

truth has no event in windows 1–3 and no step on any sensor. The scan's top candidate, both coils at window 2, is the record's own errors at window 1, which sit on one side together in standard-deviation units: +0.71 on the return air, +0.59 and +0.91 on the two discharges, −1.21 on the meter and +1.14 on rack 1's inlet. It is a false alarm of the kind the fresh records show in 1.5 % of records. The truth's drift is linear from window 20 at 0.60 K/yr, and unit 1's step candidate at window 24 would put it at 1.28 K/yr.

12.5 The report (step 6)

On the coils' conductance, something probably happened in March of year 2, with probability 0.72, but the record does not decide it. If it did, it was coil 2 (0.79): a loss of 6.9 % of its conductance (3.2 to 10.4 %), in March or April (0.54, 0.37), not in February, where the dashboard put it. The alternative is unit 2's discharge sensor reading 0.38 K high from March (0.15); the other readings favor the coil. The step changes Chapter 10's answer: with the step allowed for, coil 1 is losing 4.4 %/yr and coil 2 4.1 %/yr; given the step, coil 2 is losing 3.1, not 6.1; and with the step as probable as it is, which coil is faster is open (0.46, from 0.82). The step is 43 % of coil 2's loss since commissioning, 0.42 kW/K that correcting its cause would restore. The record's largest candidate is a rise on both coils into February of year 1 (Bayes factor 41) that rests on window 1's readings alone; without the work order naming March, the scan over three years would not say that anything happened (0.41).

12.5.1 The answers and their bands

Table 12.12 gives the answers in the order of Table 12.1. Each row is a question; the second column gives the reported value, from the Laplace evidences and the Laplace normal, or the mixture for the rates; the third the full posterior's value beside it where it was computed; and the fourth the band: the 5–95 % band of a step or a rate, and for a probability its range over the variants of Table 12.8.

The answers rest on the evidences of the family and on the step drawing's posterior. The step's size is stable across every prior and choice of Table 12.8; its probability is not, and the named window is what lifts it above the scan's 0.41.

12.5.2 The entry in the maintenance log

The two-unit hall's maintenance log holds the commissioning test (window 0), the condition assessment and the effort entry of window 18, Chapter 7's assessment of window 25 with its side-by-side check of unit 1's discharge sensor asked for and not made, and two entries at window 36: Chapter 10's rates, which noted the path's departure from a straight line around windows 13–16 untested, and Chapter 11's trend assessment, which left coil 1's acceleration undecided and coil 2's lean to a slowing rate untested against a single event. This chapter's entry follows them at window 36. It uses nothing later than window 36, and in the values of Table 12.12 it reads:

> *Window 36 (Dec Y3), change-point detection over the condition record (windows 0–36), for the chilled-water pump change of Mar Y2 (work order).*
> A step in March of year 2 (window 15): probable, not decided, $P = 0.72$. If it happened: coil 2 (0.79), a loss of 6.9 % of conductance (3.2 to 10.4 %), March or April (0.54, 0.37). Alternative: unit 2's discharge sensor 0.38 K high from March (0.15).
> Rates with the step allowed for: coil 1 −4.4 %/yr (−7.0 to −1.9), coil 2 −4.1 %/yr (−7.4 to −0.8); given the step, coil 2 −3.1 %/yr. Which coil is faster: open (0.46). Window 36's rate entry (coil 2 faster, 0.82) folded the step into the rate.
> Window 36's trend entry, coil 2's lean to a slowing rate, read against a single event: it is the step (coil 2's curvature +0.016 without the step, +0.001 with it).

Table 12.12: The report for the two-unit hall at window 36 (Dec Y3), change-point detection over windows 0–36 for the pump change of Mar Y2. Reported values from the Laplace evidences and the Laplace normal of $M_{2,15}$, the rates from the mixture over the named-window family; bands are 5–95 %, or for a probability the range over the variants of Table 12.8.

Asset manager's question	Reported value	Full posterior	Band
Did something happen in March?	$\Pr(t_0 = 15 \mid \mathbf{y}) = 0.716$ [–]; coil 2 at 15, 0.5637 [–]; $\bar{B}_{15} = 2.5206$ [–]; $\ln B_{2,15} = +2.7648$ [–]	$\ln B_{2,15} = +2.7599$ [–]	over the variants: 0.4517 to 0.9950 [–]
What stepped: a coil, or a sensor reading it?	coil 2 0.787; unit 2's discharge sensor 0.147 [–]	–	coil 2's share over the variants: 0.0139 to 0.9674 [–]
Was it March, or the month either side?	March 0.5447, April 0.3677, February 0.0876 [–]	–	step over the three windows: −0.10833 to −0.02947, median −0.06906 [–]
How large was the step, and how much of coil 2's loss is it?	$\delta_2 = -0.07139$ [–], a loss of 6.89 [%]; share of the loss 0.431 [–]; 0.4247 [kW/K]	median −0.07124 [–]	−0.11003 to −0.03274 [–]; 3.22 to 10.42 [%]
With the step allowed for, how fast is each coil losing conductance, and which is faster?	mixture: coil 1 −4.38, coil 2 −4.14 [%/yr]; $\Pr(\gamma_2 < \gamma_1) = 0.464$ [–]; given the step: −4.303, −3.090 [%/yr], 0.305 [–]	given the step: coil 2 −3.120 [%/yr], 0.31179 [–]	mixture: coil 1 −7.04 to −1.94, coil 2 −7.36 to −0.83 [%/yr]
Did a step happen anywhere in the three years?	$\Pr(\text{a step somewhere}) = 0.407$ [–]; top: both coils at window 2, $B = 41.33$ [–]	–	top candidate with window 1's readings left out: $\ln B = -0.11$ [–]

Elsewhere in the record: a rise on both coils into Feb Y1 (Bayes factor 41) rests on window 1's readings alone, with no work on record; not entered.
Action: inspect coil 2's water side for a change left by the pump work (flow, valve positions, strainer); the step, if confirmed, is recoverable and is not fouling. No cleaning decision from the step. The rates for the capacity forecast wait for the archive.
Next measurement: read the logger's archive for Dec Y1 to Jun Y2, four more steady hours a month; the side-by-side check of unit 1's discharge sensor (still open).

The entry names the work and asks for an inspection; it does not say what the pump change did to the coil, which the condition record cannot show. A cleaning decision needs the decision level that Chapter 19, in Volume II, calibrates.

Common mistake

Reading a scan's best window as an event, with the P of a window chosen in advance. A window found by scanning is the best of many, and the evidence for a step somewhere is the average over the candidates; a step in a raw-performance tile is a step in everything the tile depends on, and only the family that carries the sensors says whether the coil stepped.

12.6 The next measurement (step 7)

The report leaves the March step probable and undecided, and its size anywhere from 3.2 to 10.4 %. A further measurement is one more row or block of rows of the residual vector, and what it would do to the step and to its evidence is known before its value is.

12.6.1 The expected evidence of a candidate

A candidate measurement a adds rows with Jacobian lines $\mathbf{J}_a$ on the winner's state, extended where a candidate reads future windows to windows 37–48 as in Chapter 10, 152 entries; with no new rows the extension leaves the spreads of the step and of coil 2's rate at 0.02349 and 0.01673. Its preposterior covariance is equation (4.65), $\mathbf{\Sigma}_a = (\mathbf{\Sigma}^{-1} + \mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a)^{-1}$, and the width it takes off a quantity is $1 - \mathrm{sd}_a/\mathrm{sd}$. For the step the question is also its evidence, and two expectations of it answer two different questions. Both rest on the closed form (12.7) written in the data's own estimate of the step. With $\mathrm{sd}_a(\delta_2)$ the step's spread after the candidate, the data-only variance $\sigma^2_{\mathrm{dat},a}$ takes the prior out, $1/\sigma^2_{\mathrm{dat},a} = 1/\mathrm{sd}_a(\delta_2)^2 - 1/\sigma^2_\delta$, and with $\tilde\delta$ the mean of the step's likelihood, the estimate the rows alone give, the closed form is $\ln B = \frac{1}{2}[\tilde\delta^2\sigma^2_\delta/(\sigma^2_{\mathrm{dat},a}(\sigma^2_{\mathrm{dat},a} + \sigma^2_\delta)) - \ln(1+\psi_a)]$, with $\psi_a = \sigma^2_\delta/\sigma^2_{\mathrm{dat},a}$.

A fresh record of the candidate's design. The first expectation is that of a fresh record built as the present one plus the candidate, with the step at its present estimate $\hat\delta$ and the other seven kinds' Bayes factors held at their present values. Such a record gives $\tilde\delta \sim \mathcal{N}(\hat\delta,\ \sigma^2_{\mathrm{dat},a})$, so $\mathbb{E}[\tilde\delta^2] = \hat\delta^2 + \sigma^2_{\mathrm{dat},a}$, and

$$\mathbb{E}[\ln B] = \frac{1}{2}\left[\frac{\hat\delta^2\,\sigma^2_\delta}{\sigma^2_{\mathrm{dat},a}\,(\sigma^2_{\mathrm{dat},a} + \sigma^2_\delta)} + \frac{\psi_a}{1+\psi_a} - \ln(1+\psi_a)\right]. \tag{12.13}$$

With no new rows it is +3.0266, against this record's +2.7648: a fresh record's expectation is not this record's value.

The preposterior of this record. The second holds the present record fixed and draws only the candidate's rows. The present data-only variance is σ^2_{dat}, with $1/\sigma^2_{\mathrm{dat}} = 1/\mathrm{sd}(\delta_2)^2 - 1/\sigma^2_\delta$, and the present data-only estimate $\tilde\delta_0 = \hat\delta\,(\sigma^2_{\mathrm{dat}} + \sigma^2_\delta)/\sigma^2_\delta$ is kept. The candidate's rows alone have the data variance σ^2_{new}, $1/\sigma^2_{\mathrm{new}} = 1/\sigma^2_{\mathrm{dat},a} - 1/\sigma^2_{\mathrm{dat}}$, and under the posterior predictive of $M_{2,15}$ their estimate is $\tilde\delta_{\mathrm{new}} \sim \mathcal{N}(\hat\delta,\ \mathrm{sd}(\delta_2)^2 + \sigma^2_{\mathrm{new}})$. The combined estimate, $\tilde\delta = \sigma^2_{\mathrm{dat},a}(\tilde\delta_0/\sigma^2_{\mathrm{dat}} + \tilde\delta_{\mathrm{new}}/\sigma^2_{\mathrm{new}})$, then has mean $\tilde\mu$ and variance $\tilde v$, and

$$\begin{aligned} \mathbb{E}[\ln B \mid \mathbf{y}] &= \frac{1}{2}\left[\frac{\sigma^2_\delta\,(\tilde\mu^2 + \tilde v)}{\sigma^2_{\mathrm{dat},a}\,(\sigma^2_{\mathrm{dat},a} + \sigma^2_\delta)} - \ln(1+\psi_a)\right], \\ \tilde\mu = \sigma^2_{\mathrm{dat},a}\left(\frac{\tilde\delta_0}{\sigma^2_{\mathrm{dat}}} + \frac{\hat\delta}{\sigma^2_{\mathrm{new}}}\right), &\qquad \tilde v = \frac{\sigma^4_{\mathrm{dat},a}\,(\mathrm{sd}(\delta_2)^2 + \sigma^2_{\mathrm{new}})}{\sigma^4_{\mathrm{new}}}. \end{aligned} \tag{12.14}$$

With no new rows $\tilde v = 0$, $\tilde\mu = \tilde\delta_0$, and the preposterior is this record's $\ln B$ by the Savage–Dickey ratio, +2.7625. The recommendation stands on this preposterior.

The level a verdict needs. The named-window verdict at 0.95 needs a mean Bayes factor of 19 over the eight kinds. With the other seven kinds held at their present Bayes factors, whose sum is 4.2887, coil 2's Bayes factor must reach $8 \times 19 - 4.2887 = 147.71$, whose logarithm is 4.9953. At the present step, equation (12.13) reaches it at $\mathrm{sd}_a(\delta_2) = 0.01939$.

12.6.2 The candidates

Table 12.13 ranks the candidates. Each row is a measurement; the columns give the step's spread after it and the share taken off, coil 2's rate's spread after it and the share taken off, the expected $\ln B$ of a fresh record, equation (12.13), and the preposterior of this record's $\ln B$, equation (12.14). The logger's archive is more logged hours of past windows, each hour with its own load entry and the same coils; the duty test and the reference test read today's window;

the fourth year adds windows 37–48, with or without permanent heat meters on both coils, the permanent instrument at the place of the question. The last row is the limit of the path known exactly at windows 0–36. The two expectations are compared as bars in Figure 12.7, the horizontal axis the expected $\ln B$, with the level 4.9953 the verdict needs dashed and this record's $\ln B$ by the Savage–Dickey ratio, +2.7625, dotted.

Table 12.13: Candidate measurements for the March step, on the winner's state: the spreads of the step and of coil 2's rate after each, from the present 0.02349 and 0.01673; the expected $\ln B$ of a fresh record of the candidate's design, equation (12.13); and the preposterior of this record's $\ln B$, equation (12.14).

Candidate [–]	sd δ_2 [–]	off [%]	sd γ_2 [yr^{-1}]	off [%]	fresh record [–]	this record [–]
the logger's archive, Dec Y1–Jun Y2 (windows 12–18), four more steady hours each	0.01412	39.9	0.01467	12.3	+10.696	+11.311
the logger's archive, Feb–Apr Y2 (windows 14–16), four more steady hours each	0.01578	32.8	0.01491	10.9	+8.245	+8.583
the logger's archive, one more steady hour of every window 0–36	0.01818	22.6	0.01419	15.2	+5.868	+5.936
side-by-side check of unit 1's discharge sensor now, ten pairs (0.1 K)	0.02349	0.0	0.01632	2.4	+3.029	+2.765
reference condition test now (0.1 K on the four temperatures)	0.02304	2.0	0.01500	10.4	+3.193	+2.948
a water-side duty test of coil 2 now (3.0 kW)	0.02341	0.4	0.01637	2.2	+3.056	+2.795
a fourth year of the record (windows 37–48)	0.02156	8.2	0.01234	26.2	+3.806	+3.634
the fourth year with water-side heat meters on both coils (3.0 kW), permanent	0.02144	8.7	0.01171	30.0	+3.864	+3.699
the path known exactly at windows 0–36 (the floor)	0.00507	78.4	0.01014	39.4	+96.088	+106.060

A step is in the past. No new instrument sees it: the duty test, the reference test and the side-by-side check leave its spread within 2 % of where it is, and a fourth year with permanent meters takes off 8.7 %, through the rate with which the step shares its windows. Only the archive of the months around it narrows it, and four more hours a month for December of year 1 to June of year 2 would decide it: the preposterior of this record's $\ln B$ is 11.3 (10.7 for a fresh record of the same design) against the 4.995 the named-window question needs. The checks and tests made today leave the preposterior near this record's +2.76, because they add nothing the step's row sees. The floor says why only the past months matter. With the path known exactly, the step is $\delta = g_{2,15} - g_{2,14} - \gamma_2 \Delta t$ less one month's wander, with γ_2 read from the other 35 trend rows, so its variance is $\sigma_q^2 + \sigma_q^2/35 = 2.5000 \times 10^{-5} + 7.1429 \times 10^{-7}$ and its spread 0.00507, the prior on the step neglected: the step is then known to the wander of one month.

12.6.3 The recommendation

Read the logger's archive for December of year 1 to June of year 2, four more steady hours a month, an afternoon at a desk with no site visit; it takes 39.9 % off the step's spread, and the preposterior of this record's $\ln B$, 11.3, passes the 4.995 the verdict needs. Inspect coil 2's water side, which the graph does not hold: the inspection settles the cause, the archive the size. The side-by-side check of unit 1's discharge sensor stays open as Chapter 7's entry asked; it does nothing for the step. Of the permanent instruments, the heat meters serve the rate, 30.0 % off

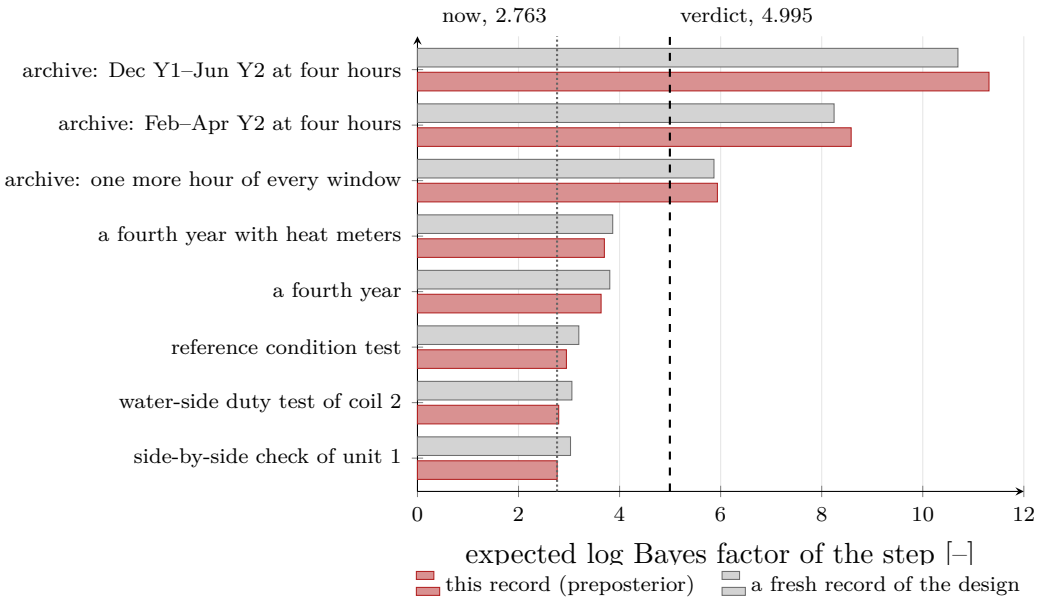


Figure 12.7: The candidate measurements of Table 12.13 ranked by the expected log Bayes factor of the step after each: the preposterior of this record's $\ln B$, equation (12.14) (red), and the expectation for a fresh record of the candidate's design, equation (12.13) (gray). Dashed: $\ln 147.71 = 4.9953$, the level coil 2's Bayes factor needs for the named-window verdict at 0.95; dotted: this record's $\ln B$ by the Savage–Dickey ratio, +2.7625. Not drawn: the floor of the path known exactly, +106.060 and +96.088.

its spread, not the step; their value for the crossing date is Chapter 23's question, in Volume II.

12.6.4 What the analysis showed

The dashboard's step finder put a drop of 0.177 on coil 2 in February and a rise of 0.1656 on coil 1 in May; read on the coils' conductance, the record shows a probable loss of 6.9 % on coil 2 in March or April and nothing on coil 1. The physics takes the load, the season and the other coil out of each window; the step is an entry of the path, and a trend row's normalized residual is its test; and the family carries the sensors that could mimic it, which is why the answer is a probability that a step happened and not a verdict. What the probability adds to the dashboard's "$P = 0.0002$" is the price of the windows it did not name and of the parts it did not consider: named in advance, March is probable; found by scanning, nothing in the record is. The step also changes the rates of Chapter 10, halving coil 2's and leaving open which coil is faster; the slowing that a changing rate on coil 2 would read is this step. Read jointly with coil 1's curvature, as Chapter 14, in Volume II, reads the two open questions in one family, the step's probability is 0.786, because the step and the curvature support each other. The analysis does not settle whether coil 1's path bends, which is Chapter 11's question, or the step itself, which the archive would.

12.6.5 Limits of the model

A constructed record. The condition record is generated from a stated truth, with the stated sensor errors, and step 5 compares the answers with that truth (Section 12.4.5). It is not field data, and the agreement is a check of the method on its own model, not of the model against a real hall.

One permanent step, at the month's resolution. The family holds at most one step, permanent, entering between two monthly windows. Two events, an event later reversed, or an event's day within its month lie outside it.

The priors set the evidence. The step's prior width sets its evidence (Lindley's effect), and the family's prior, equal kinds and even odds, sets the probabilities. The named window rests on a work order outside the record; without it the scan's reading holds.

Chapter 10's limits. The wander σ_q is chosen, not read from the record; the drift is linear from window 20; coil 1's path is a straight line with wander; the readings are independent across windows; and Chapter 5's limits, steady state, pinned airflow, known inputs and a shared β_1, hold unchanged.

The next chapter. The dashboard's February step was found on a raw-performance tile, which mixes the season with the coils; Chapter 13 reads what the season does to such a tile.

12.7 What to retain

Definition 12.1 (Change-point detection)**.** A step is one entry of the fit of the whole record: on a parameter's path, a bias δ on the one trend row into the candidate window, equation (12.1); on a sensor, an offset on its readings from that window on. A family of candidates, the kinds of part that could have stepped times the windows, is ranked by the Bayes factor of each step drawing against the drawing without a step, equation (12.6), and because the step enters the rows linearly the base fit alone gives each in closed form, equation (12.7), through $\psi = \sigma_\delta^2\, \mathbf{a}^\mathsf{T}(\mathbf{I}-\mathbf{K})\mathbf{a}$ and the score $\mathbf{a}^\mathsf{T}\hat{\mathbf{r}}$; for a coil step that is the trend row's normalized residual. At a window named in advance the probability of a step is $\bar{B}_s/(1+\bar{B}_s)$ over the kinds at that window; over the record it is $\bar{B}/(1+\bar{B})$ over every candidate, and the gap between the best candidate and that average is the look-elsewhere charge.

Three conclusions carry beyond the hall. A step in a raw-performance tile is not a step in the asset: the parameter's path, with the inputs taken out and the sensors in the family, is where a step is read. A window named by a record outside the condition record, a work order, is one hypothesis, and only such a window is free of the charge for looking everywhere; a scan has to pay it. And a step shares its windows with the rate, so freeing it moves the rate, and the only rows that narrow a past step are more readings of the months around it.

Table 12.14 lists the terms of the chapter with their expressions, units and the subsection that defines each.

Table 12.14: The terms of change-point detection and where each is defined. Units for the two-unit hall.

Term	Expression	Units	Meaning	Defined in
Step as an entry of the path	eq. (12.1)	–	a bias on one trend row	12.1.1
Candidate family	Table 12.2, eq. (12.2)	–, K, kW	kinds times windows, one step each	12.1.4
Location prior	Π_s in eq. (12.5)	–	named window or uniform	12.1.4
Bayes factor of a step	eq. (12.6)	–	gain in fit less the Occam factor	12.2.1
Closed-form step score	eq. (12.7), ψ	–	every candidate from the base fit; the trend row's normalized residual	12.2.2
Named window	eq. (12.8)	–	one hypothesis, no charge for the others	12.3.1
Location posterior	eq. (12.10)	–	one kind's evidences over the windows	12.3.3
Split of a loss	eq. (12.4), a MAP identity	–	step and rate share the windows	12.3.4
Scan, look-elsewhere charge	eq. (12.12); $\ln B_{\max} - \ln \bar{B} \le \ln 288$	–	a step somewhere, the cost of looking	12.3.6
Influence of one window	$\ln B$ with one window's readings left out	–	what a candidate rests on	12.3.6
False alarm, time to decide	fresh records at the truth	–, window	a candidate from the errors; the first window a verdict holds	12.4.4
Expected evidence of a candidate	eq. (12.13), eq. (12.14)	–	$\ln B$ a fresh record should bring; this record's preposterior	12.6.1
Logger archive	more hours of past windows	–	the only rows that see a past step	12.6.2

12.8 Exercises

Exercise 12.1. *Formulation.* Write the question "was coil 2's step larger than 5 % of its conductance?" in the form of Table 12.1 and evaluate it under $M_{2,15}$. (A prediction under a threshold, equation (2.12) with $f = -\delta_2$ and $f^* = -\ln 0.95$: $\Pr(\delta_2 < \ln 0.95 \mid \mathbf{y}) = \Phi\big((\ln 0.95 + 0.07139)/0.02349\big) = 0.8038$; larger than 10 %, 0.0741.)

Exercise 12.2. *Derivation.* From the closed form (7.22) of a bias on one reading, show that a step on a trend row has $\psi = \sigma_\delta^2(1 - K_{kk})/\sigma_q^2$, and evaluate $\ln B$ for coil 2's row into window 15. (The step's column is $\mathbf{a} = -\mathbf{e}_k/\sigma_q$, the bias column of one reading with σ_k replaced by σ_q and the sign reversed, so $\mathbf{a}^\mathsf{T}(\mathbf{I} - \mathbf{K})\mathbf{a} = (1 - K_{kk})/\sigma_q^2$ and the sign cancels in the square; with $\mathbf{a}^\mathsf{T}(\mathbf{I} - \mathbf{K})\mathbf{a} = (1 - K_{kk})/0.005^2 = 1773.9738$, $\psi = 0.15^2 \times 1773.9738 = 39.9144$, and with $r^\mathrm{N} = -3.0819$, $\ln B = +2.7773$.)

Exercise 12.3. *Derivation.* Show that the Savage–Dickey ratio of a step with prior $\mathcal{N}(0, \sigma_\delta^2)$ and Laplace marginal $\mathcal{N}(\hat{\delta}, s^2)$ gives $\ln B = \ln(s/\sigma_\delta) + \hat{\delta}^2/(2s^2)$, and evaluate it at the winner. (By equation (7.21) $B(M_\mathrm{none} : M_{2,15})$ is the posterior density of δ at zero over the prior density there, $[\mathrm{e}^{-\hat{\delta}^2/(2s^2)}/(\sqrt{2\pi}\, s)]/[1/(\sqrt{2\pi}\, \sigma_\delta)]$; its reciprocal's logarithm is $\ln(s/\sigma_\delta) + \hat{\delta}^2/(2s^2)$, and with $s = 0.02349$ and $\hat{\delta} = -0.07139$, $\ln B = +2.7625$.)

Exercise 12.4. *Application.* Scan coil 2 alone over its 36 windows with a uniform location prior, and give the mean Bayes factor and the probability of a step on coil 2 somewhere, against

the named window. (By equation (12.12) restricted to coil 2, the mean Bayes factor is 1.6480 against $B_{2,15} = 15.8758$; named, Pr = 0.9407, and scanned, Pr = 1.6480/2.6480 = 0.6224.)

Exercise 12.5. *Application.* Evaluate $\ln B_{2,15}$ by the closed form (12.7) for step priors of 0.05, 0.10, 0.30 and 1.0, and find the width that maximizes it. (With $\mathbf{a}^\mathsf{T}\hat{\mathbf{r}} = 129.8064$ and $\mathbf{a}^\mathsf{T}(\mathbf{I} - \mathbf{K})\mathbf{a} = 1773.9738$ held, $\psi = \sigma_\delta^2 \times 1773.9738$: +3.0289, +3.0304, +2.1799 and +1.0057; the largest is +3.1236, at $\sigma_\delta = 0.0692$.)

Chapter 13

Separating Seasonal Effects from Aging

13.1 The summer drop on the dashboard

The hall is three years past its handover. Today is December of year 3, window 36, early summer in the southern hemisphere where the hall stands, and the chilled water has warmed from its winter low. The building management system shows the return air at 27.16°C, 3.52 K warmer than in July, window 31. The condition-monitoring dashboard's hall tile, the apparent conductance, divides the metered IT load by the return air's rise above the chilled-water supply. It has fallen from 12.13 to 10.63 kW/K, by 12.4 % in five months. The trend chart offers two lines through the tile: the line through the last six windows, July to December, falls at 23.7 % a year, and the line through the whole record at 4.8 % a year. The return air's line through the same six windows rises at 9.46 K/yr. The conclusion the chart invites is that the coils are fouling far faster this summer than over the record, and that their cleaning should be brought forward from the October shutdown of year 4.

The asset manager has seen the tile dip before, from July to December of each year: by 6.0 % in year 1 and by 12.2 % in year 2, and asks the analyst to take the season out. The analyst fits the practitioner's seasonal term, one annual harmonic, through the whole record of the tile. The slope becomes −4.7 % a year, and this summer's drop beyond the fitted season is −10.1 %, with a band from −18.2 to −1.1 %: still an anomaly on that reading. Neither chart can say how much of the summer's drop is the ambient, how much the IT load, how much the coils and how much the errors of two readings; whether a fitted season is the season; or whether the coils themselves perform worse in summer. The question of this chapter is

how much of the summer drop in the hall's performance is the coils' loss of conductance, read on the parameter after the known inputs have taken the season out, and what a regression on raw performance, with and without a seasonal term, reads instead.

It is answered on the fit of the whole condition record of Chapter 10, with the raw figures written as quantities of its state, as Chapter 5 wrote the dashboard's index (Section 5.3.3) and Chapter 6 split a held reading.

Jargon: seasonal confound, apparent conductance, summer drop, seasonal term of a regression, the season's leak into a slope, seasonal conductance

A *seasonal confound* is a change of a performance figure that the season makes and a trend reads as aging. The *apparent conductance* of the hall is its metered IT load over the return air's rise above the chilled-water supply, $G_t^{\text{app}} = y_{Q,t}/(y_{T_r,t} - T_s(t))$: a load-normalized raw-performance figure, from which the supply is subtracted and the load divided out, while the other heat, which no meter reads, stays in. The *summer drop* is the change of a raw figure from the winter's coldest supply, in July, to a summer window. A *seasonal term* in a regression, the practitioner's weather normalization, is one annual harmonic (a cosine and a sine) or twelve month-of-year levels added to the design of a trend line, so that the regression estimates the season from the readings. The *season's leak into a slope* is the part of a regression's slope that the season alone produces over the span. A *seasonal conductance* is a season in the parameter itself: a coil whose conductance swings with the time of year.

Key idea

The season is a known input. The physics puts the chilled-water supply and the other heat into every window's rows, so the season leaves the parameter without spending a single entry of the state on it, and the trend is read on what is left. A regression must estimate the season from the same readings that carry the aging, and on a short span, or across a single summer, it cannot tell the two apart.

13.1.1 The season as a known input

Chapter 5 took the season out of one window against commissioning: the supply and the other heat enter each window's physical rows as known inputs, and a coil's conductance, a property of its tubes and fins, does not depend on them. Chapter 6 split a held reading into the season's term, the load's and the coils', for one pair of windows, equation (6.8). Chapter 10 read the coils' rates over the whole record, and its split of the dashboard's slope kept the season together with the load (Section 10.3.3). Here the season is separated from the load and from the coils across the whole record and across the summer itself, and the practitioner's two regressions are set beside the separation.

A raw figure read in two windows changes for four reasons. The known inputs moved, because the season did; the IT load moved; the coils' conductance moved; and each reading carries its own error. The physics computes the first two from the window's inputs and load with the coils held at their commissioned value, and the third from the coils as fitted; the fourth is the readings less the model. A regression on the raw figure sees only their sum. Without a seasonal term it reads the season as part of the trend; with one, it estimates the season as two more columns of its design, from the same readings that also carry the aging and the errors.

> **Principle 13.1.** *A change of performance between two windows is a sum of the known inputs' change, the load's change, the parameter's change and the readings' errors. Only the parameter's change is aging, and only a model that carries the inputs can read it.*

13.1.2 The two-unit hall through its seasons (step 0)

The plant is the two-unit hall of Chapter 5 (Section 5.1.2) through its whole condition record, the stacked graph of Chapter 10 (Figure 10.1), which is cited here and not drawn again in full: 37 copies of the window, each with its five physical rows (5.2) at its own inputs and its free IT load Q_t and log-derates $g_{1,t}$ and $g_{2,t}$; one trend row per coil between consecutive windows; the trend entries γ_1 and γ_2; the recirculation fraction β_1 of rack 1, shared by every window; and the drift rate $\dot{b}_{d,1}$ of unit 1's discharge sensor from window 20.

Modeling choices at the graph. The choices that make the graph are Chapter 5's paragraph (Section 5.1.2): steady state within each window, the airflow pinned at the units' fixed fan speed with nothing controlling the discharge air, the chilled-water supply and the other heat as known inputs, and one effective conductance per coil. This chapter rests on the third, and names it here once. The supply $T_s(t)$ and the other heat $Q_o(t)$ are taken as known exactly in every window: the supply from a precise instrument, the other heat from the schedule of the hall's fans and lighting with the envelope gain. This is what lets the season leave the parameter without spending an entry of the state on it, and it fails if either input is wrong, because the error then goes into the coils' log-derates as aging, or as its opposite. Section 13.4.3 frees each in turn: the other heat to 1.0 and 3.0 kW per window and with its season estimated, and the supply to 0.2 K per window.

The season of the record. The supply and the other heat follow one annual cosine, the same in every year, warmest between January and February and coldest between July and August. The supply runs from 12.55 to 15.45°C and the other heat from 5.10 to 10.90 kW, and the commissioning test ran at their mean, 14.00°C and 8.00 kW. A season of one fixed shape is the best case for a regression that fits one harmonic; Section 13.6.5 returns to it.

The record. The chapter reads all 37 windows of the condition record, the commissioning test and the 36 months from January of year 1 to December of year 3. Two raw figures are what a practitioner trends, read from the hall's own readings and the supply instrument. The *return air* y_{T_r} is the building management system's hall temperature. The *apparent conductance* $G_t^{\text{app}} = y_{Q,t}/(y_{T_r,t} - T_s(t))$ is the metered IT load over the return approach $A_{r,t}$ of Chapter 6, taken from the readings. Neither figure passes through unit 1's discharge sensor, so the drift that Chapter 10 carries does not enter them; it enters the fit, through the split of the hall's conductance between the two coils. Table 13.1 gives seven of the windows, the commissioning test and each July and December. Each row is a window; the first two columns are its known inputs, the next two its return-air and meter readings, and the last two the raw figures formed from them. The supply and the other heat repeat from year to year, while the apparent conductance of each December lies below that of the July before it: by 6.04 %, 12.16 % and 12.44 % in the three years.

Table 13.1: The windows of the condition record shown in this chapter: the commissioning test and each July and December. The return approach is $A_r = y_{T_r} - T_s$, and the apparent conductance $G^{\text{app}} = y_Q/A_r$.

Window	$T_s(t)$ [°C]	$Q_o(t)$ [kW]	y_{T_r} [°C]	y_Q [kW]	A_r [K]	G^{app} [kW/K]
0 (commissioning)	14.00	8.00	22.71	110.92	8.71	12.7348
7 (Jul Y1)	12.55	5.10	21.79	117.74	9.24	12.7440
12 (Dec Y1)	15.06	10.12	25.44	124.28	10.38	11.9738
19 (Jul Y2)	12.55	5.10	22.36	125.08	9.81	12.7517
24 (Dec Y2)	15.06	10.12	26.02	122.75	10.96	11.2005
31 (Jul Y3)	12.55	5.10	23.64	134.56	11.09	12.1347
36 (Dec Y3), today	15.06	10.12	27.16	128.56	12.10	10.6254

The two-unit hall's maintenance log holds, so far, the commissioning test at window 0; Chapter 5's condition assessment and Chapter 6's effort entry at window 18; Chapter 7's assessment at window 25; and three entries at window 36: Chapter 10's rate entry (Section 10.5.2), with coil 1 losing 4.2 and coil 2 6.1 % a year, the drift of unit 1's discharge sensor required, and the side-by-side check of that sensor still open; Chapter 11's entry on the form of coil 1's path; and Chapter 12's entry on coil 2's path. This chapter's entry is the fourth at window 36.

The graph. What this chapter draws on Chapter 10's graph is shown in Figure 13.1, around one window t of the record and its two neighbors. Above each window stand its two known inputs, the chilled-water supply (gray) and the other heat (white), and one season node at the top, the calendar's phase ϕ_t, feeds all of them. Between the windows, one black square per coil is a trend row; the trend entries γ_1, γ_2 (green, below left) act on every trend row. Below window t are the two readings the raw figures use, the return air and the load meter, and the apparent conductance formed from them, both raw figures drawn in purple. The dashed purple node, below right, is question 6's seasonal conductance, which enters the same trend rows.

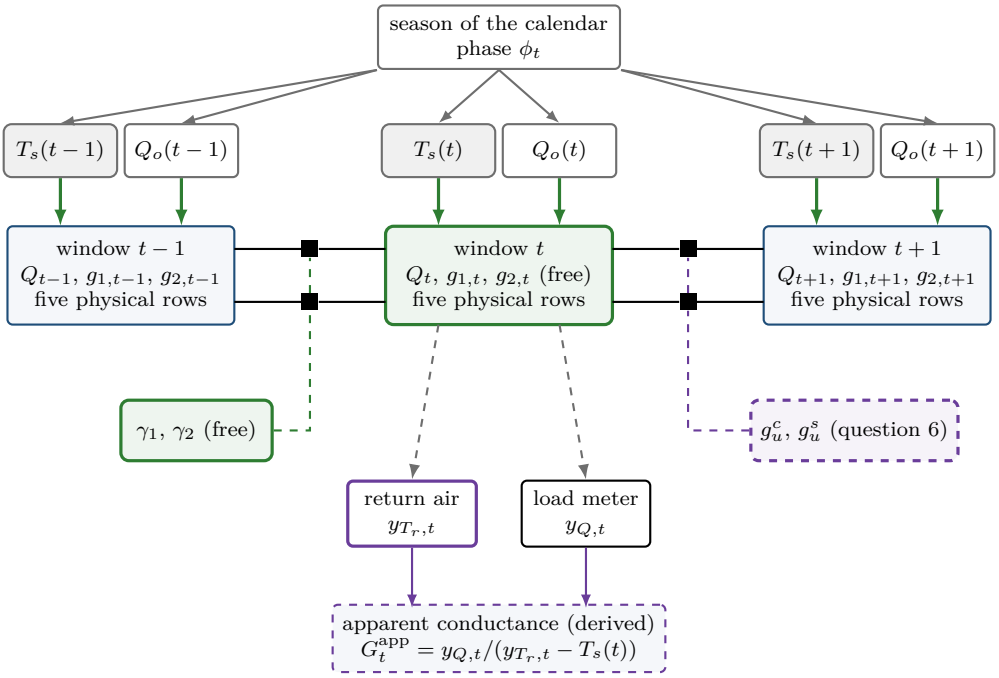


Figure 13.1: The graph of the condition record as this chapter reads it, around one window t; the whole record is Figure 10.1. Free (green): window t's IT load and log-derates and the trend entries γ_1, γ_2; the neighboring windows are drawn as blue blocks. Known inputs, pinned in each window's fit: the chilled-water supply $T_s(t)$ (gray) and the other heat $Q_o(t)$ (white), all fed by one season node, the calendar's phase ϕ_t (thin gray arrows): the season enters every window and costs no entry. Black squares: the trend rows (10.5), one per coil between consecutive windows, coil 1's above and coil 2's below; γ_u acts on every trend row of coil u (one gap's links drawn). Observed (dashed gray): the two readings of window t that the raw figures use; the other three readings and the drift on unit 1's discharge reading are those of Figure 10.1. Purple: the raw figures, the return air and the apparent conductance formed from it and the meter. Dashed purple: question 6's seasonal conductance, an annual cosine and sine of each coil's log-derate on top of its trend, which enters every trend row of its coil (Section 13.1.4).

The state. The state is Chapter 10's stacked state (10.3), 115 entries: each window's $(Q_t, g_{1,t}, g_{2,t})$, then β_1, γ_1, γ_2 and $\dot{b}_{d,1}$. The raw figures add no entry, because each is a function of one window's three entries at that window's inputs (Section 13.1.4). The hall figures see the coils through their sum, the hall's total conductance

$$G_{\text{tot},t} = G_{1,t} + G_{2,t} = \bar{G}_1 \, \mathrm{e}^{g_{1,t}} + \bar{G}_2 \, \mathrm{e}^{g_{2,t}}, \tag{13.1}$$

in kW/K, with $\bar{G}_u = 7.0\,\text{kW/K}$ the ratings. Question 6 adds four entries, in a second drawing of the record (Section 13.1.4).

13.1.3 The asset manager's questions (step 1)

The asset manager asks six questions of the record, and step 1 writes each as a probability over the stacked state, with the unknown on the left of the bar and the record's readings and prior rows on the right. Table 13.2 lists them. The first column is the question in the asset manager's words, the first of them the drill question; the second the general form of its category from Chapter 2; the third its mathematical form for the hall; the fourth its category; the fifth the computation of Section 2.3.6 that answers it; and the last the subsection where it is answered. "The summer drop" is the change from July of year 3, window 31, to today, and "performance" is each of the two raw figures in turn. The first two questions are answered together. Three questions are not rows. Each coil's rate is Chapter 10's, cited in one line of Section 13.3.4; whether coil 1's rate is changing is Chapter 11's question, and whether coil 2's path left its line Chapter 12's; and when the capacity margin closes, where the season returns through the capacity scenario's summer supply, is Chapter 14's. Which measurement would sharpen the summer's answer asks nothing new of the record; it is the next measurement, Section 13.6.

Table 13.2: The questions asked of the two-unit hall's condition record at window 36. $\mathbf{y}$ is the 185 readings and the prior rows of Chapter 10's base fit; y is a raw figure, y_{T_r} or $\ln G^{\text{app}}$; the terms of its change are its season, IT load, coils and reading errors between the two windows, defined in step 2; $\mathbf{w}^{\text{line}}$ and $\mathbf{w}^{\text{seas}}$ are the rows of weights that make a least-squares slope, without and with an annual harmonic, from the readings $\mathbf{y}^{\text{raw}}$ of windows 0 to t_{as}, also defined in step 2; M_{base} and M_{sea} the base drawing and the seasonal-conductance drawing.

Asset manager's question	General form	Mathematical form	Category	Computation	Answered in
Is the summer drop in performance degradation or just the ambient?	moments and credible interval of a derived quantity, (2.10), (2.11)	$y_{36} - y_{31} = s_{\text{season}} + s_{\text{load}} + s_{\text{coils}} + s_{\text{err}}$, each term's $p(\cdot \mid \mathbf{y})$; on $\ln G^{\text{app}}$, $s_{\text{coils}} = \ln(G_{\text{tot},36}/G_{\text{tot},31})$	estimation	trend	13.3.1
Is more of the drop the coils than the season?	$\Pr(f(\mathbf{z}) < f^* \mid \mathbf{y})$, (2.12)	$\Pr(s_{\text{coils}} < s_{\text{season}} \mid \mathbf{y})$ on $\ln G^{\text{app}}$	prediction under a threshold	trend	13.3.1
What does the trend line through the raw performance read, and how much of it is the season?	moments of a derived quantity, (2.10)	$s^{\text{line}} = \mathbf{w}^{\mathsf{T}}\mathbf{y}^{\text{raw}}$, $\mathbf{w} = \mathbf{w}^{\text{line}}$, over windows 0 to t_{as}, split by the same map into season (the leak), load, coils and errors	estimation	trend	13.3.2
Does a seasonal term in the regression take the season out?	moments of a derived quantity, (2.10)	$s^{\text{seas}} = \mathbf{w}^{\mathsf{T}}\mathbf{y}^{\text{raw}}$, $\mathbf{w} = \mathbf{w}^{\text{seas}}$, split the same way; its fitted harmonic against the known inputs'	estimation	trend	13.3.3
How fast is the hall's conductance falling, with the season taken out by the physics?	$\Pr(\gamma > 0 \mid \mathbf{y})$, (2.21) (first line); (2.11)	$p(\Gamma \mid \mathbf{y})$, $\Gamma = \ln(G_{\text{tot},36}/G_{\text{tot},0})$ over $36\,\Delta t$, as $100(\mathrm{e}^{\Gamma} - 1)$ %/yr	estimation	trend	13.3.4
Does the coils' conductance itself fall in summer?	$B_{AB} = p(\mathbf{y} \mid M_A)/p(\mathbf{y} \mid M_B)$, (2.14)	$B(M_{\text{sea}} : M_{\text{base}})$; $p(g_u^c, g_u^s \mid \mathbf{y})$	hypothesis comparison	evidence	13.3.5

13.1.4 The evidence as rows (step 2)

The rows are those of Chapter 10's base fit, unchanged: 261 rows on the 115 entries. They are the five readings of each of the 37 windows on the reduced state at that window's inputs, 185 rows, with unit 1's discharge row carrying the drift (10.2) from window 20; the rating prior on the two commissioning entries and the survey prior on β_1; the 72 trend rows (10.5) at $\sigma_q = 0.005$ per window; and the drift's prior row, of spread $\sigma_{\dot b} = 2.0\,\mathrm{K/yr}$. This chapter adds no row to them. It reads the season out of them, and it writes the raw figures as quantities of their state.

Where the season enters the rows. Each window's predicted readings are Chapter 5's forward equations (5.3) at that window's $T_s(t)$ and $Q_o(t)$. For the two raw figures, with the hall's total conductance (13.1) and the second form following from the first,

$$T_{r,t}(\mathbf{z}) = T_s(t) + \frac{Q_t + Q_o(t)}{G_{\mathrm{tot},t}}, \qquad \ln G_t^{\mathrm{app}}(\mathbf{z}) = \ln \frac{Q_t}{T_{r,t} - T_s(t)} = \ln G_{\mathrm{tot},t} + \ln Q_t - \ln\bigl(Q_t + Q_o(t)\bigr). \tag{13.2}$$

Neither input is an entry of the state, and neither has a column of $\mathbf{J}$: the season is in the value of each row and in its slopes, not in what the fit estimates. The supply cancels from the apparent conductance, which is why the practitioner forms it; the other heat does not: the tile's season is the other heat alone, which lowers it by the share $Q_o/(Q + Q_o)$ of the heat that the meter does not see.

The split of a change between two windows. Write Y_t for the model value of a raw figure at window t, $Y_t(\mathbf{z})$ of equation (13.2), and y_t for its reading. Write $\check{Y}_t$ for its value with the coils held at the commissioned ones, $g_{u,t} = g_{u,0}$, at the window's own load and inputs, and $\check{Y}_t^{\mathrm{m}}$ for that value at the season's mean inputs, 14.00°C and 8.00 kW, window 0's. In Chapter 6's order, season and load at the commissioned coils and the coils last, the change of a raw figure from window t_1 to window t_2 splits into

$$\begin{gathered} y_{t_2} - y_{t_1} = s_{\mathrm{season}} + s_{\mathrm{load}} + s_{\mathrm{coils}} + s_{\mathrm{err}}, \\ s_{\mathrm{load}} = \bigl[\check{Y}^{\mathrm{m}}\bigr]_{t_1}^{t_2}, \quad s_{\mathrm{season}} = \bigl[\check{Y} - \check{Y}^{\mathrm{m}}\bigr]_{t_1}^{t_2}, \quad s_{\mathrm{coils}} = \bigl[Y - \check{Y}\bigr]_{t_1}^{t_2}, \quad s_{\mathrm{err}} = \bigl[y - Y\bigr]_{t_1}^{t_2}, \end{gathered} \tag{13.3}$$

with $[\,\cdot\,]_{t_1}^{t_2}$ the value at window t_2 less the value at window t_1. The first three terms are quantities of the state; the fourth is the readings less the model. On the return air, by equation (13.2),

$$\begin{gathered} s_{\mathrm{season}} = \bigl[T_s(t_2) - T_s(t_1)\bigr] + \frac{Q_o(t_2) - Q_o(t_1)}{G_{\mathrm{tot},0}}, \qquad s_{\mathrm{load}} = \frac{Q_{t_2} - Q_{t_1}}{G_{\mathrm{tot},0}}, \\ s_{\mathrm{coils}} = \bigl(Q_{t_2} + Q_o(t_2)\bigr)\Bigl(\frac{1}{G_{\mathrm{tot},t_2}} - \frac{1}{G_{\mathrm{tot},0}}\Bigr) - \bigl(Q_{t_1} + Q_o(t_1)\bigr)\Bigl(\frac{1}{G_{\mathrm{tot},t_1}} - \frac{1}{G_{\mathrm{tot},0}}\Bigr), \end{gathered} \tag{13.4}$$

all in K, the season's term the known inputs' change at the commissioned total $G_{\mathrm{tot},0}$. On the apparent conductance, with $\bar{Q}_o = 8.00\,\mathrm{kW}$ the mean other heat,

$$\begin{gathered} s_{\mathrm{season}} = \ln \frac{Q_{t_2} + \bar{Q}_o}{Q_{t_2} + Q_o(t_2)} - \ln \frac{Q_{t_1} + \bar{Q}_o}{Q_{t_1} + Q_o(t_1)}, \qquad s_{\mathrm{load}} = \ln \frac{Q_{t_2}}{Q_{t_2} + \bar{Q}_o} - \ln \frac{Q_{t_1}}{Q_{t_1} + \bar{Q}_o}, \\ s_{\mathrm{coils}} = \ln \frac{G_{\mathrm{tot},t_2}}{G_{\mathrm{tot},t_1}}, \end{gathered} \tag{13.5}$$

dimensionless. On $\ln G^{\mathrm{app}}$ the split is additive in the logarithm, and the coils' term does not depend on the order: it is the hall's log-conductance change, whatever the load and the season did. The season's and the load's terms do depend on it, on the tile through the load at which the season is taken, and on the return air through the conductance as well. With both taken at the commissioned coils, the return air's season term is a quantity of the known inputs and of $G_{\mathrm{tot},0}$ alone, the same in every year.

A regression as a fixed linear map. The practitioner's trend of a raw figure is a least-squares fit over windows 0 to t_{as}. Write $\mathbf{y}^{\text{raw}}$ for the figure's readings in those windows, as a column, and $\mathbf{X}$ for the design, whose row for window t is

$$\begin{gathered}(1,\ t\,\Delta t) \quad \text{(the line)}, \qquad (1,\ t\,\Delta t,\ \cos\phi_t,\ \sin\phi_t) \quad \text{(the seasonal regression)},\\ \phi_t = \frac{2\pi\,(\text{mo}_t - 0.5)}{12}, \qquad \hat{\mathbf{k}} = \mathbf{X}^{+}\mathbf{y}^{\text{raw}}, \quad \mathbf{X}^{+} = (\mathbf{X}^{\mathsf{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathsf{T}}, \qquad s^{\text{reg}} = \mathbf{w}^{\mathsf{T}}\mathbf{y}^{\text{raw}},\end{gathered} \tag{13.6}$$

with mo_t the month index of window t (0 for January), so that $\cos\phi_t = 1$ between January and February, the plant's own phase; window 0, the commissioning test at the season's mean inputs, takes zero seasonal regressors. The coefficients $\hat{\mathbf{k}}$ are the intercept, the slope s in units per year, and for the seasonal regression the cosine and sine coefficients k_c and k_s; the slope row $\mathbf{w}^{\mathsf{T}}$ is the second row of $\mathbf{X}^{+}$, written $\mathbf{w}^{\text{line}}$ or $\mathbf{w}^{\text{seas}}$, and s^{line} or s^{seas} is the slope. The slope's standard error is the practitioner's, from the residual scatter, $\hat{\sigma}^2\,[(\mathbf{X}^{\mathsf{T}}\mathbf{X})^{-1}]_{22}$ under the root, with $\hat{\sigma}^2$ the residual sum of squares over its degrees of freedom. Because $\mathbf{X}^{+}\mathbf{X}$ is the identity, the slope row takes every column of the design to zero but the slope's own: a constant leaves the slope unchanged, and so, in the seasonal regression, does any exact harmonic $k_c\cos\phi_t + k_s\sin\phi_t$. Applied to the split, written window by window,

$$s^{\text{reg}} = \mathbf{w}^{\mathsf{T}}\mathbf{s}_{\text{season}} + \mathbf{w}^{\mathsf{T}}\mathbf{s}_{\text{load}} + \mathbf{w}^{\mathsf{T}}\mathbf{s}_{\text{coils}} + \mathbf{w}^{\mathsf{T}}\mathbf{s}_{\text{err}}, \tag{13.7}$$

with $\mathbf{s}_{\text{season}}$ the column of $\check{Y}_t - \check{Y}_t^{\text{m}}$ over the windows, $\mathbf{s}_{\text{load}}$ of $\check{Y}_t^{\text{m}}$, $\mathbf{s}_{\text{coils}}$ of $Y_t - \check{Y}_t$ and $\mathbf{s}_{\text{err}}$ of $y_t - Y_t$. The first term is the season's leak into the slope: a number of the known inputs and of the commissioned state, with no part of the coils' aging in it.

Which row acts on which entry. The Jacobian's pattern is Chapter 10's, equation (10.6): each window's reading rows touch only its own $(Q_t, g_{1,t}, g_{2,t})$, rack 1's also β_1 and unit 1's discharge row from window 20 also $\dot{b}_{d,1}$; each trend row touches $g_{u,t}$, $g_{u,t+1}$ and γ_u. The questions' gradients are read off the forms above. With $*$ for an entry that can be nonzero and the columns grouped by window,

	Q_0	$g_{1,0}$	$g_{2,0}$	1–30	Q_{31}	$g_{1,31}$	$g_{2,31}$	32–35	Q_{36}	$g_{1,36}$	$g_{2,36}$	β_1	γ_1	γ_2	$\dot{b}_{d,1}$
s_{coils} on $\ln G^{\text{app}}$	0	0	0	0	0	$*$	$*$	0	0	$*$	$*$	0	0	0	0
s_{season} on $\ln G^{\text{app}}$	0	0	0	0	$*$	0	0	0	$*$	0	0	0	0	0	0
s_{season} on T_r	0	$*$	$*$	0	0	0	0	0	0	0	0	0	0	0	0
Γ	0	$*$	$*$	0	0	0	0	0	0	$*$	$*$	0	0	0	0
a regression's terms	$*$	$*$	$*$	$*$	$*$	$*$	$*$	$*$	$*$	$*$	$*$	0	0	0	0

(13.8)

These lines are not rows: they are the gradients of the quantities the questions ask for, with Γ the hall's rate of question 5. Two facts are read here, before any number. The coils' term of the summer drop on the tile touches four entries, the two coils in the drop's two windows, so its variance is a 4×4 block of $\boldsymbol{\Sigma}$; the season's term touches no log-derate of those windows, so the coils' uncertainty does not enter it at all. And a regression's terms touch every window's entries, weighted by $\mathbf{w}$. No question's gradient touches γ_u, β_1 or $\dot{b}_{d,1}$, which reach the answers only through their covariances with the windows' entries.

Modeling choices at the rows. Four choices enter here. The summer drop is read from window 31, July of year 3 and the winter's coldest supply, 12.55°C, to today: this is the drop the dashboard shows, and it fails if the start is picked to suit the answer; Section 13.4.3 moves it to windows 30 and 32. The practitioner's regressions are fitted by ordinary least squares on the readings, with standard errors from their own residual scatter, the way a dashboard fits them, and never on the state: this keeps the naive baseline what the asset manager sees, and it is the

baseline the analysis is set against, not a rival estimate of the state. The seasonal regression is given the plant's own phase, with window 0 at zero regressors: this is the regression's best case, and a phase it had to estimate would leave it worse off. And question 6's seasonal conductance is a second drawing of the record, in which each coil's log-derate carries $g_u^c \cos\phi_t + g_u^s \sin\phi_t$ on top of its trend, entered through the trend rows' expected increment,

$$r_{q,u,t}(\mathbf{z}) = \frac{g_{u,t+1} - g_{u,t} - \gamma_u\,\Delta t - g_u^c(\cos\phi_{t+1} - \cos\phi_t) - g_u^s(\sin\phi_{t+1} - \sin\phi_t)}{\sigma_q}, \qquad \frac{g_u^c}{\sigma_{\text{sea}}},\ \frac{g_u^s}{\sigma_{\text{sea}}}, \tag{13.9}$$

for $u = 1, 2$ and $t = 0, \ldots, 35$, with four prior rows of spread $\sigma_{\text{sea}} = 0.05$ on the amplitudes. The prior allows a summer derate of up to about 5 %, a wet coil with condensate on its fins; the cosine amplitude is the summer's part, a derate when it is negative. This is what lets the record test whether the coils, and not only the inputs, carry a season, and it fails if a coil's season has another shape than one annual harmonic. Section 13.3.5 varies the prior from 0.005 to 0.1. The base fit's own choices, the straight line with its wander at $\sigma_q = 0.005$, the drift's onset at window 20 and the drift's prior, are Chapter 10's, and are not chosen again.

13.2 The posterior of the state (step 3)

All six questions of Table 13.2 are read from one posterior, that of Chapter 10's base fit, and question 6 also from the posterior of one more drawing of the same record.

13.2.1 The base fit read for the seasons

The posterior is Chapter 10's, computed there once and read here: the residual vector of the 261 rows on the stacked state (10.3), the Gauss–Newton iterations of the normal equations (5.9) on it (Table 10.3), and the Laplace covariance $\boldsymbol{\Sigma} = (\mathbf{J}^\mathsf{T}\mathbf{J})^{-1}$ at the MAP state, the Gauss–Newton form that drops the curvature term $\mathbf{C}$ of equation (4.44); Section 13.4.2 sets the full Hessian beside it. With every row normalized the information matrix is $\boldsymbol{\Lambda} = \mathbf{J}^\mathsf{T}\mathbf{J}$; with unnormalized rows and their weights it is $\mathbf{J}^\mathsf{T}\mathbf{W}\mathbf{J}$. At the MAP state twice the objective is $2\mathcal{L}(\hat{\mathbf{z}}) = 154.9755$, and the entries of Table 10.4 that the answers lean on are the two rates, $\hat{\gamma}_1 = -0.04240\,\text{yr}^{-1}$ with standard deviation 0.01463 and $\hat{\gamma}_2 = -0.06253\,\text{yr}^{-1}$ with 0.01323, and the drift rate, 0.6837 K/yr.

The hall's conductance. Carried from the entries through its gradient, $G_{u,t}$ in the column of $g_{u,t}$, the hall's total conductance (13.1) at the three windows the answers read is $G_{\text{tot},0} = 13.7893\,\text{kW/K}$ with standard deviation 0.1212, $G_{\text{tot},31} = 12.0925$ with 0.0941 and $G_{\text{tot},36} = 11.7897$ with 0.1193 kW/K. The four log-derates of the summer drop's two windows have the covariance block and correlations

$$\boldsymbol{\Sigma}_{31,36} = \text{Cov}\big((g_{1,31}, g_{2,31}, g_{1,36}, g_{2,36}) \mid \mathbf{y}\big) = 10^{-4}\begin{bmatrix} 3.4686 & -1.4401 & 3.9229 & -1.8013 \\ -1.4401 & 1.6784 & -1.7969 & 1.6585 \\ 3.9229 & -1.7969 & 5.6927 & -2.3487 \\ -1.8013 & 1.6585 & -2.3487 & 2.8577 \end{bmatrix},$$

$$\text{correlations} \quad \begin{bmatrix} 1 & -0.5969 & 0.8828 & -0.5721 \\ -0.5969 & 1 & -0.5813 & 0.7573 \\ 0.8828 & -0.5813 & 1 & -0.5823 \\ -0.5721 & 0.7573 & -0.5823 & 1 \end{bmatrix}, \tag{13.10}$$

dimensionless. The correlations read by sign, and the rows fix each sign. The two coils of one window are correlated negatively, −0.5969 at window 31 and −0.5823 at window 36: the return

air sees the sum $G_1 + G_2$, so a lower coil 1 is paid for by a higher coil 2 at the same reading, and the record knows the hall's total better than its split. One coil's two windows are correlated positively, +0.8828 for coil 1 and +0.7573 for coil 2: five trend rows tie them, each allowing a month's wander of σ_q, so most of each window's uncertainty is shared, and a difference of the two removes it. And one coil at window 31 against the other at window 36 is correlated negatively, −0.5721 and −0.5813: the within-window anticorrelation, carried along the trend rows. The hall's change over the five months takes each window's two coils with the same sign and each coil's two windows with opposite signs, so both kinds of correlation narrow it; a single coil's change keeps the within-window uncertainty of the split. Section 13.3.1 computes both.

The record against the fitted path. The record is set against the posterior in Figure 13.2, window by window. The upper panel is the return air: the dots are the readings, the solid line the model value at the MAP state, equation (13.2), and the dashed line the value $\check{T}_{r,t}$ with the coils held at the commissioned ones, at each window's own load and inputs, which carries the season and the IT load alone. The lower panel is the apparent conductance in the same three forms, with the hall's total conductance $G_{\mathrm{tot},t}$ and its 5–95 % band beside them. Thin vertical lines mark each July and December. The dashed curves swing with the season every year, and the readings follow them; the solid curves leave them as the coils age, the return air reaching 27.04°C at window 36 against 25.30°C for commissioned coils. The hall's conductance falls from 13.79 to 11.79 kW/K with no season in it at all: the season is in the inputs and in the tile, not in the parameter.

13.2.2 The second drawing and the full posterior

Question 6's drawing M_{sea} is the base drawing M_{base} with the four amplitudes of equation (13.9) freed: 119 entries, the base fit's rows with its trend rows replaced by those of equation (13.9), and four prior rows of spread 0.05. It is fitted by the same normal equations, from the base fit's MAP state with the amplitudes at zero. At its MAP state $2\mathcal{L} = 150.6304$, against the base drawing's 154.9755, and its Laplace log evidence, equation (7.7), is −73.9044, against the base drawing's −68.9335. Its amplitudes and their ratio are read in Section 13.3.5.

The full posterior by reweighted draws. Beside the Laplace normal, the reported posterior, stands the full posterior of the same rows, by the reweighted draws of Chapter 5 (Section 5.2.4): draws from the Laplace normal of the base fit, each weighted by equation (5.13) with the objective evaluated through the forward equations, and the evidence of each drawing by equation (7.12). For this record $N_{\mathrm{d}} = 100\,000$ draws from a seeded generator give an effective sample size of 98 841, 98.8 % of N_{d}, and a largest weight of 2.51×10^{-5}; the seasonal drawing's draws keep 98.6 %. The Laplace normal is close to the posterior in both drawings, and no draw dominates. Section 13.4.1 states the Monte Carlo errors.

13.3 The answers (step 4)

Every answer is read from the pair $(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ of Section 13.2.1, and that value is the one reported. A derived quantity f is carried by its gradient $\mathbf{g} = \partial f / \partial \mathbf{z}$ at $\hat{\mathbf{z}}$, formed by differences on the state where it has no closed form, with variance $\mathbf{g}^{\mathsf{T}} \boldsymbol{\Sigma} \mathbf{g}$, the delta method (4.53). The full posterior of Section 13.2.2 stands beside each value, a median beside a median. The hall's rate is reported as $100(\mathrm{e}^{\Gamma} - 1)$ %/yr, Chapter 10's form (10.1), and a regression's slope s of $\ln G^{\mathrm{app}}$ the same way; a change over the five months is reported as $100(\mathrm{e}^{x} - 1)$ % of conductance.

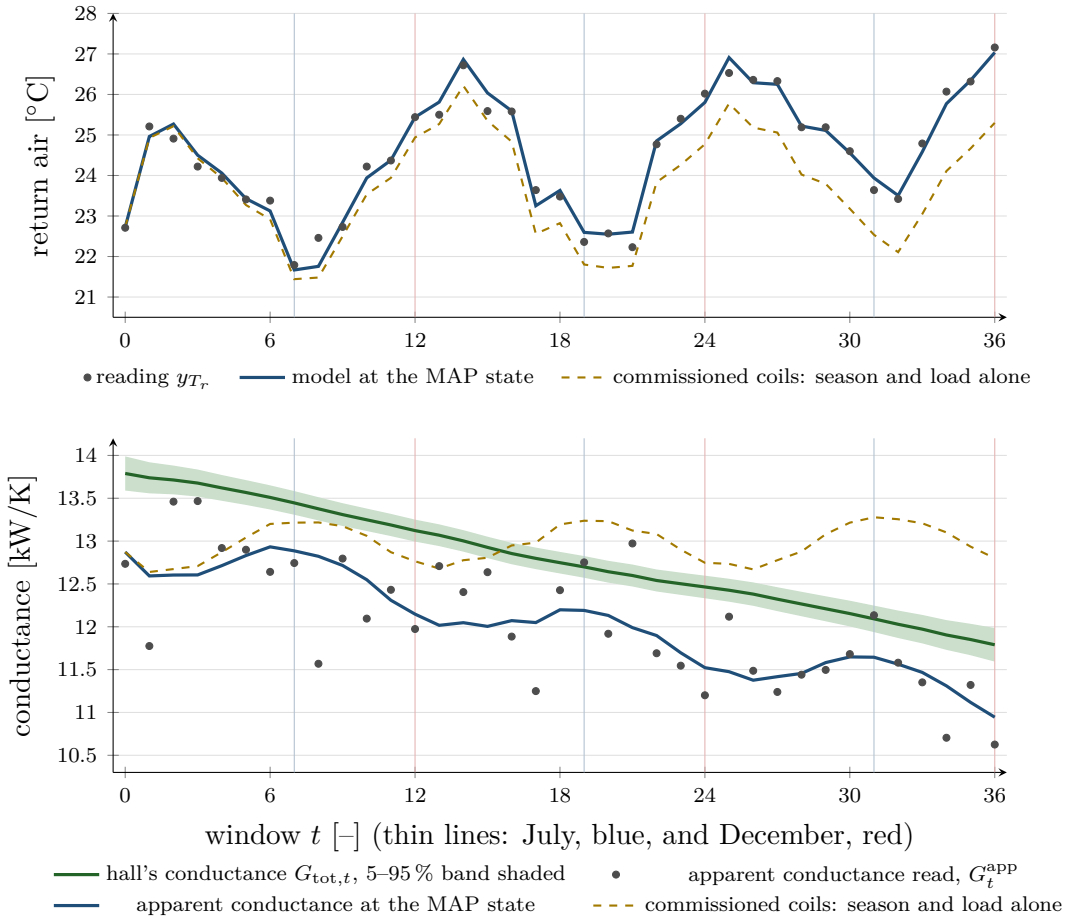


Figure 13.2: The condition record against the posterior of the base fit, window by window. Upper: the return air, read (dots), at the MAP state (solid) and with the coils at their commissioned values at each window's load and inputs (dashed). Lower: the apparent conductance in the same three forms, and the hall's total conductance $G_{\mathrm{tot},t}$ at the MAP state with its 5–95 % band (± 1.6449 standard deviations, shaded). Thin vertical lines: each July (blue) and December (red).

13.3.1 The summer drop, season against coils

Asset manager's questions. Is the summer drop in performance degradation or just the ambient? Is more of the drop the coils than the season? The first asks how the drop of each raw figure from July to December divides among its four sources; the second whether the coils' part of the tile's drop is larger than the season's.

General form. An estimation question about a derived quantity is summarized by its moments, equation (2.10),

$$\mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}] = \int f(\mathbf{z})\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}, \qquad \mathrm{Var}(f(\mathbf{z}) \mid \mathbf{y}) = \int \big(f(\mathbf{z}) - \mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}]\big)^2\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z},$$

and by its credible interval, equation (2.11), at the levels 0.05 and 0.95. A prediction under a threshold is a tail of a derived quantity under the posterior, equation (2.12),

$$\Pr(f(\mathbf{z}) < f^* \mid \mathbf{y}) = \int \mathbf{1}\big[f(\mathbf{z}) < f^*\big]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}.$$

Specialized form. Question 1 is equations (2.10) and (2.11) with $f(\mathbf{z})$ each of the first three terms of the split (13.3) with $t_1 = 31$ and $t_2 = 36$, on the return air by equation (13.4) and on the tile by equation (13.5), and $p(\mathbf{z} \mid \mathbf{y})$ the Laplace normal $\mathcal{N}(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ of Section 13.2.1. The reading errors' term is the readings' change less the model's; its variance is that of the model's change, because the readings are fixed numbers. Each term's gradient lies in the columns the pattern (13.8) allows. On the tile the coils' term has a closed gradient: with $\partial G_{u,t}/\partial g_{u,t} = G_{u,t}$,

$$\begin{gathered} s_{\text{coils}} = \ln G_{\text{tot},36} - \ln G_{\text{tot},31}, \qquad \mathbf{g}_{31,36} = \Big(-\frac{G_{1,31}}{G_{\text{tot},31}},\ -\frac{G_{2,31}}{G_{\text{tot},31}},\ \frac{G_{1,36}}{G_{\text{tot},36}},\ \frac{G_{2,36}}{G_{\text{tot},36}}\Big)^{\mathsf{T}}, \\ \mathrm{Var}(s_{\text{coils}} \mid \mathbf{y}) = \mathbf{g}_{31,36}^{\mathsf{T}} \boldsymbol{\Sigma}_{31,36}\, \mathbf{g}_{31,36}, \end{gathered} \tag{13.11}$$

on the entries $(g_{1,31}, g_{2,31}, g_{1,36}, g_{2,36})$ of equation (13.10), each window's coils entering by their shares of the hall's conductance, and the gradient zero elsewhere. Its 5–95 % band is $\hat{s}_{\text{coils}} \pm 1.6449\, \mathrm{sd}(s_{\text{coils}} \mid \mathbf{y})$, mapped to a change of conductance by $100(\mathrm{e}^{s} - 1)$ %, which passes quantiles unchanged. Question 2 is equation (2.12) with $f(\mathbf{z}) = s_{\text{coils}} - s_{\text{season}}$ on the tile and $f^* = 0$: both terms are losses, and the event is that the coils' term is the more negative, with

$$\Pr(s_{\text{coils}} < s_{\text{season}} \mid \mathbf{y}) = \Phi\Big(\frac{-(\hat{s}_{\text{coils}} - \hat{s}_{\text{season}})}{\mathrm{sd}(s_{\text{coils}} - s_{\text{season}} \mid \mathbf{y})}\Big).$$

One choice enters here: the comparison is made on the tile the asset manager reads, where both terms are parts of one logarithm and the order of the split does not matter. On the return air the season's term is a temperature at the commissioned coils, and a comparison there would carry the split's order with it.

Building blocks. The pair $(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$ and the block $\boldsymbol{\Sigma}_{31,36}$ of Section 13.2.1; the conductances at the MAP state, the entries of $\mathbf{g}_{31,36}$; the split (13.3) and its two forms; the variance (13.11); the normal distribution function Φ.

The four parts of the drop. Table 13.3 splits each raw figure's change from window 31 to window 36. Each row is a term of equation (13.3); the columns give its value on the return air and on the tile, each with its posterior standard deviation and its share of the reading's change.

On the return air, +3.5200 K is +2.8735 K of season, 81.6 % of the change: the supply's known +2.5095 K and the other heat's +0.3640 K through the commissioned coils. The season's

Table 13.3: The summer drop, window 31 (Jul Y3) to window 36 (Dec Y3), split by equation (13.3) at the MAP state of the base fit, with posterior standard deviations in parentheses.

Term	Return air [K]	Share [%]	$\ln G^{\text{app}}$ [–]	Share [%]
Reading change, $y_{36} - y_{31}$	+3.5200	100	−0.13282	100
Season, s_{season}	+2.8735 (0.0032)	81.6	−0.03596 (0.00042)	27.1
of it the supply, known; the other heat	+2.5095; +0.3640	–	–	–
IT load, s_{load}	−0.1112 (0.2186)	−3.2	−0.00067 (0.00131)	0.5
Coils, s_{coils}	+0.3354 (0.0852)	9.5	−0.02536 (0.00736)	19.1
Reading errors, s_{err}	+0.4223 (0.2447)	12.0	−0.07083 (0.00785)	53.3

standard deviation is only 0.0032 K, because by equation (13.4) it depends on the state through $G_{\text{tot},0}$ alone. The IT load fell a little between the two windows, −0.1112 K, with the loads' own uncertainty; the coils are +0.3354 K, band +0.195 to +0.475 K; and the two windows' reading errors are +0.4223 K. On the tile the −0.13282 of the reading, −12.44 %, is −0.03596 of season, 27.1 %, which depends on the two loads only; −0.02536 of coils, 19.1 %; and −0.07083 of reading errors, 53.3 %.

The coils' five months. By equation (13.11), the gradient at the MAP state is $\mathbf{g}_{31,36} = (-0.52022,\ -0.47978,\ +0.51973,\ +0.48027)^{\mathsf{T}}$, and with the block (13.10),

$$\operatorname{Var}(s_{\text{coils}} \mid \mathbf{y}) = \mathbf{g}_{31,36}^{\mathsf{T}} \boldsymbol{\Sigma}_{31,36}\, \mathbf{g}_{31,36} = 5.4106 \times 10^{-5} = 0.00736^2.$$

The hall's conductance changed over the five months by $\ln(G_{\text{tot},36}/G_{\text{tot},31}) = -0.02536$, standard deviation 0.00736, 5–95 % band [−0.03746, −0.01326]: a loss of 2.50 % of conductance, with the band from 1.32 to 3.68 %. The full posterior gives −2.50 % and a band from −3.67 to −1.30 %. The probability that the hall lost conductance over the five months is 0.9997. Each coil's own five months is a difference along one coil: coil 1's $g_{1,36} - g_{1,31} = -0.02631$, standard deviation 0.01147, −2.60 %, and coil 2's −0.02433, standard deviation 0.01104, −2.40 %. Both are wider than the hall's 0.00736, because each keeps the within-window uncertainty of the split that the hall's total cancels, as Section 13.2.1 read from the signs of $\boldsymbol{\Sigma}_{31,36}$. These are the base drawing's readings: its straight line with wander pulls a five-month change toward the record's rate, and Section 13.4.5 shows that on this record it reads the hall's five months short of the truth and the split between the coils further off still.

Season or coils. The difference of the tile's two terms is $\hat{s}_{\text{coils}} - \hat{s}_{\text{season}} = +0.01060$, standard deviation 0.00732, so

$$\Pr(s_{\text{coils}} < s_{\text{season}} \mid \mathbf{y}) = \Phi(-1.4491) = 0.074,$$

and the full posterior gives 0.0724 (standard error 0.0007). The season is the larger part of the tile's drop with probability 0.93. The coils are 0.41 of the model's drop, coils and season together, with standard deviation 0.07.

The same season every summer. The season's term on the return air is +2.8735 K in each of the three spans from July to December, because it is the known inputs at the commissioned coils, and the inputs repeat. On the tile it is −0.0391, −0.0384 and −0.0360 in years 1, 2 and 3, moving only with the loads, and the coils' term at the MAP state is −0.0243, −0.0186 and −0.0254. The readings' changes were −6.04, −12.16 and −12.44 %. From year 1 to year 2 the season's and the coils' terms moved little while the reading's change doubled: the difference is the errors of the two windows' readings.

The two splits are drawn as waterfalls in Figure 13.3, the return air on the left and the tile on the right. Each bar starts where the previous one ended, from zero, the July value: the season (gold), the IT load (gray), the coils (blue, with a bar of one standard deviation at its end) and the reading errors (light gray); the last bar is the whole change to December. The dotted marker on the coils' bar is the stated truth's coils term, a check marker (Section 13.4.5). On the return air the gold bar is most of the change; on the tile the reading errors' bar is the longest, and the coils' bar is shorter than the season's.

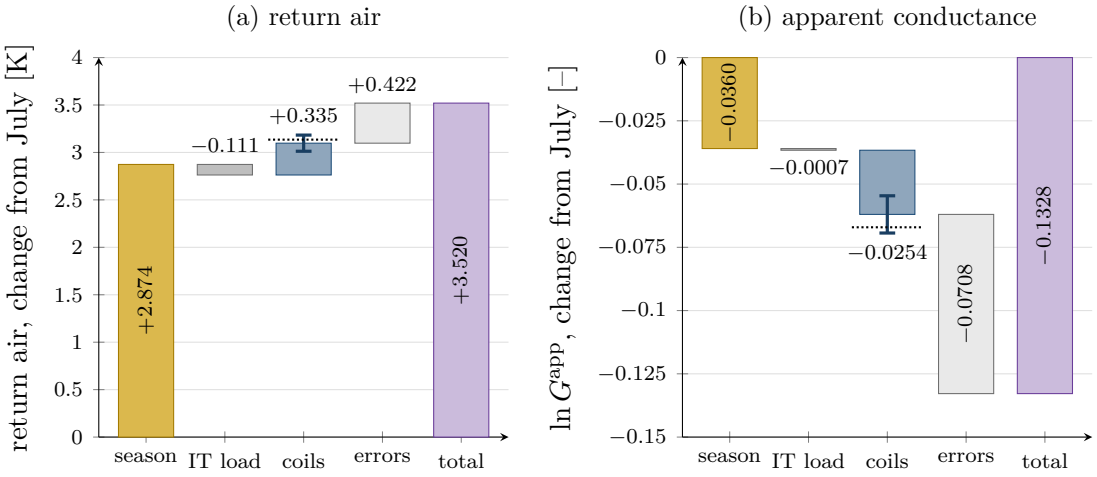


Figure 13.3: The summer drop from window 31 to window 36 as waterfalls, equation (13.3) at the MAP state (Table 13.3): (a) the return air, (b) the logarithm of the apparent conductance. Each term's bar starts where the previous one ended; the last bar is the reading's change. Gold: the season; gray: the IT load; blue: the coils, with ±1 posterior standard deviation; light gray: the reading errors; purple: the total. Dotted: the stated truth's coils term, a check marker (Section 13.4.5).

The return air's 3.52 K since July is 82 % season, the same 2.87 K every summer; the coils are 0.34 K of it. The tile's 12.4 % is half the errors of two windows' readings and a quarter season, and on the base drawing the coils lost 2.50 % of their conductance over the five months.

13.3.2 The line through the raw performance

Asset manager's question. What does the trend line through the raw performance read, and how much of it is the season? The dashboard draws two lines, through the last six windows and through the whole record, and the question asks what each slope is made of.

General form. An estimation question about a derived quantity is summarized by its moments, equation (2.10),

$$\mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}] = \int f(\mathbf{z})\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}, \qquad \mathrm{Var}(f(\mathbf{z}) \mid \mathbf{y}) = \int \big(f(\mathbf{z}) - \mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}]\big)^2\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}.$$

Specialized form. Question 3 is equation (2.10) with $f(\mathbf{z})$ each term of the slope's split (13.7), with $\mathbf{w} = \mathbf{w}^{\mathrm{line}}$ of equation (13.6) over the span, and each term's series the split (13.3) of every window against window 0, by equations (13.4) and (13.5) with $t_1 = 0$ and $t_2 = t$. Each term is linear in its series, so its gradient is $\mathbf{w}^{\mathsf{T}}$ applied to the series' gradient, in every window's columns, the last line of the pattern (13.8), and its variance is $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$. The readings' slope itself is a number of the readings; the split says which parts of the state make it. The spans are the dashboard's two, windows 0 to 36 and 31 to 36, and every span from window 0 to t_{as}, for $t_{\mathrm{as}} = 6$ to 36: the line as the record grew.

Building blocks. The pair $(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$; the slope row $\mathbf{w}^{\text{line}}$ of equation (13.6); the series of the split (13.3) at every window; the split (13.7).

The two lines taken apart. Table 13.4 splits the dashboard's two lines, on each raw figure. Each row is a line; the columns give the readings' slope, with the practitioner's standard error, and its four parts by equation (13.7), with the posterior standard deviations of the load's and the coils' parts.

Table 13.4: The least-squares lines through the raw figures, split by equation (13.7) at the MAP state. The readings' slope carries the practitioner's standard error, the load and coils terms their posterior standard deviations, in parentheses; the season's term is its leak into the slope. Slopes of $\ln G^{\text{app}}$ in yr^{-1}, of the return air in K/yr.

Line	Readings' slope [yr^{-1} or K/yr]	Season (leak) [same]	IT load [same]	Coils [same]	Reading errors [same]
$\ln G^{\text{app}}$, windows 0–36	−0.04934 (0.00757)	+0.00215	+0.00184 (0.00025)	−0.05255 (0.00398)	−0.00078
$\ln G^{\text{app}}$, windows 31–36	−0.27114 (0.08437)	−0.09302	+0.00232 (0.00271)	−0.06075 (0.01807)	−0.11968
T_r, windows 0–36	+0.6929 (0.2452)	−0.1699	+0.2799 (0.0388)	+0.5747 (0.0423)	+0.0083
T_r, windows 31–36	+9.456 (1.242)	+7.354	+0.379 (0.451)	+0.910 (0.206)	+0.814

The six-window line through the tile, $-0.27114\,\text{yr}^{-1}$ or $-23.75\,\%$/yr, is 34.3 % season, 44.1 % the errors of six windows' readings and 22.4 % coils: read as aging, it is a summer's worth of ambient and noise. On the return air's six-window line the season is 77.8 %. Over the whole record the season leaks little into the tile's line, $+0.00215\,\text{yr}^{-1}$, and the line is 106.5 % coils, the load pulling it the other way. On the return air's whole-record line the season's leak is −24.5 %, the IT load 40.4 % and the coils 82.9 %: 0.2799 K/yr of the line is the load's growth. On the return air, where the season is large, it widens the line even where it leaks little: the line has a residual standard deviation of 1.33 K and a standard error of 0.245 K/yr, against 0.49 K and 0.094 K/yr when the regression carries a seasonal term. On the tile, where the season is small beside the noise, a seasonal term barely changes the residual standard deviation, 0.0409 against 0.0408, and its standard error rises, from 0.00757 to $0.00780\,\text{yr}^{-1}$, for the two coefficients it spends.

The leak as the record grows. The line as the record grew is shown in Figure 13.4, whose horizontal axis is the last window t_{as} of each fit. The upper panel holds the slopes of the tile through windows 0 to t_{as}, the line's and the seasonal regression's, and the season's leak into the line's slope by equation (13.7); the lower panel, read in Section 13.3.4, sets the hall's rate on the parameter against both. In the upper panel the leak rises to $+0.0654\,\text{yr}^{-1}$ at $t_{\text{as}} = 8$, August of year 1: a line from commissioning into the first winter reads the season as an improvement of the coils. From $t_{\text{as}} = 12$ on it stays between −0.0032 and $+0.0179\,\text{yr}^{-1}$. The return air's line runs from −3.03 K/yr at $t_{\text{as}} = 8$ to +1.60 K/yr at $t_{\text{as}} = 16$.

The leak's fall with the span. A season that is an exact harmonic, $k_c \cos\phi_t + k_s \sin\phi_t$ in every window, leaks into the line's slope as k_c times the line's slope of a unit cosine over the span, plus k_s times that of a unit sine (Exercise 13.2). Over 12, 24 and 36 windows the unit cosine's slope is −0.764, −0.214 and $-0.099\,\text{yr}^{-1}$, and the sine's −1.324, −0.371 and $-0.171\,\text{yr}^{-1}$: over whole years they fall about as the inverse square of the span. The tile's season is close to such a harmonic, of amplitude 0.0229 at the MAP state, and its least-squares cosine coefficient is −0.02245, so at $t_{\text{as}} = 36$ the cosine's part of the leak is $(-0.02245)(-0.09893) = +0.00222\,\text{yr}^{-1}$, beside the exact +0.00215.

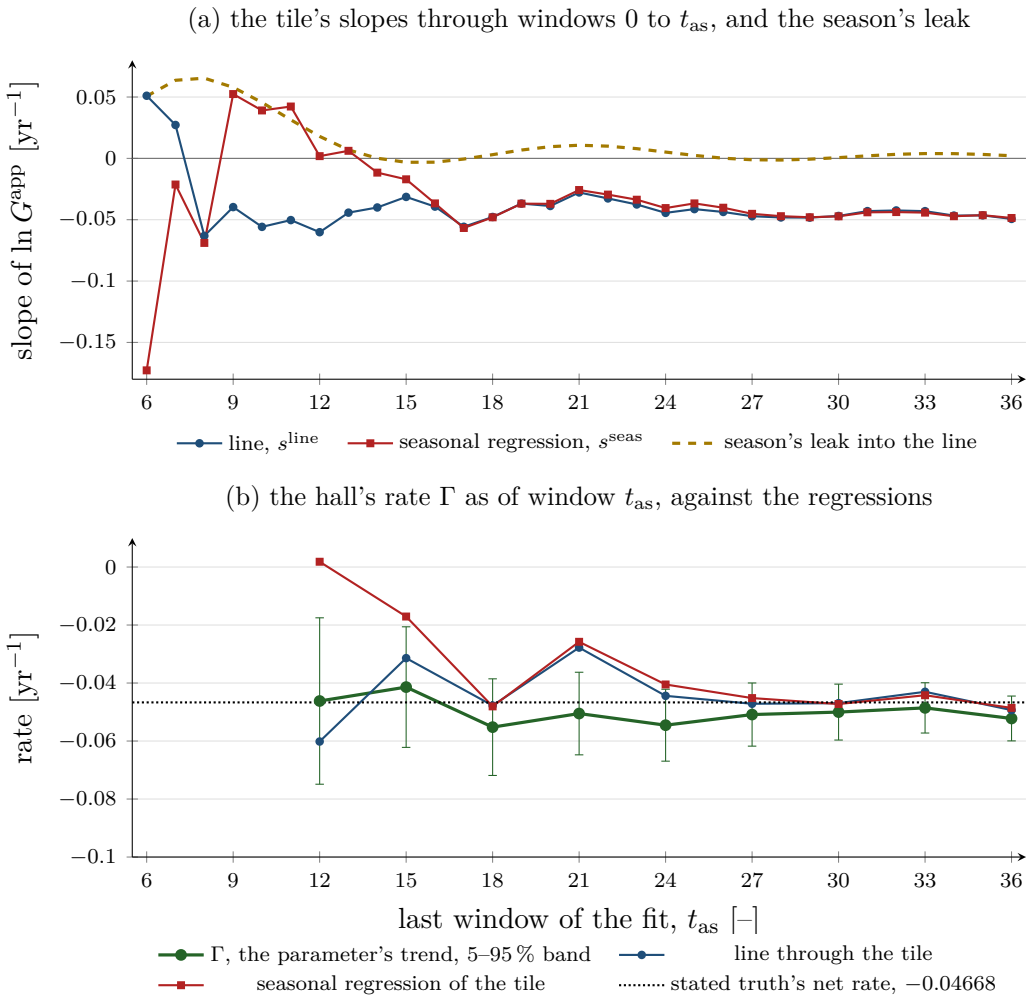


Figure 13.4: As of each window t_{as}, from the record cut at t_{as}. (a) The slope of $\ln G^{app}$ through windows 0 to t_{as} by the line and by the seasonal regression, and the season's exact leak into the line's slope, equation (13.7) at the MAP state, for $t_{as} = 6$ to 36. (b) The hall's rate Γ, the change of the logarithm of the hall's total conductance from window 0 to window t_{as} per year (Section 13.3.4), from the base fit cut at t_{as}, with its 5–95 % band, against the two regressions' slopes, for $t_{as} = 12$ to 36 in steps of three. Dotted: the stated truth's net rate, a check marker (Section 13.4.5).

Principle 13.2. *A line through raw performance reads the season as aging in proportion to the season's amplitude, and over whole years in about inverse proportion to the square of the span; on a span shorter than a year it reads it as aging entirely.*

13.3.3 The regression with a seasonal term

Asset manager's question. Does a seasonal term in the regression take the season out? The analyst's second chart fits one annual harmonic beside the line, and the question asks what that slope and that harmonic are made of.

General form. An estimation question about a derived quantity is summarized by its moments, equation (2.10),

$$\mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}] = \int f(\mathbf{z})\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}, \qquad \mathrm{Var}(f(\mathbf{z}) \mid \mathbf{y}) = \int \big(f(\mathbf{z}) - \mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}]\big)^2\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z}.$$

Specialized form. Question 4 is equation (2.10) with $f(\mathbf{z})$ each term of the split (13.7) with $\mathbf{w} = \mathbf{w}^{\mathrm{seas}}$, the slope row of the harmonic design of equation (13.6) over windows 0 to 36, its gradient formed as in Section 13.3.2. The fitted harmonic $(\hat{k}_c, \hat{k}_s)$ is the third and fourth rows of $\mathbf{X}^+$ applied to $\mathbf{y}^{\mathrm{raw}}$, split the same way, and the same rows applied to $\mathbf{s}_{\mathrm{season}}$ alone give the harmonic that the known inputs make. Because $\mathbf{X}^+\mathbf{X}$ is the identity, an exact harmonic season is fitted exactly and leaks nothing into the slope; what leaks is the season's departure from a harmonic. On the return air the season (13.4) is an exact harmonic, the supply's and the other heat's cosines at the commissioned conductance; on the tile it moves with the load as well. The regression's reading of the summer drop is its fitted trend over the five months plus the part of the drop its fit leaves in the residuals,

$$s^{\mathrm{seas}}_{\mathrm{drop}} = (36 - 31)\,\Delta t\, s^{\mathrm{seas}} + (\hat{\nu}_{36} - \hat{\nu}_{31}), \qquad \mathrm{sd}(s^{\mathrm{seas}}_{\mathrm{drop}}) = \sqrt{\big(5\,\Delta t\,\mathrm{se}(s^{\mathrm{seas}})\big)^2 + 2\,\hat{\sigma}^2}, \qquad (13.12)$$

with $\hat{\nu}_t = y^{\mathrm{raw}}_t - [\mathbf{X}\hat{\mathbf{k}}]_t$ the regression's residual at window t and $\hat{\sigma}$ the residuals' standard deviation: the practitioner's "aging beyond the fitted season", with the trend's standard error and two windows' residual scatter as its spread.

Building blocks. The pair $(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$; the rows of $\mathbf{X}^+$ for the harmonic design; the series of the split (13.3); the split (13.7); the regression's reading (13.12) of the summer drop.

The seasonal regression taken apart. Table 13.5 splits the seasonal regression over the whole record on each raw figure. Each row is a quantity: the readings' slope with its standard error, its four parts by equation (13.7), the fitted cosine and sine coefficients with their standard errors, and the cosine and sine coefficients of the season the known inputs make, the same rows of $\mathbf{X}^+$ applied to $\mathbf{s}_{\mathrm{season}}$; the columns are the two raw figures.

The seasonal term's gains. On the return air, where the supply makes the season large, the regression finds it: a fitted cosine of 1.741 K against the known 1.718 K, a fitted sine of −0.014 (standard error 0.119) against zero, and a leak of zero. Over three whole years on the tile it leaves the slope close to the line's, −0.04858 against $-0.04934\,\mathrm{yr}^{-1}$, −4.74 against −4.81 %/yr, with a leak of −0.00008. The practitioner's other seasonal term, twelve month-of-year levels and a trend over windows 1 to 36, reads −4.60 %/yr on the tile (standard error 0.0079) and +0.813 K/yr on the return air (standard error 0.093).

Table 13.5: The seasonal regression over windows 0 to 36, split by equation (13.7) at the MAP state, and its fitted harmonic against the season's. Standard errors and posterior standard deviations in parentheses.

Quantity	$\ln G^{\text{app}}$ [yr^{-1}; coefficients –]	Return air T_r [K/yr; coefficients K]
Readings' slope, s^{seas}	−0.04858 (0.00780)	+0.8628 (0.0937)
season (the leak)	−0.00008	+0.0000
IT load	+0.00196 (0.00025)	+0.2970 (0.0396)
coils	−0.05242 (0.00400)	+0.5751 (0.0425)
reading errors	+0.00196	−0.0094
Fitted cosine $\hat{k}_c$; the season's	−0.0101 (0.0097); −0.0225	+1.7406 (0.1166); +1.7176
Fitted sine $\hat{k}_s$; the season's	+0.0102 (0.0099); −0.0001	−0.0138 (0.1185); 0.0000

The seasonal term's remaining errors. Four things the regression cannot do without the physics, each with its number.

1. *The IT load read as aging.* On the return air, 0.297 of the slope's 0.863 K/yr, 34.4 %, is the load's growth. Converted the practitioner's way, the slope over the mean return approach of windows 1 to 36, 10.598 K, the seasonal slope is −7.82 %/yr of conductance; with the load's part removed by the physics it is −5.20 %/yr (Exercise 13.5).
2. *A small season barely seen.* On the tile the season is the unmetered other heat alone, a cosine coefficient of −0.0225, and the regression fits −0.0101 with a standard error of 0.0097: of its fitted cosine, −0.02245 is the season and +0.01168 the reading errors. The known inputs give the season exactly; the regression barely resolves it.
3. *The summer charged with the readings' errors.* By equation (13.12), the regression's fitted change over the five months, −0.04695, is −0.02671 of harmonic (standard error 0.01859) and −0.02024 of trend (standard error 0.00325), and the residual change is −0.08587; its "aging beyond the fitted season" is $s^{\text{seas}}_{\text{drop}} = -0.10611$, standard deviation 0.05783, a loss of 10.07 % with a band from 18.23 to 1.09 %, against the parameter's 2.50 %, band 1.32 to 3.68 %, of Section 13.3.1. The residuals of two windows are mostly their readings' errors, and the regression has no other place to put them.
4. *No coil.* A hall figure reads the total, and the dashboard's per-coil tile mixes the coils with each other and with unit 1's sensor (Section 10.3.3).

A seasonal term takes the season out of a slope over whole years, but not out of a summer, and not out of a figure in which the season is small beside the noise.

13.3.4 The hall's rate on the parameter

Asset manager's question. How fast is the hall's conductance falling, with the season taken out by the physics? The regressions read the hall's figures, so the rate set beside them is the hall's.

General form. The sign of a trend is the first line of the trend form, equation (2.21),

$$\Pr(\gamma > 0 \mid \mathbf{y}) = \int \mathbf{1}[\gamma > 0]\, p(\mathbf{z} \mid \mathbf{y})\, \mathrm{d}\mathbf{z},$$

and its size an estimate with its credible interval, equation (2.11), here at the levels 0.05 and 0.95.

Specialized form. The question is the first line of equation (2.21) and equation (2.11) with the trend the hall's conductance rate,

$$\Gamma = \frac{\ln G_{\mathrm{tot},36} - \ln G_{\mathrm{tot},0}}{36\,\Delta t}, \qquad \mathbf{g}_\Gamma = \frac{1}{36\,\Delta t}\Big(-\frac{G_{1,0}}{G_{\mathrm{tot},0}},\ -\frac{G_{2,0}}{G_{\mathrm{tot},0}},\ \frac{G_{1,36}}{G_{\mathrm{tot},36}},\ \frac{G_{2,36}}{G_{\mathrm{tot},36}}\Big)^{\mathsf{T}},$$
$$\Pr(\Gamma < 0 \mid \mathbf{y}) = \Phi\left(\frac{-\hat{\Gamma}}{\mathrm{sd}(\Gamma \mid \mathbf{y})}\right), \qquad \mathrm{sd}(\Gamma \mid \mathbf{y}) = \sqrt{\mathbf{g}_\Gamma^{\mathsf{T}} \boldsymbol{\Sigma}\, \mathbf{g}_\Gamma}, \tag{13.13}$$

in yr^{-1}, the gradient on $(g_{1,0}, g_{2,0}, g_{1,36}, g_{2,36})$ and zero elsewhere, the fourth line of the pattern (13.8), and $p(\mathbf{z} \mid \mathbf{y})$ the Laplace normal. With γ_u free the trend rows make each coil's fitted change its rate times the span (Chapter 10), so to first order Γ is the conductance-weighted mean of the two rates. One choice enters here: the hall's Γ is the like-with-like reference for regressions on hall figures, which see the total and never a coil; each coil's rate stays Chapter 10's answer.

Building blocks. The pair $(\hat{\mathbf{z}}, \boldsymbol{\Sigma})$; the conductances at the MAP state, the entries of $\mathbf{g}_\Gamma$; the rate (13.13); Φ; the regressions' slopes of Sections 13.3.2 and 13.3.3.

The hall's rate. At the MAP state, by equation (13.13),

$$\mathbf{g}_\Gamma = (-0.16821,\ -0.16512,\ +0.17324,\ +0.16009)^{\mathsf{T}}\,\mathrm{yr}^{-1},$$
$$\hat{\Gamma} = -0.05222\,\mathrm{yr}^{-1}, \qquad \mathrm{sd}(\Gamma \mid \mathbf{y}) = 0.00470\,\mathrm{yr}^{-1},$$

a rate of $-5.09\,\%$/yr with 5–95 % band -5.82 to $-4.35\,\%$/yr, and $\Pr(\Gamma < 0 \mid \mathbf{y}) = 1$ to six decimals. The full posterior gives a median of -5.08 and a band of -5.81 to $-4.34\,\%$/yr. As the coils' rates weighted by coil 1's mean share of the hall's conductance over the record, 0.5143, the first-order form gives $0.5143\,(-0.04240) + 0.4857\,(-0.06253) = -0.05218\,\mathrm{yr}^{-1}$. On the straight line, the limit $\sigma_q \to 0$, the rate is $-5.075\,\%$/yr with standard deviation 0.00394. Each coil's own rate is Chapter 10's, -4.15 and $-6.06\,\%$/yr (Section 10.3.1).

The routes compared. Table 13.6 sets the hall's rate beside the regressions, like with like: each row is a route, a slope of the logarithm of a conductance per year, with its standard deviation or standard error, its rate and its 5–95 % band. The six-window line is a summer, not a rate. The line through the whole record and the seasonal regression read -4.81 and $-4.74\,\%$/yr on this record, beside the parameter's -5.09, with wider bands: the parameter's standard deviation is 0.603 of the seasonal regression's standard error. The parameter reads all five readings of every window through the physics; the regression reads two, through their ratio. Whether the narrower band also reads the truth better is the question of fresh records, Section 13.4.4.

This summer against the record's pace. The coils' five months of Section 13.3.1, -0.02536, are $-0.06087\,\mathrm{yr}^{-1}$ as a rate, $-5.91\,\%$/yr, against the record's Γ. Their difference from five months of Γ, $s_{\mathrm{coils}} - 5\,\Delta t\,\Gamma = -0.00360$ with standard deviation 0.00653, is 0.55 standard deviations, and the probability that the summer ran faster than the record's rate is 0.709 (Exercise 13.1): on the base drawing's reading the summer shows no departure from the record's pace. That reading is the drawing's, not a test of the pace. Its straight line with wander pulls a five-month change toward the record's rate, and on the truth's own readings it gives a pace of -0.00188, while the true summer ran faster than the true rate by -0.01100, -1.68 posterior standard deviations (Section 13.4.5).

Table 13.6: The hall's conductance rate by route, windows 0 to 36: the slope of $\ln G^{\text{app}}$ for the regressions, of $\ln G_{\text{tot}}$ for the parameter. Bands are 5–95 %.

Route	Slope (sd) [yr^{-1}]	Rate [%/yr]	Band [%/yr]
Line through the tile, last six windows	−0.27114 (0.08437)	−23.75	−33.63 to −12.40
Line through the tile, whole record	−0.04934 (0.00757)	−4.81	−5.99 to −3.62
Seasonal regression of the tile, whole record	−0.04858 (0.00780)	−4.74	−5.96 to −3.51
The parameter's trend Γ, base fit (reported)	−0.05222 (0.00470)	−5.09	−5.82 to −4.35
The parameter's trend, straight line	−0.05208 (0.00394)	−5.08	−5.69 to −4.46

The winter to come. At July of year 4, window 43, with both windows taken at the forecast load of 137.66 kW, the season's term returns +0.03455 to the tile's logarithm, while the coils, at the rate $\hat{\Gamma}$, lose another 0.03046 over the seven months. On the record's posterior, with the load held at that forecast, the tile's model value at window 43 is 11.03 kW/K, band 10.74 to 11.32, against today's model value of 10.94 kW/K and July of year 3's 11.64: the asset manager should expect the tile next winter near today's level, not near this July's.

The rate as of each window. Table 13.7 gives the hall's rate from the base fit cut at window t_{as}, using nothing later, beside the two regressions' slopes over the same span; each row is an as-of window. The lower panel of Figure 13.4 plots the same from $t_{\text{as}} = 12$, with Γ's band as error bars and the stated truth's net rate dotted. From $t_{\text{as}} = 12$ on, every as-of Γ lies between −0.0552 and −0.0414 yr^{-1}, while the two regressions' slopes over the same spans run from −0.0602 to +0.0018 yr^{-1}, and the seasonal regression reads a gain of +0.0018 yr^{-1} (standard error 0.0537) at $t_{\text{as}} = 12$, where Γ reads −0.0462 (standard deviation 0.0174). As of June of year 1, $t_{\text{as}} = 6$, the seasonal regression reads −0.1728 (standard error 0.2116) and Γ +0.0467 (standard deviation 0.0436). On a short record the harmonic and the trend share the span; the known inputs share nothing with the trend. Coil 2's steep months, windows 13 to 16, move Γ from −0.0414 at $t_{\text{as}} = 15$ to −0.0552 at $t_{\text{as}} = 18$, reported here as the record shows them; coil 2's path is Chapter 12's question.

Table 13.7: The hall's rate Γ from the base fit cut at window t_{as}, against the slopes of the line and of the seasonal regression through the tile over windows 0 to t_{as}.

As of window t_{as}	Γ [yr^{-1}]	sd [yr^{-1}]	Line [yr^{-1}]	se [yr^{-1}]	Seasonal [yr^{-1}]	se [yr^{-1}]
6 (Jun Y1)	+0.04670	0.04364	+0.05104	0.11042	−0.17277	0.21159
9 (Sep Y1)	−0.02482	0.02622	−0.03971	0.06763	+0.05242	0.11451
12 (Dec Y1)	−0.04620	0.01744	−0.06017	0.03988	+0.00182	0.05369
15 (Mar Y2)	−0.04140	0.01264	−0.03141	0.02732	−0.01707	0.02653
18 (Jun Y2)	−0.05521	0.01013	−0.04777	0.02175	−0.04805	0.02227
21 (Sep Y2)	−0.05052	0.00866	−0.02778	0.01809	−0.02579	0.02012
24 (Dec Y2)	−0.05454	0.00754	−0.04442	0.01490	−0.04050	0.01641
27 (Mar Y3)	−0.05087	0.00662	−0.04715	0.01215	−0.04519	0.01278
30 (Jun Y3)	−0.05002	0.00586	−0.04694	0.00992	−0.04724	0.01018
33 (Sep Y3)	−0.04857	0.00527	−0.04302	0.00848	−0.04421	0.00880
36 (Dec Y3)	−0.05222	0.00470	−0.04934	0.00757	−0.04858	0.00780

The season in the capacity scenario. The capacity requirement $G^{*} = 10.761$ kW/K is set at the scenario's summer supply of 15.5°C, so the season returns in the forecast of the margin

through the requirement, not through the trend; that forecast is Chapter 14's.

13.3.5 A season of the coils themselves

Asset manager's question. Does the coils' conductance itself fall in summer? A wet coil, with condensate on its fins, could pass less heat in the humid months, and the known inputs would not carry that.

General form. A hypothesis comparison is the ratio of the evidences of two complete drawings fitted to the same evidence, equation (2.14),

$$B_{AB} = \frac{p(\mathbf{y} \mid M_A)}{p(\mathbf{y} \mid M_B)}.$$

Specialized form. The question is equation (2.14) with $M_A = M_{\text{sea}}$ and $M_B = M_{\text{base}}$ of Section 13.2.2, each evidence the Laplace evidence (7.7) of its own fit. The two drawings are nested: the seasonal drawing with its four amplitudes pinned at zero is the base drawing. By equation (7.8) with four entries freed, each of prior spread σ_{sea},

$$\ln B(M_{\text{sea}} : M_{\text{base}}) = \left[\mathcal{L}_{\text{base}}(\hat{\mathbf{z}}_{\text{base}}) - \mathcal{L}_{\text{sea}}(\hat{\mathbf{z}}_{\text{sea}})\right] - \frac{1}{2} \ln \frac{\sigma_{\text{sea}}^{8} \det \mathbf{\Lambda}_{\text{sea}}}{\det \mathbf{\Lambda}_{\text{base}}}, \tag{13.14}$$

the gain in fit less the Occam factor. Beside it stands the Savage–Dickey ratio (7.21) of the four amplitudes together,

$$B(M_{\text{base}} : M_{\text{sea}}) = \frac{p(\mathbf{g}^{\text{sea}} = \mathbf{0} \mid \mathbf{y}, M_{\text{sea}})}{p(\mathbf{g}^{\text{sea}} = \mathbf{0} \mid M_{\text{sea}})}, \qquad \mathbf{g}^{\text{sea}} = (g_1^c,\ g_1^s,\ g_2^c,\ g_2^s)^{\mathsf{T}}, \tag{13.15}$$

the Laplace normal's density of the amplitudes at zero over their prior density there, $(2\pi\sigma_{\text{sea}}^2)^{-2}$. Each amplitude is an entry of M_{sea}'s state, read off with its standard deviation; a summer derate is a negative cosine amplitude, because $\cos\phi_t = 1$ at the summer peak.

Building blocks. Each drawing's objective at its MAP state and the logarithm of the determinant of its information (Section 13.2.2); the prior spread σ_{sea} of equation (13.9); the amplitudes' block of M_{sea}'s covariance.

The evidence. By equation (13.14),

$$\ln B(M_{\text{sea}} : M_{\text{base}}) = 2.1726 - 7.1435 = -4.9710, \qquad B = 0.0069,$$

the gain in fit, half of $154.9755 - 150.6304$, less an Occam factor of 7.1435 for four freed entries. The Savage–Dickey ratio (13.15) gives $\ln p(\mathbf{0} \mid \mathbf{y}) - \ln p(\mathbf{0}) = 13.2890 - 8.3072 = +4.9818$ in favor of the base drawing, that is -4.9818; importance sampling gives -4.9698. The record favors the drawing without a seasonal conductance: the known inputs carry the season.

The amplitudes. Under the seasonal drawing $g_1^c = +0.0010$ (standard deviation 0.0086, band -0.0132 to $+0.0152$), $g_1^s = +0.0175$ (0.0087, band $+0.0032$ to $+0.0318$), $g_2^c = +0.0037$ (0.0091, band -0.0112 to $+0.0186$) and $g_2^s = -0.0078$ (0.0092, band -0.0228 to $+0.0073$). The summer components g_u^c sit at zero: a summer derate larger than 1.3 % for coil 1 or 1.1 % for coil 2 lies outside the 5–95 % band. The total seasonal amplitude of each coil, $\sqrt{(g_u^c)^2 + (g_u^s)^2}$, has a 95 % quantile of 0.0333 for coil 1 and 0.0264 for coil 2 in the full posterior. Coil 1's sine term, a swing between autumn and spring whose band excludes zero, is reported as the record shows it; Section 13.4.5 reads it against the truth. Under the seasonal drawing the rates barely move, $\gamma_1 = -0.03885$ and $\gamma_2 = -0.06318\,\text{yr}^{-1}$, against the base fit's -0.04240 and -0.06253.

The prior on the amplitude. The size of the factor depends on the amplitudes' prior spread, and Table 13.8 repeats the comparison over five spreads. Each row is a spread; the columns give $\ln B$, coil 1's sine amplitude with its standard deviation, and the gain in fit. The factor's sign does not change: $\ln B$ runs from -0.10 at a spread of 0.005 to -7.64 at 0.10, because a wider prior pays a larger Occam factor for a gain in fit that stops growing. The record rules out a seasonal conductance of several percent; it cannot rule out one under one percent, which the narrowest prior describes and which the readings could not see.

Table 13.8: The evidence for a seasonal conductance over the prior spread σ_{sea} of its amplitudes, by equation (13.14).

Prior spread σ_{sea} [–]	$\ln B(M_{\text{sea}} : M_{\text{base}})$ [–]	g_1^s (sd) [–]	Gain in fit [–]
0.005	-0.1007	$+0.00424$ (0.00426)	0.5394
0.010	-0.5754	$+0.00968$ (0.00644)	1.2181
0.020	-1.9398	$+0.01476$ (0.00798)	1.8406
0.050 (the chapter's)	-4.9710	$+0.01751$ (0.00871)	2.1726
0.100	-7.6365	$+0.01800$ (0.00883)	2.2316

Principle 13.3. *The season is in the inputs, not in the coils; the evidence against a seasonal conductance is the check that the known inputs carry all of it.*

13.4 The checks (step 5)

Each answer of Section 13.3 rests on approximations, on inputs taken as known and on one record's draws, and this hall lets each be tested against the calculation it stands for.

13.4.1 The full posterior against the Laplace answers

The reweighted draws of Section 13.2.2 have been set beside every answer of step 4. Their Monte Carlo standard errors, by batch means over 20 batches, are at most 0.00004 on the medians and band ends of the five months and of Γ, 0.00067 on the probability of question 2, and 0.00028 on the seasonal drawing's log evidence by importance sampling. The largest gap between the full posterior and the Laplace answer is 0.00013 in the logarithm across the medians and band ends: larger than its standard error, so resolved, and small beside the bands of Section 13.3. The probability of question 2 is 0.07239 against 0.07366, a gap of 0.00127, 1.9 standard errors, at the edge of what the draws resolve; the log Bayes factor is -4.9698 against -4.9710, a gap of 0.0011. Every answer is reported to a precision that the Laplace normal meets.

13.4.2 The limit of exact rows and the curvature term

The fit runs on the reduced state, which equation (4.48) makes the limit $\sigma_p \to 0$ of the full state. The check fits the full state of the record, 300 entries, the 115 of the stacked state and each window's $(T_r, T_{d,1}, T_{d,2}, T_c, T_1)$, under the 185 physical rows (5.2) at a common tolerance beside the readings and the priors, and lets the tolerance shrink. At $\sigma_p = 0.1$, 0.01 and 0.001 the coils' five months read -0.02530 (standard deviation 0.00738), -0.02536 (0.00736) and -0.02536 (0.00736), and Γ reads -0.05211 (0.00476), -0.05222 (0.00470) and $-0.05222\,\text{yr}^{-1}$ (0.00470), in 7, 5 and 13 iterations, against the reduced state's -0.02536 (0.00736) and -0.05222 (0.00470). At 0.01 and below the full state reproduces the reduced state to the fifth decimal.

The covariance of Section 13.2.1 is the Gauss–Newton form, without the curvature term $\mathbf{C} = \mathbf{H} - \mathbf{J}^{\mathsf{T}}\mathbf{J}$ of equation (4.44). The full Hessian at $\hat{\mathbf{z}}$, by differences, gives the Newton

covariance, in which the five months' standard deviation is 0.00735 and Γ's 0.00469, ratios of 0.9988 and 0.9975 to Gauss–Newton's. The largest entry of $\mathbf{C}$ is 8.81×10^{-4} of the largest entry of $\mathbf{J}^{\mathsf{T}}\mathbf{J}$. The Gauss–Newton covariance is the one reported.

13.4.3 The known inputs and the priors

Table 13.9 repeats the two reported answers under other choices. Each row is a variant: the other heat made an entry of each window with a prior row of 1.0 or 3.0 kW about its schedule, or with its mean and amplitude freed; the supply made an entry of each window with a prior row of 0.2 K about its instrument; two other wanders and the straight line; the base drawing without its drift entry; and question 6's seasonal drawing. The columns give the coils' five months and Γ, each with its standard deviation, the rate, and the change of Γ's standard deviation from the base fit's.

Table 13.9: The coils' five months, $\ln(G_{\text{tot},36}/G_{\text{tot},31})$, and the hall's rate Γ under other choices of the inputs and the priors.

Variant	Five months [–]	sd [–]	Γ [yr^{-1}]	sd [yr^{-1}]	Γ [%/yr]	sd of Γ change [%]
The base fit (reported)	−0.02536	0.00736	−0.05222	0.00470	−5.088	0
Other heat known to 1.0 kW per window	−0.02518	0.00739	−0.05215	0.00477	−5.081	+1.5
Other heat known to 3.0 kW per window	−0.02413	0.00760	−0.05168	0.00524	−5.037	+11.5
Other heat's season estimated	−0.02579	0.00739	−0.05194	0.00486	−5.061	+3.5
Supply known to 0.2 K per window	−0.02396	0.00784	−0.05263	0.00596	−5.127	+26.8
$\sigma_q = 0.0025$	−0.02273	0.00425	−0.05217	0.00419	−5.083	−10.9
$\sigma_q = 0.01$	−0.03053	0.01184	−0.05165	0.00566	−5.034	+20.5
Straight line, $\sigma_q \to 0$	−0.02160	0.00168	−0.05208	0.00394	−5.075	−16.3
No drift entry	−0.02991	0.00725	−0.06059	0.00420	−5.879	−10.7
Seasonal-conductance drawing	−0.02494	0.01065	−0.05143	0.00500	−5.013	+6.3

The other heat and the supply. With its season estimated from the record, the other heat reads a mean of 7.72 kW (standard deviation 1.45) and an amplitude of 1.99 kW (1.07), against the schedule's 8.0 and 3.0 kW, and Γ's standard deviation grows by 3.5 %. The reason is in the forward equations (5.3): $T_r - T_{d,u} = G_{u,t}(T_r - T_s)/C_u$, so the ratio $(T_r - T_{d,u})/(T_r - T_s) = G_{u,t}/C_u$ fixes each coil's conductance from three temperatures whatever the heat. The supply enters that ratio directly, and known only to 0.2 K per window it widens Γ's standard deviation by 26.8 %. The instrument that takes the season out is the supply's.

The wander in the five months' band. The five months' standard deviation is 0.00168 on the straight line, 0.00425 at $\sigma_q = 0.0025$, 0.00736 at the base fit's 0.005 and 0.01184 at 0.01, while Γ's stays between 0.00394 and 0.00566. Over five windows the trend rows allow each coil a wander of $\sqrt{5}\,\sigma_q = 0.01118$, and most of the summer's band is that allowance, which Chapter 10 chose and this chapter keeps. The rate over three years hardly depends on it.

The drift and the drop's start. Without the drift entry the hall's rate steepens to −5.88 %/yr: unit 1's sensor reaches the hall's total too, through coil 1's share. Measured from June of year 3, window 30, a drop of six months, the coils' term reads −0.03044 (standard deviation 0.00787) against a season of −0.03005 on the tile, a tie; measured from August, window 32, a drop of four months, it reads −0.02010 (0.00673) against −0.03689. The return air rose 2.56 K from window 30, of which 2.4290 K is season, and 3.74 K from window 32, of

which 2.8735 K. The verdict that the season is the larger part of the tile's drop holds from windows 31 and 32, and is a tie from window 30.

13.4.4 Fresh records

The answers are one record's. Two hundred fresh records at the stated truth, the same loads, path and drift with fresh sensor errors, are each read by the base fit, the three regressions and the seasonal drawing. Table 13.10 summarizes them: each row is a route, with the mean and standard deviation of its answer over the records and its root-mean-square error against the truth's value, the net rate $-0.04668\,\mathrm{yr}^{-1}$ of the true hall conductance over the record, or its true five months, -0.03045 (Section 13.4.5).

Table 13.10: The routes over 200 fresh records at the stated truth. Rates in yr^{-1}, the five months in [–]; the root-mean-square error is taken against the truth's net rate or its five months.

Route	Mean [yr^{-1} or –]	sd [same]	rms against the truth [same]
Γ, the base fit	−0.04577	0.00426	0.00435
Line through the tile, whole record	−0.04297	0.00679	0.00772
Seasonal regression of the tile	−0.04534	0.00698	0.00709
Line through the last six windows	−0.16862	0.10935	0.16360
Five months, the base fit	−0.02069	0.00385	0.01048
Five months, the seasonal regression's reading	−0.03841	0.05402	0.05447

Γ's 5–95 % band covers its pseudo-true value, Chapter 10's like-with-like reference, in 0.925 of the records and the truth's net rate in 0.945; the seasonal regression's band covers its own pseudo-true value in 0.900 and the truth's net rate in 0.890. Γ's posterior standard deviation, 0.00473 on average, is close to its scatter over the records, 0.00426. Over records the parameter reads the truth's rate better than the seasonal regression by a factor of 1.63 in root-mean-square error. On this record the order is the other way: its Γ, -0.05222, sits at the 0.080 quantile of the fresh records, while its seasonal-regression slope sits at the 0.300 quantile, and the regression lands 0.24 standard errors from the truth's net rate against Γ's 1.18 standard deviations (Section 13.4.5). The comparison of the two routes rests on the fresh records, not on this record's distance from the truth. The base fit's five months average -0.02069, near the drawing's pseudo-true -0.02096 and short of the truth's -0.03045, with a scatter of 0.00385; the regression's reading of the summer scatters with a standard deviation of 0.05402. The five months' 5–95 % band, of standard deviation 0.00738 on average, covers the drawing's pseudo-true value in every record but the truth's five months in only 0.745 of them, against the 0.90 a band of that level promises. The pace of Section 13.3.4, the five months less five months of Γ, averages -0.00163 over the records (standard deviation 0.00289), and in none of them does it reach the truth's -0.01100: the base drawing does not read a summer faster than its own line.

The seasonal drawing decides against a seasonal conductance, $\ln B < -2.944$, a posterior probability of 0.95 or more for the base drawing at even prior odds, in 0.895 of the records, and favors it, $\ln B > 0$, in 0.010. This record's -4.9710 sits at the 0.510 quantile.

13.4.5 The answers against the stated truth

The last check sets the answers against the stated truth of the record (Section 13.6.5). Table 13.11 collects the routes: each row is a route or a reference, and the columns give the coils' five

months with its band or standard deviation, the probability of question 2, Γ with its band or standard deviation, and the log Bayes factor of question 6.

At the true state the summer drop splits on the return air into a season of +2.8790 K, a load term of −0.1990 K, coils of +0.3727 K and reading errors of +0.4674 K, and on the tile into −0.03636 of season, −0.00120 of load, −0.03045 of coils and −0.06482 of errors, beside the posterior's season of +2.8735 K and −0.03596: the season's split of the drop, 2.9 of the return air's 3.5 K, is the known inputs' and holds against the truth. The coils' term does not hold as well. The coils' true five months are −0.03045, and this record's posterior, −0.02536, lies +0.69 standard deviations from them. The drawing's own pseudo-true value, the base fit on the truth's noise-free readings, is −0.02096, short of the truth by 0.00949, 1.29 posterior standard deviations: the truth's path is not a straight line, and over these five months the straight line with wander pulls the change toward the record's rate. The pace says the same. The truth's summer ran faster than its own rate, $-0.03045 - (-0.01945) = -0.01100$, −1.68 posterior standard deviations, while the drawing reads −0.00188 on the truth's readings and −0.00360 on this record. The finding of Section 13.3.4, no departure from the record's pace, is the base drawing's reading, and the truth contradicts it. Per coil the gap is larger: coil 1's true five months are −0.0521 against the posterior's −0.0263, +2.25 standard deviations, and coil 2's −0.0083 against −0.0243, −1.45 standard deviations, with pseudo-true values of −0.0208 and −0.0211. This is the drawing's limit, not the record's draws: on the truth's own readings the straight-line drawing gives the two coils nearly equal shares of the hall's five months, which the truth does not have (Section 13.6.5); whether coil 1's rate changes is Chapter 11's question. On this record the hall's five months fall within their band of the truth, the split between the coils does not, and over fresh records the band covers the truth's five months in 0.745 of them (Section 13.4.4).

The truth's net rate is $-0.04668\,\mathrm{yr}^{-1}$, −4.56 %/yr; the posterior's Γ lies −1.18 standard deviations from it and −1.37 from its pseudo-true −0.04579, which is this record's draw, the 0.080 quantile of Section 13.4.4. The seasonal regression's slope, −0.04858, lies −0.24 standard errors from the truth's net rate, beside its pseudo-true −0.04574. On the truth's noise-free readings the seasonal drawing gives $\ln B = -6.6164$ and amplitudes of at most 0.009: $g_1^c = +0.00015$, $g_1^s = +0.00356$, $g_2^c = +0.00221$ and $g_2^s = -0.00882$. Coil 1's sine amplitude of +0.0175 on this record is its readings' draw. The truth has no seasonal conductance.

Table 13.11: The routes to the summer drop and the hall's rate, the pseudo-true values and the stated truth. Bands are 5–95 %; P is the probability of question 2; $\ln B$ is that of the seasonal drawing against the base drawing.

Route	Five months [–]	band or sd [–]	P [–]	Γ [yr^{-1}]	band or sd [yr^{-1}]	$\ln B$ [–]
Laplace on the base fit (reported)	−0.02536	−0.03746 to −0.01326	0.07366	−0.05222	−0.05995 to −0.04449	−4.9710
Full posterior, reweighted draws	−0.02531	−0.03739 to −0.01313	0.07239	−0.05215	−0.05984 to −0.04437	−4.9698
Full state, $\sigma_p = 0.1$	−0.02530	sd 0.00738	–	−0.05211	sd 0.00476	–
Newton covariance	–	sd 0.00735	–	–	sd 0.00469	–
Savage–Dickey ratio	–	–	–	–	–	−4.9818
Drawing's pseudo-true value	−0.02096	–	–	−0.04579	–	−6.6164
Stated truth	−0.03045	–	–	−0.04668	–	no season

13.5 The report (step 6)

The return air rose 3.5 K since July, and 2.9 K of it is the season, the supply's 2.5 K and the other heat's 0.4 K, the same 2.9 K every summer; the coils are 0.3 K. The tile's 12 % is half the

errors of two windows' readings and a quarter season. On the base drawing the coils lost 2.5 % of their conductance over the five months (1.3 to 3.7 %), and the season is the larger part of the tile's drop with probability 0.93. The season's part is the known inputs' and holds; the coils' part is the drawing's reading, which pulls a summer toward the record's rate, and the check against the truth shows the true summer faster than the true rate and the coils' loss larger, −0.0305 in the logarithm against the posterior's −0.0254. The line through the last six months, −24 % a year, is a summer, not a rate. A seasonal term fixes the slope over three whole years, −4.7 against the parameter's −5.1 % a year, but not the summer, which it reads as a 10 % loss, and not a short record. The coils carry no season of their own ($B = 0.007$ against one).

13.5.1 The answers and their bands

Table 13.12 gives the answers in the order of Table 13.2. Each row is a question; the second column gives the reported value from the Laplace normal, the third the full posterior's value beside it where one exists, and the fourth the band: the 5–95 % band of a change or a rate, a regression's own band for its slope, and for the Bayes factor its range over the amplitudes' prior.

Table 13.12: The report for the two-unit hall at window 36 (Dec Y3) on the summer drop from window 31 (Jul Y3) and the whole record. Reported values from the Laplace normal of the base fit; bands are 5–95 %.

Asset manager's question	Reported value	Full posterior	Band
Is the summer drop in performance degradation or just the ambient?	return air +3.52 [K]: season +2.87, IT load −0.11, coils +0.34, reading errors +0.42 [K]; hall conductance −2.50 [%] over the five months on the base drawing	−2.50 [%]	coils +0.195 to +0.475 [K]; conductance −3.68 to −1.32 [%]
Is more of the drop the coils than the season?	no: P(coils larger) = 0.074 [–]; season the larger, 0.93 [–]	0.072 [–]	coils' share 0.41, sd 0.07 [–]
What does the trend line through the raw performance read, and how much of it is the season?	six windows −23.75 [%/yr], 34.3 [%] season and 44.1 [%] reading errors; whole record −4.81 [%/yr], leak +0.00215 [yr^{-1}]	–	regression's own: −33.63 to −12.40; −5.99 to −3.62 [%/yr]
Does a seasonal term in the regression take the season out?	over three years, −4.74 [%/yr]; the summer beyond the fitted season −10.07 [%]	–	regression's own: −5.96 to −3.51 [%/yr]; −18.23 to −1.09 [%]
How fast is the hall's conductance falling, with the season taken out by the physics?	Γ: −5.09 [%/yr]	−5.08 [%/yr]	−5.82 to −4.35 [%/yr]
Does the coils' conductance itself fall in summer?	no: $\ln B = -4.97$, $B = 0.0069$ [–]; summer derate above 1.3 [%] excluded	$\ln B = -4.97$ [–]	$\ln B$ −0.10 to −7.64 over prior spreads 0.005 to 0.1 [–]

The answers rest on the known inputs. The supply takes the season out of the parameter, and step 5 shows that the supply instrument, not the other heat's schedule, is the input the answers lean on; the probability adds the band, which separates the coils' 2.5 % from the season's part and from the errors of two windows.

13.5.2 The entry in the maintenance log

The entry is the two-unit hall's fourth at window 36, after Chapter 10's rate entry (Section 10.5.2) and the entries of Chapters 11 and 12. It uses nothing later than window 36, and in the values of Table 13.12 it reads:

Window 36 (Dec Y3), the summer drop from window 31 (Jul Y3), from the condition record (windows 0–36).
Return air +3.5 K: season 2.9 K (supply 2.5 K, other heat 0.4 K), IT load −0.1 K, coils 0.3 K (0.2 to 0.5), reading errors 0.4 K.
Hall conductance over the five months, on the straight-line drawing: −2.5 % (−3.7 to −1.3 %); the season the larger part of the tile's drop ($P = 0.93$). Hall rate −5.1 %/yr (−5.8 to −4.4). Whether the summer ran faster than the rate is not settled by this drawing (Chapter 11).
No seasonal change in the coils' own conductance (Bayes factor 0.007 against; a summer derate above 1.3 % excluded).
Dashboard hall tile and six-month trend line: season-sensitive; not a condition reading.
Action: none from the summer drop; the cleaning plan stays with the October shutdown of year 4 (window 46).
Next measurement: reference-grade readings (0.1 K) at Jul Y4 and Dec Y4; the supply instrument kept in calibration; the side-by-side check of unit 1's discharge sensor (Chapter 10's entry, still open).

The entry records what its numbers rest on: the supply and the other heat as known inputs of every window, and the base fit's straight line with its monthly wander, which sets most of the five months' band. When to clean is not settled by the summer's drop; it needs the crossing date of Chapter 14 and the decision level of Chapter 19.

Common mistake

Trending a raw-performance figure across seasons as condition. The return air and the apparent conductance move with the supply and the other heat; a line through them reads the season as aging, and a fitted harmonic estimates from noisy readings what the known inputs give exactly. The condition is the parameter, with the season removed by the physics.

13.6 The next measurement (step 7)

The report leaves the coils' five months known to a standard deviation of 0.00736 in the logarithm, and the asset manager will ask the same question of next summer. A candidate measurement is one more row or more of the residual vector, and what it removes from the band is known before its value is.

13.6.1 The preposterior of a candidate row

A candidate a is one or more rows with Jacobian lines $\mathbf{J}_a$ in the columns of the state, each the gradient of the quantity it reads over its standard deviation. Its preposterior covariance is equation (4.65),

$$\boldsymbol{\Sigma}_a = \left(\boldsymbol{\Sigma}^{-1} + \mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a\right)^{-1}, \qquad \mathrm{sd}_a = \sqrt{\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}_a\,\mathbf{g}}, \tag{13.16}$$

with $\mathbf{g}$ the gradient of the five months (13.11), and the width it takes off is $1 - \mathrm{sd}_a/\mathrm{sd}$. Two questions are asked. Today's five months use the base state as it stands. Next summer's, $\ln(G_{\mathrm{tot},48}/G_{\mathrm{tot},43})$ from July to December of year 4, use the state extended to window 48 as Chapter 10's step 7 extends it: windows 37 to 48 join the state with their loads at the forecast growth and their path on the fitted line, tied by their own trend rows, and their readings are candidate rows. With no new rows, next summer's five months have the forecast's standard deviation, 0.00868 (0.01273 and 0.01247 for the two coils' own). The candidates open to the asset manager are: reference-grade readings, 0.1 K on the return air, both discharges and rack 1's inlet; the side-by-side check of unit 1's discharge sensor, ten pairs at 0.1 K, a row on the drift alone; a second logged hour of past windows from the logger's archive; the record continued through year 4; and, on the continued record, the permanent instruments at the place of the

question, water-side heat meters on both coils ($\sigma_W = 3.0\,\mathrm{kW}$) or a second discharge sensor on unit 1 (0.3 K).

13.6.2 The candidates

Table 13.13 ranks the candidates for the two questions. Each row is a candidate; the columns give the standard deviation of the hall's five months after it by equation (13.16), the share of the present or forecast width it removes, each coil's own five months, and for today's candidates the standard deviation of Γ. The widths removed are compared as bars in Figure 13.5, today's candidates in blue and next summer's in red.

Table 13.13: Candidate measurements for today's five months (present standard deviation 0.00736; coils 0.01147 and 0.01104; Γ 0.00470) and for next summer's (forecast 0.00868; coils 0.01273 and 0.01247), by the preposterior (13.16).

Candidate	sd_a, hall [–]	removed [%]	coil 1 [–]	coil 2 [–]	sd Γ [yr^{-1}]
Today's five months, July to December of year 3					
Reference-grade readings now (0.1 K: return air, both discharges, rack 1)	0.00652	11.3	0.00947	0.00966	0.00382
Side-by-side check of unit 1's discharge sensor now	0.00727	1.2	0.01086	0.01092	0.00428
A second logged hour of windows 31–36, from the archive	0.00680	7.5	0.01069	0.01039	0.00418
A second logged hour of every window 0–36	0.00670	8.9	0.01047	0.01026	0.00363
The archive's hours 31–36 and the reference-grade readings now	0.00621	15.5	0.00914	0.00937	0.00363
Next summer's five months, July to December of year 4					
The record continued, windows 37–48	0.00712	18.0	0.01070	0.01069	–
with reference-grade readings at Jul Y4 and Dec Y4	0.00620	28.6	0.00887	0.00920	–
with water-side heat meters on both coils, permanent	0.00697	19.7	0.01013	0.01027	–
with a second discharge sensor on unit 1, permanent	0.00694	20.1	0.01002	0.01062	–
with the side-by-side check now	0.00710	18.3	0.01055	0.01065	–

Nothing narrows today's answer much: the best pair of candidates takes 15.5 % off, because most of today's band is the wander the trend rows allow between July and December, which no reading in today's window removes (Section 13.4.3). The side-by-side check, a row on the drift alone, does almost nothing here, 1.2 %: the hall figures do not pass through unit 1's sensor. For next summer the record continued takes 18.0 % off the forecast. The ranking follows each row's slope on the entries of the question's gradient, the four log-derates of windows 43 and 48. The reference-grade readings read both ends of next summer's drop at a third of the logger's spread, 0.1 against 0.3 K, in exactly those columns, and take 28.6 % off. The heat meters read the split of the heat between the coils, with equal and opposite slopes on the two log-derates (Chapter 5), which the hall's total does not need, and take 19.7 %; a second discharge sensor on unit 1 takes 20.1 %. Their value for the coils' own rates is Chapter 10's, and for the crossing date Chapter 23's.

13.6.3 The recommendation

For next summer's question the next measurement is a pair of reference-grade readings at July and December of year 4, the season's two ends, the largest step, 28.6 % off the forecast band. The supply instrument should be kept in calibration, because it is the known input that takes

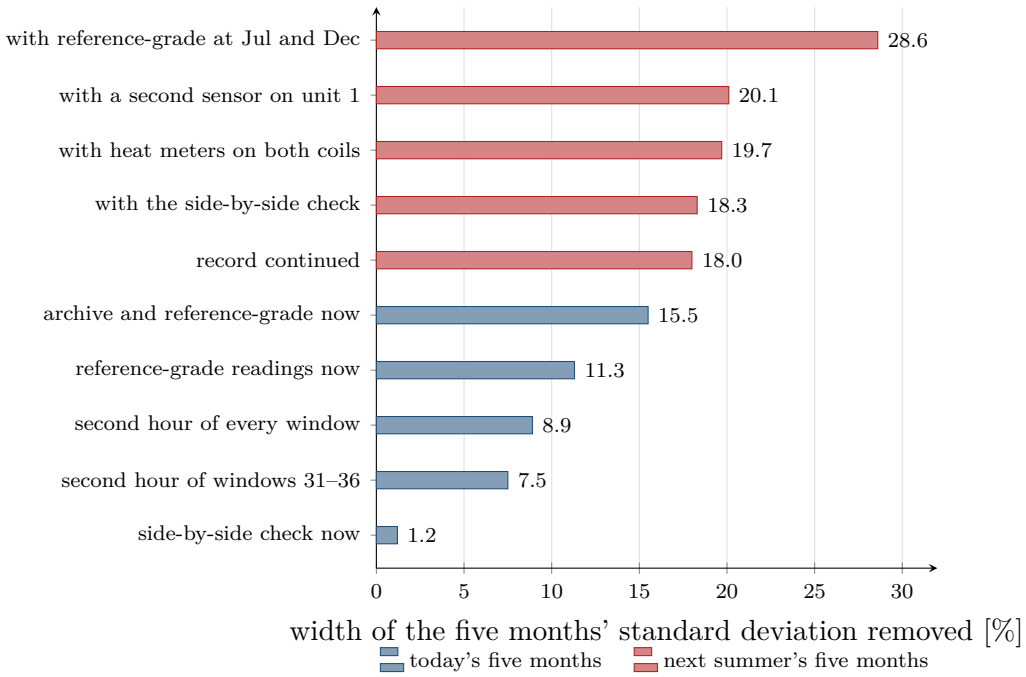


Figure 13.5: The candidate measurements of Table 13.13, ranked within each question by the width they remove from the hall's five months, by the preposterior (13.16): today's five months against their present 0.00736 (blue), and next summer's against their forecast 0.00868 (red); next summer's candidates after the first are added to the continued record.

the season out: 0.2 K of error per window would widen the hall's rate by 26.8 % (Section 13.4.3). Heat meters, if they are bought, are for the split between the coils, not for the season. The side-by-side check of unit 1's discharge sensor remains Chapter 10's recommendation for the coils' rates; it does not narrow the hall's summer.

13.6.4 What the analysis showed

The dashboard read 3.5 K and −12 % in five months, a six-month line of −24 % a year, and with a fitted season still −10 % beyond it; the season is 2.9 of the 3.5 K, a split the truth confirms, and the base drawing reads the coils at 0.3 K and 2.5 % of conductance, a summer at the record's pace, where the truth's summer was faster. The difference is the physics before it is the probability: with the supply and the other heat as inputs of every window, the season leaves the parameter without an entry spent on it, while a regression must estimate the season from the same readings that carry the aging and the noise, and can do so only over whole years and only where the season is large beside the noise. The probability adds the band, the probability that the season is the larger part, and the evidence that the coils carry no season of their own. The analysis did not test three things: the summer's pace, which the straight-line drawing pulls toward the record's rate; the split of the five months between the coils, which it cannot follow (the drawing's limit, Section 13.6.5); and a season of another shape than one annual cosine.

13.6.5 Limits of the model

A constructed record. The condition record is generated from a stated truth, with the stated sensor errors, and step 5 compares the answers with that truth (Section 13.4.5). It is not field data. Its season is one annual cosine, the same in every year, which is the regression's best case; a warm December or a changed reset of the chiller would move a fitted season and leave the known inputs right.

The known inputs. The supply and the other heat are taken as known exactly. Step 5 shows that the other heat's schedule matters little and the supply instrument a great deal; a supply sensor out of calibration would put its error into every window's log-derates.

The straight line with wander. The base drawing's path is a straight line with a monthly wander of $\sigma_q = 0.005$, which sets most of the five months' band. A line with wander pulls any five-month change toward the record's rate, so the drawing cannot read a summer faster or slower than its own line: on the truth's readings it gives the five months as -0.02096 against the true -0.03045, and its band covers the truth's five months in 0.745 of fresh records. The split between the coils is further off still. The season's terms do not touch the coils and are not affected. Whether coil 1's rate changes is Chapter 11's question; a summer's own pace needs a drawing whose rate can change.

The base fit's own limits. The drift is linear from window 20, the readings are independent across windows (Chapter 24 asks when they are not), the hall is in steady state within each window, and the airflow is pinned: Chapter 10's limits carry over.

The next chapter. Chapter 14, in Volume II, asks when the capacity margin closes, as a date with its band; the season enters there through the capacity scenario's summer supply.

13.7 What to retain

Definition 13.1 (The season separated by the known inputs)**.** A raw figure's change from window t_1 to window t_2 is the sum of four terms, equation (13.3): the season's, the known inputs' change at the commissioned coils; the IT load's, at the commissioned coils and the season's mean inputs; the coils', the fitted coils against the commissioned ones at the record's inputs; and the readings' errors, the readings less the model. Each of the first three is a quantity of the state with variance $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$. A regression's slope is a fixed linear map of the readings, $s^{\text{reg}} = \mathbf{w}^{\mathsf{T}}\mathbf{y}^{\text{raw}}$, and the same map splits it; its season term is the leak $\mathbf{w}^{\mathsf{T}}\mathbf{s}_{\text{season}}$, a number of the known inputs. The parameter's trend is read after the physics has taken the season out, at no cost in entries.

Three conclusions carry beyond the hall. A season enters a physical model as a known input, and leaves the parameter there; a regression on raw performance has to estimate it, and on a short span or across a single summer it cannot tell it from aging. A change of a raw figure between two windows can be taken apart into season, load, parameter and the readings' errors as quantities of the same posterior, which says how much of a dashboard's drop is condition. And the instrument that reads the known input, here the supply's, is the season's instrument: its calibration matters more to a trend than any sensor added for the trend itself.

Table 13.14 lists the terms of the chapter with their expressions, units and the subsection that defines each.

Table 13.14: The terms of the season's separation from aging and where each is defined. Units for the two-unit hall.

Term	Expression	Units	Meaning	Defined in
Known inputs	$T_s(t)$, $Q_o(t)$ in eq. (13.2)	°C, kW	pinned in every window; the season at no cost in entries	13.1.2
Hall's total conductance	$G_{\text{tot},t}$, eq. (13.1)	kW/K	what the hall figures see	13.1.2
Apparent conductance	$G_t^{\text{app}} = y_{Q,t}/(y_{T_r,t} - T_s(t))$, eq. (13.2)	kW/K	the load-normalized raw figure	13.1.4
Rows and maps				
Split of a change	eq. (13.3), (13.4), (13.5)	K or –	season, load, coils, reading errors	13.1.4
Regression as a linear map	$s^{\text{reg}} = \mathbf{w}^{\mathsf{T}}\mathbf{y}^{\text{raw}}$, eq. (13.6)	yr^{-1}, K/yr	the dashboard's slope	13.1.4
Season's leak into a slope	$\mathbf{w}^{\mathsf{T}}\mathbf{s}_{\text{season}}$, eq. (13.7)	yr^{-1}, K/yr	the slope the season alone makes	13.1.4
Seasonal conductance	eq. (13.9)	–	a season on the parameter itself	13.1.4
Questions and answers				
The coils' five months	eq. (13.11)	–	the hall's change, from a 4×4 block of $\boldsymbol{\Sigma}$	13.3.1
Regression's summer aging	eq. (13.12)	–	trend plus the residuals' change	13.3.3
The hall's rate	Γ, eq. (13.13)	yr^{-1}	the like-with-like reference for hall regressions	13.3.4
Evidence of a seasonal conductance	eq. (13.14), (13.15)	–	the check that the inputs carry the season	13.3.5
The season's instrument	the supply's calibration	K	widens the rate most when it errs	13.4.3
Preposterior of a candidate	eq. (13.16)	–	the band after a measurement	13.6.1

13.8 Exercises

Exercise 13.1. *Formulation.* Write the question "was this summer's loss faster than the record's rate?" in the form of Table 13.2 and evaluate it from the base fit. (A prediction under a threshold, equation (2.12) with $f = s_{\text{coils}} - 5\,\Delta t\,\Gamma$ and $f^* = 0$, the gradient that of equation (13.11) less $5\,\Delta t$ times that of equation (13.13): $\hat{f} = -0.02536 - (-0.02176) = -0.00360$; with $\text{Cov}(s_{\text{coils}}, \Gamma) = 1.8316 \times 10^{-5}$, $\text{sd}(f) = 0.00653$, and $\Pr(f < 0 \mid \mathbf{y}) = \Phi(0.5513) = 0.7093$.)

Exercise 13.2. *Derivation.* Show that a season that is an exact harmonic leaks into the line's slope as $k_c\,\mathbf{w}^{\mathsf{T}}\cos\boldsymbol{\phi} + k_s\,\mathbf{w}^{\mathsf{T}}\sin\boldsymbol{\phi}$, and evaluate the unit cosine's and sine's slopes over windows 0 to 12, 24 and 36. (By equation (13.7) the leak is $\mathbf{w}^{\mathsf{T}}\mathbf{s}_{\text{season}}$, linear in the series; with $s_{\text{season},t} = k_c\cos\phi_t + k_s\sin\phi_t$ plus a constant, which $\mathbf{w}$ takes to zero, it is $k_c\,\mathbf{w}^{\mathsf{T}}\cos\boldsymbol{\phi} + k_s\,\mathbf{w}^{\mathsf{T}}\sin\boldsymbol{\phi}$. The unit cosine's slope is -0.76425, -0.21399 and -0.09893 per yr, the sine's -1.32372, -0.37064 and -0.17135; with the tile's $k_c = -0.02245$ the leak over 36 windows is $+0.00222$ per yr.)

Exercise 13.3. *Derivation.* From equation (13.2), show that the return air's season between windows 31 and 36 is $\Delta T_s + \Delta Q_o/G_{\text{tot},0}$, and evaluate it. (With $g_{u,t} = g_{u,0}$, $\check{T}_{r,t} = T_s(t) + (Q_t + Q_o(t))/G_{\text{tot},0}$ and at the mean inputs $\check{T}^{\text{m}}_{r,t} = 14.00 + (Q_t + 8.00)/G_{\text{tot},0}$, so $\check{T}_{r,t} - \check{T}^{\text{m}}_{r,t} = T_s(t) - 14.00 + (Q_o(t) - 8.00)/G_{\text{tot},0}$, free of the load; between the two windows $2.5095 + 5.0191/13.7893 = 2.5095 + 0.3640 = 2.8735$ K.)

Exercise 13.4. *Application.* Refit the seasonal regression of $\ln G^{\text{app}}$ over windows 1 to 36, leaving the commissioning test out, and set the slope beside the whole record's -4.742 %/yr.

(Slope −0.04959 per yr, standard error 0.00825, fitted cosine −0.01023 and sine +0.00994: −4.838 %/yr.)

Exercise 13.5. *Application.* Remove the IT load's part from the return air's seasonal slope, and convert the rest to a rate of conductance the practitioner's way, over the mean return approach of windows 1 to 36. (By Table 13.5, $0.8628 - 0.2970 = 0.5657\,\mathrm{K/yr}$; over $10.5983\,\mathrm{K}$ it is −0.05338 per yr, −5.198 %/yr.)

Appendix A

Atlas of asset-management questions

This appendix can be read independently of the chapters. Each row is a question an asset manager asks. The first column gives the question in the asset manager's words, as the chapter's questions table gives it; the second the general form of its category, written out or cited by equation number; the third its mathematical form for the plant of that chapter, once the plant graph and its state $\mathbf{z}$ are fixed, in that chapter's symbols; the fourth its category and the fifth its computation, the five categories and the five computations of Chapter 2; and the last the section in which the question is answered, or the chapter when a whole chapter answers it. The rows grow as chapters are added.

Asset manager's question	General form	Mathematical form	Category	Computation	Answered in
Chapter 1. Physical Assets as Graphs					
What temperature is the hall running at now?	$\mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}]$, $\mathrm{Var}(f(\mathbf{z}) \mid \mathbf{y})$	$p(T \mid \mathbf{y})$, $T = z_1$	estimation	posterior	1.3.1
How far is the as-commissioned conductance from its design value?	$\mathbb{E}[f(\mathbf{z}) \mid \mathbf{y}]$, $\mathrm{Var}(f(\mathbf{z}) \mid \mathbf{y})$, $\Pr(f(\mathbf{z}) \le f^* \mid \mathbf{y})$	$p(h \mid \mathbf{y})$, $h = G/G_{\mathrm{ref}} = \mathrm{e}^{z_2}$; $\Pr(h < 1 \mid \mathbf{y})$	estimation of a derived quantity	posterior	1.3.1
Is the hall within its 24°C limit now?	$\Pr(f(\mathbf{z}) < f^* \mid \mathbf{y})$	$\Pr(T < 24^\circ\mathrm{C} \mid \mathbf{y})$	prediction under a threshold	posterior	1.3.2
Can the path carry 60 kW within 27.5°C?	$\Pr(f(\mathbf{z}) < f^* \mid \mathbf{y})$	$\Pr(T_2 < 27.5^\circ\mathrm{C} \mid \mathbf{y})$, $T_2 = T_c + (Q_2/G_{\mathrm{ref}})\,\mathrm{e}^{-z_2}$	prediction under a threshold	posterior	1.3.2
Which added sensor would narrow the health index?	preposterior covariance $(\boldsymbol{\Sigma}^{-1} + \mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a)^{-1}$	$\boldsymbol{\Sigma}_a$ for each candidate row	design	preposterior	1.6.1
Chapter 2. The Language of Asset-Management Questions					
Where does the air path stand against its design sheet?	(2.9), (2.10)	$p(h \mid \mathbf{y})$, $h = \mathrm{e}^{z_3}$: $\hat{z}_3$ and $\sqrt{\Sigma_{33}}$ through $\mathbf{g} = (0,\ 0,\ h)^{\mathsf{T}}$	estimation	posterior	2.4.1
Can the hall carry the requested 60 kW below 27.5°C?	(2.12)	$\Pr(T_2 < 27.5^\circ\mathrm{C} \mid \mathbf{y})$, the normal of the entry T_2	prediction under a threshold	posterior	2.4.2
Will the summer season pass without a corrective work order?	(2.13)	$\mathbb{E}[\mathrm{e}^{-\lambda\tau} \mid \mathbf{y}]$ over $\mathcal{N}(\hat{\omega}, \sigma_\omega^2)$; $(b'/(b'+\tau))^{a'}$ exact	prediction under a threshold	posterior	2.4.4

Asset manager's question	General form	Mathematical form	Category	Computation	Answered in
How fast is the conductance falling, and when does it reach its functional limit?	(2.21)	$\Pr(\gamma < 0 \mid \mathbf{y})$; $\Pr(t^* \le s \mid \mathbf{y}) = \Pr(\gamma < 0,\ g_s \le g_F \mid \mathbf{y})$	prediction under a threshold; estimation for the crossing date	trend	2.3.4
Will both replacements be needed this year?	(2.12), an intersection	$\Pr(F_A)\ \Pr(F_B \mid F_A)$	prediction under a threshold	posterior	2.4.7
Is the air path fouled, or is the sensor drifting?	(2.14), (2.15)	$\Pr(M_j \mid \mathbf{y}) \propto p(\mathbf{y} \mid M_j)\ \Pr(M_j)$	hypothesis comparison; diagnosis	evidence	2.5.1, 2.5.3
Change the filters before the summer peak, or wait?	(2.23)	$\mathcal{R}(a) = c_a + C_e\ \Pr(e \mid \mathbf{y}, a)$, $a^* = \arg\min_a \mathcal{R}(a)$	design	expected consequence	2.5.4, 2.5.5
Is another condition test or an inspection worth scheduling?	(2.16)	$\mathbf{\Sigma}_a = (\mathbf{\Sigma}^{-1} + \mathbf{J}_a^{\mathsf{T}}\mathbf{J}_a)^{-1}$; EVSI	design	preposterior; expected consequence	2.6.1, 2.6.2
Chapter 3. Uncertainty and Conditional Probability					
Will one test read the coil below 5.0 kW/K when nothing has changed?	(2.13)	$\Pr(y_T > T_c + Q/G^* \mid T)$	prediction under a threshold	posterior	3.1.1
What is known about the coil's conductance, before and after the commissioning test?	(2.9)	$p(g \mid \mathbf{y})$; $\hat{G}$, $h = \mathrm{e}^g$ and band	estimation	posterior	3.1.2
Will the next test read below 5.0 kW/K, and do more tests remove that?	(2.13)	$\int \Pr(y_{\text{new}} > c \mid T)\, p(T \mid \mathbf{y})\, \mathrm{d}T$	prediction under a threshold	posterior	3.1.3
Is a flagged coil fouled?	(2.15)	$\Pr(F \mid H)$	diagnosis	evidence	3.3.3
Which inspection grade is a flagged coil in?	(2.15)	$\Pr(S = s \mid H)$	diagnosis	evidence	3.4.1
Will the unit run the next year without a corrective failure?	(2.13)	$\int \mathrm{e}^{-\lambda\tau} p(\lambda \mid \mathbf{y})\, \mathrm{d}\lambda$	prediction under a threshold	posterior	3.4.2
A fouled coil was found; should the crew open the fan?	(2.12)	$\Pr(F_B \mid F_A)$	prediction under a threshold	posterior	3.4.6
Does the hall exceed its limit after the test?	(2.12)	$\Pr(T > T_{\max} \mid y_T)$	prediction under a threshold	posterior	3.5.5
How tight is the health index from two tests by one sensor?	(2.9)	$\sqrt{\Sigma_{gg}}$ with $\mathbf{V}_y$ full, or b_T in the state	estimation	posterior	3.5.7
How many corrective work orders next year, and does a longer log help?	(2.10)	$\mathrm{Var}(K_\tau \mid \mathbf{y})$, (3.35)	prediction under a threshold	posterior	3.6.2

Asset manager's question	General form	Mathematical form	Category	Computation	Answered in
How well is the change since commissioning known?	(2.10)	$\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$, $\mathbf{g} = (-1,\ 1)^{\mathsf{T}}$	estimation	posterior	3.6.4
Clean path or blocked filter; what is the hall's exceedance?	(2.14)	(3.39)	hypothesis comparison	evidence	3.6.6
Chapter 4. The Anatomy of Probabilistic Inference					
What is the thermal resistance of the air path at commissioning, and how well is it known?	(2.9)	$p(z_i \mid \mathbf{y})$: $\hat{z}_i$ and Σ_{ii}	estimation	posterior	4
What will the hall temperature be at full load?	(2.10)	$f(\hat{\mathbf{z}})$ and $\mathbf{g}^{\mathsf{T}}\boldsymbol{\Sigma}\mathbf{g}$ by the delta method	estimation	posterior	4
Can the hall be released for 60 kW within the 28°C acceptance limit?	(2.12)	$\Pr(T_2 < T_{\max} \mid \mathbf{y}) = \Phi\big((T_{\max} - \hat{z}_3)/\sqrt{\Sigma_{33}}\big)$	prediction under a threshold	posterior	4
Is the conductance falling between two tests, and how fast?	(2.21)	$\Pr(\gamma < 0 \mid \mathbf{y})$: a normal tail on the trend entry; $\mathrm{Var}(\gamma)$ from the two windows' variances and $n_{\mathrm{w}}\sigma_q^2$	prediction under a threshold	trend	4.6.5
Chapter 5. Health Indices from Physical Parameters					
How far has this coil's conductance fallen from its commissioned value?	(2.9)–(2.11)	$p(h_{u,18} \mid \mathbf{y})$, $\ln h_{u,18} = g_{u,18} - g_{u,0}$: median and 5–95 % band	estimation	posterior	5.3.1
Has it lost any conductance at all, beyond what the readings' scatter allows?	(2.12)	$\Pr(g_{u,18} - g_{u,0} < 0 \mid \mathbf{y})$, a normal tail	prediction under a threshold	posterior	5.3.2
Which coil has fallen more, and how sure is that?	(2.12)	$\Pr(\ln h_{2,18} < \ln h_{1,18} \mid \mathbf{y})$	prediction under a threshold	posterior	5.3.2
Is the dashboard's 21 % and 29 % the coils, or the load and the season?	(2.10), (2.11)	$p(h_u^{\mathrm{raw}} \mid \mathbf{y})$, $h_u^{\mathrm{raw}} = A_{u,0}(\mathbf{z})/A_{u,18}(\mathbf{z})$	estimation	posterior	5.3.3
Do the coils still carry the hall's capacity scenario today?	(2.12)	$\Pr(G_{1,18} + G_{2,18} \geq G^* \mid \mathbf{y})$	prediction under a threshold	posterior	5.3.4
Chapter 6. Hidden Degradation Under Control					
The unit still meets its set-point; is it working harder to do so?	(2.12), (2.10), (2.11)	$\Pr(\eta_{A,u} > 0 \mid \mathbf{y})$, $\eta_{A,u} = A_{u,18}(\mathbf{z}) - \breve{A}_u(\mathbf{z})$: median and 5–95 % band	prediction under a threshold	posterior	6.3.1
How much colder must the chilled water run, because of the coils?	(2.9)–(2.11)	$p(\eta_s \mid \mathbf{y})$, $\eta_s = T_{r,18} - \breve{T}_r$; $\Pr(\eta_s > 0 \mid \mathbf{y})$	estimation	posterior	6.3.2

Asset manager's question	General form	Mathematical form	Category	Computation	Answered in
Why does the cold aisle still read as it did at handover?	(2.10), (2.11)	$T_{c,18} - T_{c,0} = s_{\text{season}} + s_{\text{load}} + \eta_s$, each term's $p(\cdot \mid \mathbf{y})$	estimation	posterior	6.3.3
Which unit carries the heat the fouled coil no longer passes?	(2.12)	$\Pr(\eta_{D,1} > 0 \mid \mathbf{y})$, $\eta_{D,1} = -\eta_{D,2}$	prediction under a threshold	posterior	6.3.4
Chapter 7. Distinguishing Degradation from Sensor Drift					
The approach temperature has risen 1.5 K; is the exchanger fouled or is the sensor drifting?	(2.14), (2.15)	$B(M_{\text{foul}} : M_{\text{drift}})$ by the Laplace evidence; $\Pr(M \mid \mathbf{y})$ over M_{neither}, M_{foul}, M_{drift}, M_{both}	hypothesis comparison	evidence	7.3.1
How much of the 1.5 K is in question at all, once the load, the season and the other coil are taken out?	(2.10)	$A_{1,25} - A_{1,0}$ split into heat, coil 2, coil 1, sensor and reading errors	estimation	posterior	7.3.2
Could it be both, and how much of each?	(2.10), (2.11)	$p(\ln h_{1,25}, b_{d,1} \mid \mathbf{y}, M_{\text{both}})$	estimation	posterior	7.3.3
What is coil 1's health index once the drift is allowed for, and has it lost conductance?	(2.11), (2.12)	$p(h_{1,25} \mid \mathbf{y}, M_{\text{both}})$; $\Pr(\ln h_{1,25} < 0 \mid \mathbf{y}, M_{\text{both}})$; the model average beside it	estimation	posterior	7.3.4
Would the readings themselves have flagged a faulty sensor?	(2.12)	$\Pr(\chi^2_{n_r - n_z} > 2\mathcal{L})$ and the largest $\lvert r_k^{\text{N}} \rvert$ under Chapter 5's drawing	prediction under a threshold	posterior	7.3.5
Chapter 8. Battery State of Health					
What is the usable capacity now, and the internal resistance?	(2.9)–(2.11)	$p(C_{\text{bat},36}, R_{\text{bat},36} \mid \mathbf{y})$: the correlation first, then each median and 5–95 % band	estimation	posterior	8.3.1
How much capacity has the string lost since commissioning, and how far has its resistance risen?	(2.10), (2.11), (2.12)	$\ln h_C = \xi_{36} - \xi_0$, ΔR; $\Pr(\Delta R > 0 \mid \mathbf{y})$	estimation; prediction under a threshold for $\Pr(\Delta R > 0)$	posterior	8.3.2
Is the dashboard's 83.2 % the capacity?	(2.10), (2.11)	$p(T_{\text{run},36}/T_{\text{run},0} \mid \mathbf{y})$: $\ln h_C$ plus the log-ratio of the loaded heads	estimation	posterior	8.3.3
Is the string still above the replacement criterion of 80 %?	(2.12)	$\Pr(h_C < 0.80 \mid \mathbf{y})$; $\Pr(C_{\text{bat},36} < 80\,\text{Ah} \mid \mathbf{y})$	prediction under a threshold	posterior	8.3.4
Chapter 9. Fleet Pooling and Hierarchical Models					
Where does this unit sit among its fifty siblings?	(2.9)–(2.11)	$p(\mathcal{K}_0 \mid \mathbf{y})$; $p(\pi_0 \mid \mathbf{y})$, from composition draws	estimation	posterior	9.3.1
Is it fouling faster than the design's typical coil, or as slowly as the chart says?	(2.12)	$\Pr(\gamma_0 < \mu_\gamma \mid \mathbf{y})$; $\Pr(\pi_0 > 0.9 \mid \mathbf{y})$; $\Pr(\gamma_0 > \hat{\gamma}_0^{\text{chart}} \mid \mathbf{y})$	prediction under a threshold	posterior	9.3.1

Asset manager's question	General form	Mathematical form	Category	Computation	Answered in
How fast is it fouling, and what is its health index today, read with the fleet?	(2.9)–(2.11)	$p(\gamma_0 \mid \mathbf{y})$ by mixture quantiles; $d_0(3)$, $h_0(3) = \mathrm{e}^{3\gamma_0}$	estimation	posterior	9.3.2
How fast does this design foul, how alike are its coils, and what does one coil tell the fleet?	(2.9), (2.13)	$p(\mu_\gamma, \tau_\gamma, \mu_\mathrm{c}, \tau_\mathrm{c} \mid \mathbf{y})$; $p(\gamma_\mathrm{new} \mid \mathbf{y})$	estimation, prediction	posterior	9.3.3
Are the coils alike enough to share one line, or different enough to need their own?	(2.14)	$\ln p(\mathbf{y} \mid M)$ of the separate, hierarchical and fully pooled drawings	hypothesis comparison	evidence	9.3.4
Chapter 10. Degradation Rates from Condition Records					
How many percent of conductance per year is it losing?	(2.21), first line; (2.9), (2.11)	$p(\gamma_u \mid \mathbf{y})$, read as $100(\mathrm{e}^{\gamma_u} - 1)\,\%$/yr: median and 5–95 % band, $u = 1, 2$	estimation	trend	10.3.1
Is it losing conductance at all, beyond three years of scatter?	(2.21), first line	$\Pr(\gamma_u < 0 \mid \mathbf{y})$	prediction under a threshold	trend	10.3.1
Which coil is losing conductance faster, and how sure is that?	(2.12)	$\Pr(\gamma_2 < \gamma_1 \mid \mathbf{y})$	prediction under a threshold	trend	10.3.2
Is the dashboard's 15 % a year the coils, or the load, the season and the sensor?	(2.10)	$s_u^{\mathrm{raw}}(\mathbf{z})$, the slope of the line through $h_{u,t}^{\mathrm{raw}}(\mathbf{z})$, split into load and season, coils, sensor and reading errors	estimation	trend	10.3.3
How far has each coil fallen since commissioning, read with the whole record?	(2.9), (2.11)	$p(h_{u,36} \mid \mathbf{y})$, $\ln h_{u,36} = g_{u,36} - g_{u,0}$	estimation	posterior	10.3.4
Does the record still need the drift of unit 1's sensor, and how large is it now?	(2.14); (2.9)	$B(\text{drift} : \text{none})$; $p(b_{d,1}(36) \mid \mathbf{y})$	hypothesis comparison	evidence	10.3.5
Chapter 11. Accelerating Degradation					
Is the loss accelerating?	(2.12) on the curvature; (2.21), first line, one order up	$\Pr(\dot{\gamma}_u < 0 \mid \mathbf{y}, M_{12})$, $u = 1, 2$	prediction under a threshold	trend	11.3.1
Does the record need a changing rate at all, or does the constant rate do?	(2.14), (2.47)	$B(M_u : M_0)$; $\Pr(M \mid \mathbf{y})$ over $\{M_0, M_1, M_2, M_{12}\}$; accelerating, constant and slowing as the family's split	hypothesis comparison	evidence	11.3.2
How fast is it losing now, against the record's average?	(2.9)–(2.11), over the family	$\gamma_u(3) = \gamma_u + \dot{\gamma}_u\,(3 - \bar{t}_\mathrm{y})$ in %/yr, averaged over the four drawings	estimation	trend	11.3.3

Asset manager's question	General form	Mathematical form	Category	Computation	Answered in
Is the dashboard's bend in year 3 the coil?	(2.10)	$\dot{s}_u^{\mathrm{raw}}(\mathbf{z})$, the bend of the quadratic through $h_{u,t}^{\mathrm{raw}}(\mathbf{z})$, split into load and season, coils, sensor and reading errors	estimation	trend	11.3.4
With the coil allowed to accelerate, does unit 1's sensor still need its drift?	(2.14)	$B(\text{drift} : \text{none} \mid M_1)$; the drift alone, the acceleration alone, both	hypothesis comparison	evidence	11.3.5
Chapter 12. Change-Point Detection					
Did something happen in March?	(2.15) at the named window; (2.14)	$\Pr(t_0 = 15 \mid \mathbf{y}) = \bar{B}_{15}/(1+\bar{B}_{15})$, $\bar{B}_{15} = \frac{1}{8}\sum_j B_{j,15}$	diagnosis	evidence	12.3.1
What stepped: a coil, or a sensor reading it?	(2.15)	$\Pr(M_{j,15} \mid \mathbf{y}, t_0 = 15)$ over the eight kinds j	diagnosis	evidence	12.3.2
Was it March, or the month either side?	(2.15) over windows	$\Pr(t_0 = s \mid \mathbf{y}, \text{coil 2})$, $s \in \{14, 15, 16\}$	diagnosis	evidence	12.3.3
How large was the step, and how much of coil 2's loss is it?	(2.9), (2.11), (2.10)	$p(\delta_2 \mid \mathbf{y}, M_{2,15})$, as $100(\mathrm{e}^{\delta_2} - 1)\,\%$ of conductance; $\delta_2 / \ln h_{2,36}$	estimation	posterior	12.3.4
With the step allowed for, how fast is each coil losing conductance, and which is faster?	(2.21), first line; (2.12); (7.10)	$p(\gamma_u \mid \mathbf{y})$ mixed over the named-window family; $\Pr(\gamma_2 < \gamma_1 \mid \mathbf{y})$	estimation	trend	12.3.5
Did a step happen anywhere in the three years?	(2.15) over the whole record	$\Pr(\text{a step} \mid \mathbf{y}) = \bar{B}/(1+\bar{B})$ over windows 1–36; the ranking of the 288 candidates	diagnosis	evidence	12.3.6
Chapter 13. Separating Seasonal Effects from Aging					
Is the summer drop in performance degradation or just the ambient?	(2.10), (2.11)	$y_{36} - y_{31} = s_{\text{season}} + s_{\text{load}} + s_{\text{coils}} + s_{\text{err}}$, each term's $p(\cdot \mid \mathbf{y})$; on $\ln G^{\text{app}}$, $s_{\text{coils}} = \ln(G_{\text{tot},36}/G_{\text{tot},31})$	estimation	trend	13.3.1
Is more of the drop the coils than the season?	(2.12)	$\Pr(s_{\text{coils}} < s_{\text{season}} \mid \mathbf{y})$ on $\ln G^{\text{app}}$	prediction under a threshold	trend	13.3.1
What does the trend line through the raw performance read, and how much of it is the season?	(2.10)	$s^{\text{line}} = (\mathbf{w}^{\text{line}})^{\mathsf{T}}\mathbf{y}^{\text{raw}}$ over windows 0 to t_{as}, split into season (the leak), load, coils and errors	estimation	trend	13.3.2
Does a seasonal term in the regression take the season out?	(2.10)	$s^{\text{seas}} = (\mathbf{w}^{\text{seas}})^{\mathsf{T}}\mathbf{y}^{\text{raw}}$, split the same way; its fitted harmonic against the known inputs'	estimation	trend	13.3.3
How fast is the hall's conductance falling, with the season taken out by the physics?	(2.21), first line; (2.11)	$p(\Gamma \mid \mathbf{y})$, $\Gamma = \ln(G_{\text{tot},36}/G_{\text{tot},0})$ over $36\,\Delta t$, as $100(\mathrm{e}^{\Gamma} - 1)\,\%/\text{yr}$	estimation	trend	13.3.4
Does the coils' conductance itself fall in summer?	(2.14)	$B(M_{\text{sea}} : M_{\text{base}})$; $p(g_u^c, g_u^s \mid \mathbf{y})$	hypothesis comparison	evidence	13.3.5