INTERACTIVE CONTROL ENGINEERING
Principles and Applications

Volume 19 Control of Biological Systems and Human Performance

CHAPTER 14

Estimating and Steering an Epidemic From Week-Old, Partial Counts

Pages 351–372

Edited by John D. Hedengren

How to cite this chapter. "Estimating and Steering an Epidemic From Week-Old, Partial Counts," in *Control of Biological Systems and Human Performance*, John D. Hedengren, Ed., vol. 19 of *Interactive Control Engineering: Principles and Applications*. Advanced Process Solutions, LLC, 2026, ch. 14, pp. 351–372.

The page numbers in this file are those of the printed volume.

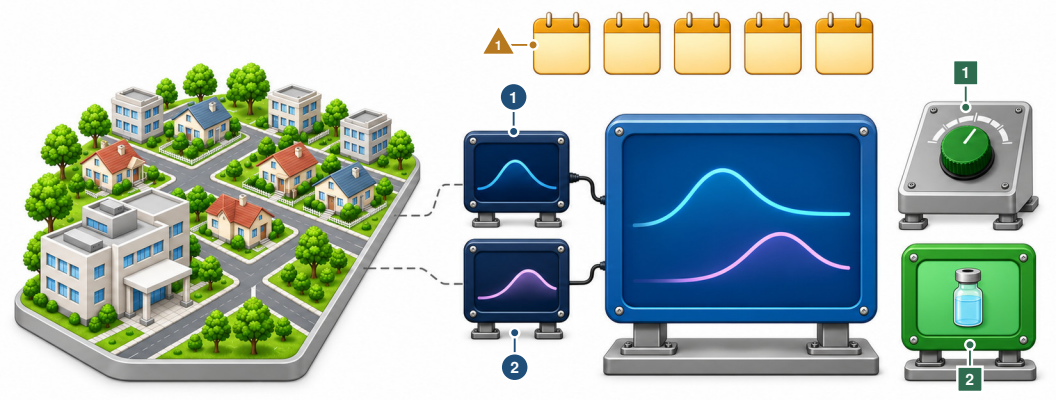
The app for this chapter
https://apmonitor.com/apps/epidemic-control/

The volume online: contents, other chapters and citation records
https://apmonitor.com/apps/books/v19-biological-human/#epidemic-control

Copyright © 2026 Advanced Process Solutions, LLC. All rights reserved. Produced by Advanced Process Solutions, LLC.

CHAPTER 14

Estimating and Steering an Epidemic From Week-Old, Partial Counts



The controller moves 1 the intervention dial, standing for every measure that cuts contact
2 vaccine doses a day (advanced)

The controller reads 1 reported cases, a week late and only a share of them
2 hospital census, the beds in use, counted daily

What disturbs it 1 a week between an onset and its report ▲ a testing change that moves the count, not the epidemic ▲ a variant or a holiday, unannounced

A teaching city publishes its case reports and its hospital census to the screens long after the infections they describe, while the dial acts on the city of today: notice the row of calendar days between what happens and what is seen. Illustration, not to scale.

IN PLAIN TERMS **Steering an Epidemic by Late, Partial Counts**

The main idea. A city learns about an epidemic from two counts. One is the cases its tests find, a week late and only a share of them. The other is the hospital beds in use. Measures that cut contact slow the spread at once, but the counts show it weeks later.

The challenge. A change in testing can halve the case count with no change in the epidemic. A more contagious strain can also arrive unannounced. Beds in use must stay below what hospitals hold, so whoever sets the measures must judge the present from old numbers.

Why it matters. Acting on old numbers means acting late, and each day of waiting fills more beds.

THREE WAYS TO READ THIS CHAPTER

Explore	12 min	the first page, the history segment on 1918 (Section 14.1), the history segment on 1927 (Section 14.2), the fact on a count given as a range (Section 14.3), the two try-it segments and the fact on a retuned threshold (Section 14.7), and the key points
Study	45 min	the whole chapter with its three pause points and the twelve exercises at its end
Practice	20 min	the first page, the cases and the graded 180 days (Section 14.7), Technology in Practice (Section 14.9), Limitations (Section 14.10) and the open challenge

The control question A city counts its cases a week after the onsets they sample, finds only a share of them, and fills its hospital beds two weeks after the infections. What can an estimator recover from those counts, and which policy keeps the census under the beds when testing changes and a variant arrives unannounced?

Learning objectives

- Derive the early growth rate, the peak and the final size of the uncontrolled city, and add up the delays between the dial and the beds.
- Explain why a relay on the reported count cycles and a PI loop on the census overshoots its target.
- Compare an EKF, a particle filter and an MHE on one augmented state, and tell where the census decides the result and where a bound does.
- Set up the economic MPC and the scenario tree, and count their moves, slacks and shared days.
- Judge whether a scenario tree buys anything a hand-set planning margin does not.

Prerequisites Sections 1.4.3 (epidemic growth rate), 1.4.5 (sparse and late samples), 1.6.2 (velocity form), 1.6.3 (MPC and MHE) and 1.6.4 (scenarios). Chapter 11 for a loop on a late laboratory result.

The app https://apmonitor.com/apps/epidemic-control/

See also Volume 5, Chapter 4 for multirate estimation on a late analyzer and laboratory, Volume 5, Chapter 13 for an extended Kalman filter that applies late fixes at the instant they describe. Volume 9, Chapter 7 for a scenario tree with non-anticipativity on an ensemble forecast and Volume 9, Chapter 6 for a predictive controller that sends heat ahead of a delay.



14.1 Introduction

About 140 people are infected in a teaching city of one million when the simulator's first case begins. Left alone, the epidemic infects 94 % of the city and fills 7,201 hospital beds on day 73, against a capacity of 1,500. The reported 7-day average of new cases passes 300 a day on day 38, and a dial set to 70 % the next day turns the true infections at once. The reported average keeps rising for 20 more days, to 1,794 a day on day 59, and the census peaks at 1,435 beds on day 62. Waiting a day longer costs about 182 beds at the peak, and waiting two days breaches the capacity for 23 days.

> ⧖ HISTORY **Measures in Seventeen Cities (1918)**
>
> A study published in 2007 gathered the dates on which 17 cities of the United States brought in 19 classes of public health measures during the influenza pandemic of 1918, among them the closing of schools, churches and theaters. Cities that put several measures in place at an early phase of their epidemic had peak death rates about 50 % lower than cities that did not, and less steep epidemic curves. The difference in total excess deaths was smaller, about 20 %, and less certain statistically. Few cities kept their measures longer than six weeks, and the findings supported the view that spread resumes once measures are relaxed [1]. When federal planning guidance for community measures was drafted in 2007, evidence of this kind, historical analyses and observations rather than controlled studies, was most of what there was [2].

The delays are not the only obstacle. A reported case is a sample of the onsets of a week before, scaled by an ascertainment that no one measures, and a change in testing moves the count without moving the epidemic. Counts that arrive late and partial are the data of every surveillance system, and the statistics of a reporting delay were worked out for events that occur before they are reported [3]. Process plants have the same shape. A laboratory assay reports a stream's composition hours after the sample, an analyzer drifts against a slower unbiased measurement, and a constraint binds on a quantity no actuator moves directly. The kraft digester of Volume 5, Chapter 4 grades its cook on exactly such an analyzer and laboratory. The estimators that treat such plants, the extended Kalman filter (EKF), the particle filter (PF) and the moving horizon estimator (MHE), have been set against each other on process examples where the states must stay physical [4], and the constrained MHE has a stability analysis of its own [5]. The belief this city overturns is that a constrained estimator

wins whenever the data change their meaning. In this model the census decides that question, and the constraint matters only where the data fall far outside what the model expects.

A single plant on which all of this can be watched at once is rare. Work from the control community fits a compartmental model to a real outbreak and designs estimation, control or both for it [6–8], and the epidemiological literature estimates the reproduction number from reported incidence without a controller in the loop [9]. This chapter sets three estimators and four policies against one declared city, with its truth logged beside every count, and measures what each can and cannot do. The compartmental model, the three estimators and model predictive control (MPC) are established methods. Four features of the implementation carry the lessons: an augmented state with a chain of onsets that lets one model and one noise description serve all three estimators, an MHE solved daily by Gauss–Newton steps that are each a quadratic program, a scenario-tree MPC solved by successive linearization, and a measurement boundary that keeps every estimator and controller away from the truth. Two findings on this teaching model shape its lessons: the census anchors every estimator through a testing cut, and the scenario tree's protection against a variant is a margin that a hand-set margin also buys, for less. Ten cases, each graded by a completion rule, carry the experiments.

Section 14.2 traces the problem's history. Section 14.3 declares the city and the two counts it publishes, and Section 14.4 derives its growth rate, its final size and the three delays between the dial and the beds. Section 14.5 sets out the three estimators on one augmented state, and Section 14.6 the relay, the PI loop, the economic MPC and the scenario tree. Section 14.7 measures the ten cases, Section 14.8 says which cases carry the lesson, Section 14.9 names the surveillance streams, the measures and the people who decide in a real system, and Sections 14.10 and 14.11 state the limitations and the conclusions. The exercises in Section 14.12 work through the cases.

14.2 Compartments, Undercounts and Planned Measures

The model the simulator integrates descends from a paper of 1927, as the history box below tells. Seven decades later a review of the field set out the threshold theorems for such models in terms of the basic reproduction number R_0, the contact number and the replacement number, for the classic SIR models and for SEIR models with an exposed class [10]. The early growth of an epidemic and its reproduction number are related through the distribution of the generation interval, and the equation that applies depends on that shape [11].

> ⧖ HISTORY **One Formula for the Final Size (1927)**
>
> A paper of 1927 by Kermack and McKendrick set out the problem in its first paragraph: one or more infected people enter a community more or less susceptible to a disease, the disease spreads by contact, and each infected person is in the end removed by recovery or by death [12]. Its formula for the final size of an epidemic rested on a simple model with exponentially distributed infectious periods. A paper of 2006 reviewed the later work and showed the formula unchanged by a latent stage, by any number of infectious stages or by a stage of isolation, though it fails with certain kinds of uneven mixing. The size of an epidemic of a new respiratory pathogen can therefore be estimated as soon as its basic reproduction number can be approximated [13], and the simulator, with its latent stage, checks its integrator against the 1927 relation.

Counting came next as a problem in its own right. A count of events that have occurred but are not yet reported can be corrected for the delay with a model of how reports arrive [3]. Only a share of infections is ever reported, and under-reporting in epidemics became a modeling problem of its own [14]. A framework with software for estimating a reproduction number that varies in time followed in 2013 [9].

State estimation reached epidemic models from signal processing and from the process industries. The bootstrap particle filter of 1993 carried a cloud of samples through a nonlinear model and weighted each by the data [15]. Regularized filters added progressive correction, which applies a sharp likelihood in stages so that the cloud does not collapse [16], and filtering was turned into maximum

Table 14.1 Each Variable's Role and Its Value When Case 1 Begins

Variable	Role	Start	Range
H	constrained state (true census)	0 beds	$\le H_{\max} = 1500$
$\tilde{H}_k$	measurement, census count	0 beds	daily
C_k	measurement, reported cases	0 a day	daily
u	MV, intervention dial	0	0–1
D	MV, vaccine doses (advanced)	0	0–5000 a day
m	DV, variant multiplier	1	1–2.5
h	DV, holiday multiplier	1	1–2
$\bar{\rho}$	DV, testing level	0.30	cut $\times 0.5$
outage	DV, reports held	none	1–10 d
S, E, I, R	states	$N-140$, 100, 40, 0	people
u_e	state, effect of the dial	0	0–1
c	ledger, cost	\$0	\$M
β, ρ	hidden parameters	0.50/d, 0.30	truth
$\hat{x}, \hat{R}_t$	estimates	prior	per estimator

likelihood estimation for partially observed nonlinear stochastic systems, with a disease's dynamics as the worked application [17]. In process control, moving horizon estimation with constraints and an arrival cost received a stability analysis in 2003 [5], and a critical evaluation of the two estimators on process examples argued for the constrained MHE where the EKF's estimates can leave the physical region [4].

Planning the measures is the newest layer. Economic MPC optimizes an economic cost rather than tracking a setpoint [18], which suits a dial whose cost is lost output. Multi-stage MPC represents uncertainty by a tree of scenarios whose decisions may differ only once the branches can be told apart, and it was applied to a semi-batch polymerization reactor under uncertainty [19]. The reservoir of Volume 9, Chapter 7 plans its release on a scenario tree built from an ensemble of inflow forecasts. More recently the process systems and control communities applied modeling, state estimation and optimal control to an epidemic [6], robust and optimal predictive control to the same kind of outbreak [7], and an optimal predictive strategy to distancing policies [8]. These studies are context only. The simulator models no real place, period or pathogen, and every value in it is a declared teaching value.

14.3 A Declared City and What It Publishes

The simulated city is one well-mixed population of $N = 10^6$ people in days, as shown in Figure 14.1. Everything inside the dashed frame is teaching truth, logged and drawn for the learner but read by no estimator or controller. Two counts leave the city once a day, the reported cases C_k and the census count $\tilde{H}_k$, and they are the only path from the city to the estimators. The intervention dial u is the manipulated variable, chosen once a day by the policy in force. Each variable's role is given in Table 14.1.

14.3.1 Compartments, Transmission and the Dial

Susceptible people are infected at the incidence f, pass a latent period in E, are infectious in I, and leave either for the hospital or for the removed class, as shown in Equation (14.1).

$$\begin{aligned} f &= \beta(1-u_e)SI/N, \qquad \dot{S} = -f - DS/N \\ \dot{E} &= f - \sigma E, \qquad \dot{I} = \sigma E - \gamma I \\ \dot{H} &= p_h \gamma I - H/L_h \\ \dot{R} &= (1-p_h)\gamma I + H/L_h + DS/N \\ \tau_u \dot{u}_e &= u - u_e, \qquad \dot{c} = c_u u^2 \end{aligned} \tag{14.1}$$

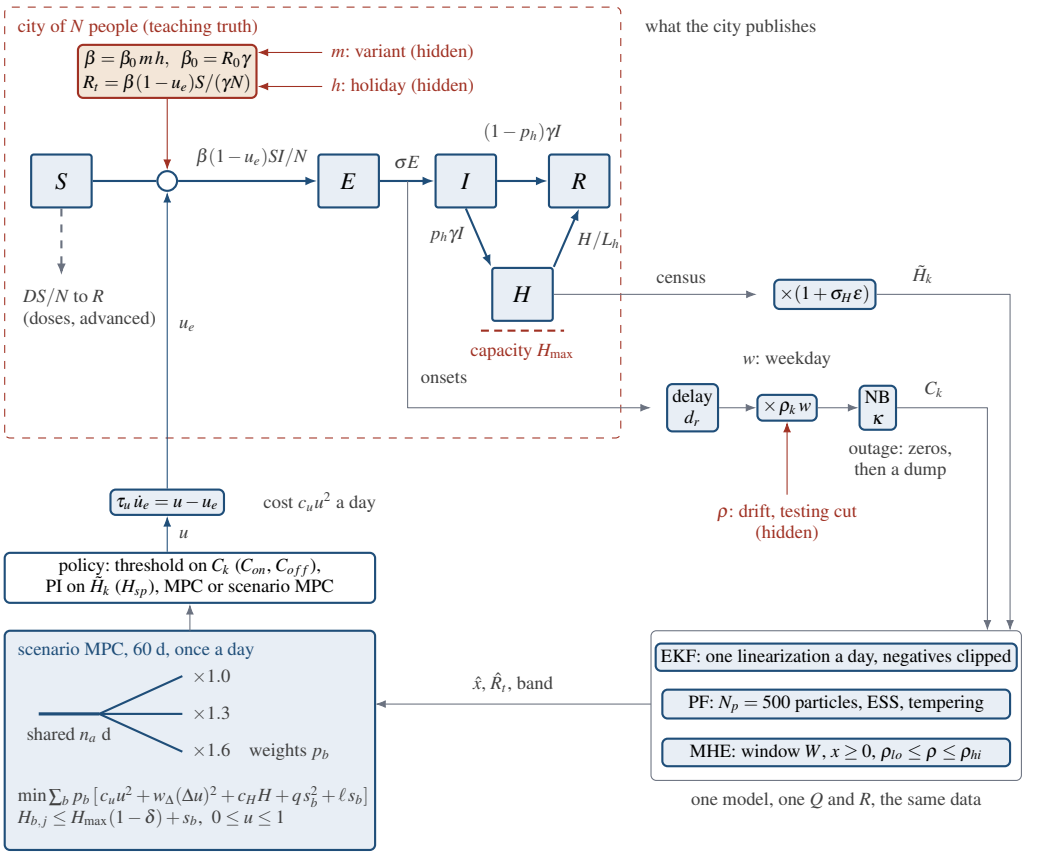


Figure 14.1 The teaching city as modeled, with what it publishes and who reads it. Inside the dashed frame (teaching truth) the compartments S, E, I, H and R exchange people at the rates shown, the transmission β carries the hidden variant and holiday multipliers m and h, and the census H stands against the capacity $H_{\max}$. The onsets σE reach the reports C_k after the delay d_r, the ascertainment ρ_k, the weekday factor w and negative-binomial noise of dispersion κ, and the census reaches $\tilde{H}_k$ with relative noise σ_H. Three estimators read only those two counts. A policy, the scenario-tree MPC among them, sets the dial u, whose effect u_e lags by τ_u. Not to scale.

where $1/\sigma$ is the latent period, $1/\gamma$ the infectious period, p_h the share of those leaving I who are admitted, L_h the stay, D the vaccine doses a day, u_e the dial's effect and c the cost ledger in millions of dollars. The doses reach whoever comes, so only the share S/N still susceptible gains immunity. The five compartments sum to N at every step. Transmission and the reproduction number are given by Equation (14.2).

$$\beta(t) = \beta_0\, m(t)\, h(t), \quad \beta_0 = R_0\gamma, \quad R_t = \frac{\beta(1-u_e)S}{\gamma N} \tag{14.2}$$

where m is the variant multiplier and h the holiday multiplier, both hidden. The dial removes a share of transmission once its effect has caught up, and its cost is quadratic, \$50 million a day at $u = 1$, so a moderate dial held long costs less than a hard one held briefly.

14.3.2 Reports, Ascertainment and the Census

The day's report counts a sample of the onsets σE that occurred during one day, d_r days earlier, as shown in Equation (14.3).

$$\mu_k = \rho_k \, w_{\mathrm{dow}(k)} \int_{k-d_r-1}^{k-d_r} \sigma E \, dt, \qquad C_k \sim \mathrm{NB}(\mu_k, \kappa) \tag{14.3}$$

where w is a declared weekday pattern (1.15, 1.10, 1.05, 1.00, 1.00, 0.85 and 0.85, mean 1) and the negative binomial has mean μ_k and variance $\mu_k + \mu_k^2/\kappa$. It is drawn as a Poisson count whose mean is μ_k/κ times a gamma deviate of shape κ, the Poisson by inversion below a mean of 30 and by a rounded normal above. The ascertainment drifts about a testing level $\bar{\rho}$ as a discrete Ornstein–Uhlenbeck deviation with a 30-day correlation, as shown in Equation (14.4).

$$\begin{aligned} \rho_k &= \mathrm{clip}(\bar{\rho} + \delta_k,\ 0.05,\ 0.95) \\ \delta_k &= a\,\delta_{k-1} + s_\rho \sqrt{1-a^2}\,\varepsilon_k, \qquad a = \mathrm{e}^{-1/30} \end{aligned} \tag{14.4}$$

where $s_\rho = 0.03$ and ε_k is a standard normal deviate. A testing cut multiplies $\bar{\rho}$ by $f_t = 0.5$ at once, clipped to 0.05 to 0.9. The census is counted once a day with $\sigma_H = 0.01$, as shown in Equation (14.5).

$$\tilde{H}_k = \max\left(0,\ \mathrm{round}[H(k)(1+\sigma_H \varepsilon'_k)]\right) \tag{14.5}$$

A reporting outage of d_{out} days publishes zeros and holds the day's reports, then adds the held reports to the next day's, the catch-up dump. Under zero noise the published report equals Equation (14.3) over the window $[k-8, k-7]$ to a relative 6.0×10^{-14}, and shifting the window a day either side misses by 14 to 16 %, which confirms the 7-day delay.

> **+ DID YOU KNOW? A Count Given as a Range**
>
> Even after an epidemic, its size is an estimate. Guidelines of the Centers for Disease Control and Prevention (CDC) put the people in the United States infected with the 2009 pandemic influenza virus between April 2009 and April 2010 at an estimated 43 to 89 million, and the deaths at about 12,000 [2].

14.3.3 Events, Starting State and Integration

Every hard event (the variant, a testing change, the outage and its dump, the holiday, a breach of the capacity) happens exactly once, at the boundary between two steps, and leaves a mark in the log, and the right-hand side of Equation (14.1) stays pure. The run starts with $E = 100$, $I = 40$, $S = N - 140$ and no one in hospital. The engine integrates in days with fourth-order Runge–Kutta at $\Delta t = 0.05$ d, runs the estimators and the policy once a day, and logs every 6 h. The onsets for Equation (14.3) are taken from a delay line of the cumulative onset ledger. Within one step the controller executes first, then the integration, the events, the outputs and the log, so the city publishes day k at $t = k$ and the controller uses it at the start of the next step. A log row at a whole day therefore shows that day's counts and the previous day's estimates, which an exported column (`est_day`) records. Every noise stream is seeded from the case, as Section 1.6.1 describes for every app. The parameters are listed in Table 14.2.

14.3.4 What the Model Leaves Out

The city has no age groups, households, regions or travel, and immunity does not wane. The latent and infectious stays are exponential, and real ones are distributions with a delay before infectiousness. Admission comes as people leave I and every patient stays 10 days, with no intensive care and no deaths. The reporting delay is fixed at 7 days, whereas a real one is a distribution that changes with testing capacity. The dial stands for every measure at once, with one lag and a quadratic cost. The estimators are told p_h, the census noise and the weekday pattern exactly, and that assumption carries

Table 14.2 Parameters, Defaults and Standing (A Adjustable in the Panel, H Hidden From the Estimators and Policies, F Not Adjustable in the Panel). All Are Teaching Values

Symbol	Meaning	Default (range)	
N	population	10^6	F
R_0	basic reproduction number	3 (1.2–6)	A, H
$1/\sigma$, $1/\gamma$	latent and infectious periods	4, 6 d	A
p_h, L_h	hospitalized share, stay	0.03, 10 d	A
τ_u	lag of the dial's effect	4 d (1–14)	A
d_r	onset-to-report delay	7 d (2–14)	A
ρ_0, s_ρ	ascertainment at start, drift sd	0.30, 0.03	A, H
κ, σ_H	report dispersion, census noise	20, 0.01	A
m_v, f_t	variant multiplier, testing cut	1.5, 0.5	A, H
h, d_{out}	holiday ($\times$, 7 d), outage	1.3, 3 d	A, H
c_u	cost of the full dial	\$50M/d	A
$H_{\max}$, δ	capacity, planning margin	1500, 5 %	A
C_{on}, C_{off}, u_{on}	threshold lines and dial	300, 100/d, 0.8	A
H_{sp}, K_c, τ_I	PI target, gain, reset	900, 0.002/bed, 20 d	A
W, ρ_{lo}, ρ_{hi}	MHE window and band	21 d, 0.1, 0.8	A
q_β, σ_ρ, q_E	random walks of β, ρ, E	0.03, 0.01, 0.03	A, F
N_p	particles	500	A
m_b, p_b, n_a	tree branches, weights, shared days	1, 1.3, 1.6, $\frac{1}{3}$, 7	A, F
w_Δ, c_H	move penalty, cost of illness	\$5M, 0	A
q, ℓ	slack prices	0.1, \$10M/bed	F

more of the estimation results than any other (Section 14.7). Each of these choices keeps the lesson on the delays and the hidden fraction, and each breaks down where a real epidemic's heterogeneity matters.

14.4 Growth, Final Size and Three Delays

With the dial at zero the model has two closed forms. Because $\mathrm{d}\ln S/\mathrm{d}t = -\beta I/N$ and $\mathrm{d}(S+E+I)/\mathrm{d}t = -\gamma I$, the quantity $s+e+i-\ln(s)/R_0$ is conserved in fractions of N, so $E+I$ peaks where $s = 1/R_0$ and the final size z solves $z = 1-\mathrm{e}^{-R_0 z}$ for a vanishing seed. At $R_0 = 3$ these give a peak $E+I$ of $0.300509N$ and a final size of 0.940490 from the 140 seeded, against 0.940480 for a vanishing seed. Early on, a growth rate r of the infections satisfies $(1+r/\sigma)(1+r/\gamma) = R_0$ for exponential latent and infectious stages, the relation the generation-interval argument gives for this model [11], so the infections double every 4.69 d at $R_0 = 3$ and every 3.43 d at $R_0 = 4$. Section 1.4.3 sets this growth beside the other growth laws of the volume.

Three delays separate the dial from the beds. The dial's effect builds over $\tau_u = 4\,\mathrm{d}$. A person infected today has onset after a mean latent period of 4 days and is reported 7 days after that, so the report describes infections about 11 days old. Admission follows the infectious period, a mean 10 days after infection, and the stay lasts 10 days more, so the census integrates the infections of the last three weeks. A policy that reads the reports acts on the epidemic of about two weeks before, and one that reads the census acts later still. The 7-day average adds about three days more. Section 1.4.5 treats sparse and late samples in general.

🕮 **Pause.** Try Exercise 14.1 before reading on.

14.5 Three Estimators on One Augmented State

All three estimators solve one problem with one model, one prior and one description of the noise, and they differ only in method. Once a day, at $t = k$, they receive day k's report and census and

estimate the augmented state of Equation (14.6).

$$\begin{aligned} x_k &= [S, E, I, H, \beta, \rho, z_1, \ldots, z_L]_k, \qquad L = d_r + 1 \\ x_k &= f_d(x_{k-1}, u, u_e, D) + w_k \\ y_k &= [\rho_k \, w_{\mathrm{dow}(k)} \, z_{L,k}, \; H_k] + v_k \end{aligned} \tag{14.6}$$

where f_d is one day of Equation (14.1) integrated with four Runge–Kutta steps of 0.25 d and the dial's lag in closed form, and z_j holds the onsets of the day j days back, so the report is a delayed sample of a state rather than an output with dead time. The navigation filter of Volume 5, Chapter 13 meets a late measurement the same way, applying each acoustic fix at the instant it describes. The dial and its effect are the controller's own and are known exactly. The process noise is a random walk on β with standard deviation $\sigma_\beta = q_\beta \beta_{ref}$ a day ($\beta_{ref} = 0.5$/d, the prior $R_0 = 3$ times γ), a random walk on ρ with $\sigma_\rho = 0.01$ a day, and a relative error $q_E = 0.03$ on E. The measurement variances are $\max(1, \mu + \mu^2/\kappa)$ for the report and $(\sigma_H H)^2 + 1$ for the census. The prior is $\beta = 0.50 \pm 0.125$/d, $\rho = 0.30 \pm 0.10$, $E = 100 \pm 100$ and $I = 50 \pm 50$ (one standard deviation). The estimators are told N, σ, γ, p_h, L_h, τ_u, d_r, κ and σ_H, and never R_0, ρ, m or h. Each reports $\hat{R}_t = \hat{\beta}(1 - u_e)\hat{S}/(\gamma N)$ with a band.

14.5.1 Extended Kalman Filter

The EKF predicts with one model day, linearized by forward differences of f_d about the estimate, as $P \leftarrow FPF^\top + Q$, then updates with both measurements in the Joseph form. Its only constraint is the common practical one: a component that comes out negative is set to zero. Its band on $\hat{R}_t$ is two standard deviations from P by the delta method.

14.5.2 Particle Filter With Tempered Updates

The PF carries $N_p = 500$ samples of the augmented state through the same model day and draws the same process noise for each [15]. Each is weighted by the negative-binomial likelihood of the report and a Gaussian likelihood of the census, and the cloud is resampled systematically when its effective sample size, $1/\sum w_i^2$, falls below $N_p/2$. A day whose likelihood alone would leave fewer than $N_p/10$ effective particles is taken in stages, each raising the likelihood to the largest power that keeps the effective size at $N_p/2$, with the cloud resampled and jittered by the stage's share of one day's process noise between stages [16]. The jitter makes this a regularized approximation rather than an exact importance-sampling step. The band is the weighted 5 to 95 % range.

14.5.3 Moving Horizon Estimation

The MHE solves, once a day, a window of the last $W = 21$ days, as shown in Equation (14.7).

$$\begin{aligned} \min_{\varphi, w} \; & \|\varphi\|^2 + \sum_j \left(w_{\beta,j}^2 + w_{\rho,j}^2 + w_{E,j}^2 \right) \\ & + \sum_j \left[\frac{(C_j - \mu_j)^2}{R_{C,j}} + \frac{(\tilde{H}_j - H_j)^2}{R_{H,j}} \right] \\ \text{s.t.} \; & x_{k_0} = \bar{x} + L_P \varphi \geq 0, \quad \rho_{lo} \leq \rho_j \leq \rho_{hi}, \\ & \beta_j \geq 0, \quad w_{E,j} \geq -0.9/q_E \end{aligned} \tag{14.7}$$

where the decision variables are the window's starting state, written through the Cholesky factor L_P of the arrival covariance, and the standardized process noises of every day, 77 variables for a full window with 99 linear constraint rows. The arrival cost is the filtering form [5], with $\bar{x}$ the MHE's own estimate made at the window's start and P carried forward by the extended Kalman recursion along its estimates. Each Gauss–Newton step linearizes the residuals by forward differences and solves $\min \frac{1}{2}\|r + J\Delta\|^2 + \frac{1}{2}\lambda\|\Delta\|^2$ with the constraints on the dense dual active-set quadratic program of the shared framework (`core/qp.js`, Section 1.6.3), with $\lambda = 10^{-6}$. A backtracking line search

on the true cost accepts the step, the variances $R_{C,j}$ are re-evaluated at the iterate, and the iteration stops when the cost falls by less than 10^{-6} of itself, the step falls below 10^{-4}, or after 12 iterations. Yesterday's solution, shifted a day, is the warm start. A window whose end state comes out non-finite or with a negative compartment is not used and the previous estimate is carried one model day instead, which happens only with several parameters at their extremes at once. The band is two standard deviations from the Gauss–Newton Hessian, ignoring the active constraints.

14.6 Dial Policies as Implemented

The policy in force sets the dial once a day and holds it until the next day, as a policy decision is held. In manual mode the learner's dial passes through, and the vaccine doses, an advanced second actuator, are given by hand in either mode up to a declared supply.

14.6.1 A Relay and a PI Loop

The threshold policy is a relay with hysteresis on the reported 7-day average $\bar{C}_7$: the dial goes to $u_{on} = 0.8$ when $\bar{C}_7$ passes $C_{on} = 300$ a day and to zero when it falls below $C_{off} = 100$, and it holds between them. The PI policy is the velocity-form controller of Section 1.6.2 with no derivative action, on the census count against a target $H_{sp} = 900$ beds. It is direct acting, $a = -1$, because a higher dial lowers the census. It has $K_c = 0.002$ of the dial per bed and $\tau_I = 20\,\mathrm{d}$, executes once a day, clamps its output to $0 \le u \le 1$ and unwinds its integral by back-calculation with a one-day tracking time T_t. Under any other policy it tracks the dial in force, so a transfer is bumpless.

🕮 **Pause.** Try Exercise 14.3 before reading on.

14.6.2 Economic MPC by Successive Linearization

The MPC plans the dial for 60 days from the selected estimate, never the truth, as shown in Equation (14.8).

$$\begin{aligned} \min_{u,s} \; & \sum_{b} p_b \Big[\sum_{j=0}^{59} \big(c_u u_{b,j}^2 + c_H H_{b,j} \big) + w_\Delta \sum (\Delta u_b)^2 \\ & \qquad + q\, s_b^2 + \ell\, s_b \Big] \\ \text{s.t.} \; & 0 \le u_{b,j} \le 1, \quad s_b \ge 0, \quad H_{b,j} \le H_{\max}(1-\delta) + s_b \\ & u_{b,j} = u_{b',j} \;\; (j < n_a) \end{aligned} \tag{14.8}$$

where b indexes the branches, p_b their weights, $H_{b,j}$ the predicted census on branch b at the end of day j, and s_b one slack per branch priced at $\ell = \$10$ million a bed, far above any day's cost. The dial moves daily for the first $n_a = 7$ days and in blocks of 7 days after, so the nominal controller has 15 moves and one slack. The prediction is the estimators' model day f_d run from $(\hat{S}, \hat{E}, \hat{I}, \hat{H})$ with $\hat{\beta}$ held, the dial's lag in closed form. Each linearization takes the sensitivities of every predicted census to every move by forward differences, solves the quadratic program on `core/qp.js` with at most 600 active-set iterations, replaces the plan and simulates it again. The plan is accepted when no move changes by more than 10^{-3}, or after 4 linearizations, a cap the status reports. Yesterday's plan, shifted a day, starts the next solve, and a failed solve holds yesterday's dial. The plan's last weeks ease the dial, because the census their infections would fill lies past the horizon. Only the first day is applied.

14.6.3 Branches That Share a Week

The scenario MPC solves Equation (14.8) with three branches whose transmission is $\hat{\beta}$ times $m_b = 1.0$, 1.3 and 1.6 from today, each with weight $p_b = \frac{1}{3}$, and the first 7 days of the dial shared by all three. That equality is the non-anticipativity constraint of multi-stage MPC [19]: the dial applied this week must keep even the 1.6 branch under the planning ceiling until the branches may differ. With 31 moves and three slacks the problem stays strictly convex, and the app's test finds 0.0667 as the least eigenvalue of its Hessian. The tree does not know when a variant will arrive. It hedges against $\times 1.3$

Table 14.3 The Ten Cases

#	Case, policy and events	Days	Teaching point
1	Manual, dial by hand	120	the two lags
2	Threshold 300/100 a day at 0.8	240	a relay on a lagged count
3	PI on the census, target 900	150	feedback that comes late
4	Manual, dial 0.66 from day 35	120	three estimators agree
5	As case 4, testing cut on day 70	120	the census anchors
6	MPC on the MHE	180	beds as a constraint
7	As case 6, variant on day 110	180	tree against a margin
8	As case 4, outage days 71–73, dump on day 74	100	where a bound matters
9	Challenge: threshold on the EKF, cut on day 60, variant on day 110, then a 28-day hand-over	180	the whole problem
10	$R_0 = 4$, threshold, no estimator	180	out of reach of a reactive rule

Table 14.4 Measured Outcomes of the Cases. Peak Is the True Census, and Cost Is the Dial's Cost Over the Case's Run

#	Configuration	Peak (beds)	Cost ($M)
1	dial 0	7201, day 73	0
	dial 0.7 on day 39 / 40 / 41	1435 / 1610 / 1799	—
2	threshold as shipped	932, day 55	4768
	reporting delay 3 / 14 d	715 / 2406	—
3	PI, target 900 / 300	1749 / 1051	—
6	MPC / scenario MPC	1421 / 1126	1788 / 2198
7	MPC, margin 5 %	1542, 33 d over	2440
	MPC, margin 10 %	1485	2599
	scenario MPC	1369	2855
10	threshold, no estimator	2098, day 51	—
	MPC on MHE / EKF / PF[a]	1436 / 1435 / 1431	2783 / 2783 / 2680

[a]Peak and cost through the 28-day hand-over with the dial held. Cases 4, 5 and 8 are estimation cases and are reported in the text.

and $\times 1.6$ every day, whether a variant is coming or not, and it branches at day 0 rather than day 7 because a tree branching a week ahead lets its recourse see the branch the moment it begins and keeps no margin. It peaks at 104.8 % of the capacity in case 7 and at 104.0 to 106.5 % on the challenge's seven records.

Three models are kept apart. The plant integrates Equation (14.1) at 0.05 d with the true β, the hidden multipliers and the reports' noise. The estimators carry the model day f_d, the random walks and the declared parameters. The MPC predicts with the same model day from an estimate and, on a branch, a multiplied $\hat{\beta}$. Over 60 days at the shipped parameters the model day reproduces the plant's census to better than 0.001 %, and with a transmission of 15 a day ($R_0 = 6$, a latent day, two infectious days, a variant of 2.5 and a holiday of 2) to within 0.97 %.

14.7 Measurements on Ten Cases

The ten cases appear in Table 14.3. Each runs on its own seed with the shipped settings, with any setting its row names typed in before Run, and the events at the days listed. All values reported below were produced by the app's engine, run without a browser but stepped exactly as the page steps it, and an estimate is compared with the truth of the day whose data it rests on, the alignment the app's own referee uses. Its record is the same whether the case is the first loaded, follows another case from the menu, or is rerun from Reset. The measured outcomes are collected in Table 14.4, and Figure 14.2 draws its peak column as bars.

14.7.1 The Count That Keeps Rising

In case 1 the dial set to 0.7 on day 39 turns the true incidence on that same day, but the reported 7-day average keeps rising for 20 days, to 1,794 a day on day 59, and the census peaks at 1,435 beds on day 61.75. The delays of Section 14.4 add up: the dial's lag, the onsets still to come from everyone

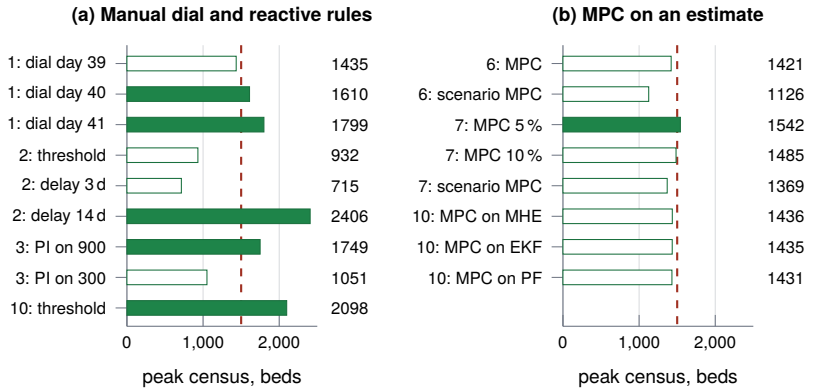


Figure 14.2 The peak census of Table 14.4, case by case (the number before each label), against the capacity of 1500 beds (dashed line). Solid bars pass the capacity, white bars stay under it. The uncontrolled peak of 7201 beds is off the scale. Notice that a day's delay on the dial, a 14-day reporting delay, the PI loop on its target of 900 and the threshold at $R_0 = 4$ carry the reactive peaks past the line, while every MPC on an estimate stays under it except the nominal MPC at its 5 % margin when the variant of case 7 arrives.

already exposed, the 7-day report delay and the 7-day average. The peak moves by about 182 beds for each day the dial waits, to 1,610 beds on day 40 and 1,799 on day 41, which stays over the capacity for 23 days, and the reported average still peaks on day 59 in every one of the three runs.

> **TRY IT IN THE APP** **1 · The Uncontrolled Wave and Its Two Lags**
>
> Load the case and press run with the Intervention dial at 0. Watch the Curves view: the reports trail the dashed true onsets by the 7-day reporting delay, and the census in the right plot climbs past the capacity line of 1,500 beds. Left alone, it peaks at 7,201 beds on day 73.

14.7.2 A Relay and a PI on Late Counts

The threshold of case 2 switches on at day 37, off at 104, on at 129, off at 181 and on at 210, periods of 92 and 81 days, with the census peaking at 932 beds on day 55 and a cost of $4,768 million over 240 days. Lengthening the reporting delay to 14 days stretches the first period to 117 days and lets the census reach 2,406 beds, and shortening it to 3 days gives periods of 80 and 69 days and a peak of 715. The period of a relay on a lagged signal follows the lag. The PI on the census in case 3 first moves the dial on day 37 and reaches the full dial on day 47, when 1,019 beds are already counted, and the census peaks at 1,749 beds on day 57, 1.94 times its target and over the capacity for 22.9 days. A target of 300 beds still peaks at 1,051, 3.5 times the target. Feedback on the census overshoots by the lag, which is built into the count.

14.7.3 Three Estimators Anchored by the Census

On the steady record of case 4, from day 30 on, every daily estimate of R_t is within 0.083 of the truth for the EKF, 0.061 for the PF and 0.082 for the MHE, with a mean error of 0.021, and each estimate of ρ is within 0.065. All three are within 0.2 on all 99 days from day 21, and over seven seeds the three bands contain the true R_t on every day.

Case 5 cuts testing in half on day 70. The reported 7-day average falls from 1,123 a day on day 69 to 490 on day 77 while the census falls only from 996 to 939. The MHE's bound on ρ might be expected to separate the estimators here, and it does not. All three read the cut as testing, with R_t within 0.076 (EKF), 0.097 (PF) and 0.073 (MHE) of the truth from the cut to day 120, and each $\hat{\rho}$ within 0.03 of the truth from day 92 (91 for the MHE). The census pins the number infected, and

in this model only ρ can jump overnight, so the explanation the data leave is a testing change. The bound acts only late and barely. On day 102 the EKF's $\hat{\rho}$ dips to 0.089 and the PF's to 0.086, below the 0.1 floor that stops the MHE's at 0.100, but the drifting true ρ is itself 0.103 that day. Over eight records the EKF crosses the floor on three, and on two of those the true ρ also falls below 0.1 (to 0.074 and 0.092), where the bound holds the MHE away from the truth. A bound is information, and here it is sometimes wrong.

Pause. Try Exercise 14.2 before reading on.

The census decides this result, and the result depends on what the estimators are told about it. With the census withheld from the estimators none finds the cut: the EKF's $\hat{\rho}$ stays above 0.25 and the PF's above 0.22 against a true ρ between 0.10 and 0.17, the MHE's wanders as low as 0.11 and back, the R_t errors grow to 0.22 (EKF), 0.32 (PF) and 0.24 (MHE), and the diverged-estimator alarm, judged on the census they no longer read, lights for all three. The bound cannot help there because the wrong answer lies inside the band. With p_h declared 20 or 30 percent low or high, every R_t stays within 0.089 while $\hat{\rho}$ scales with the error (0.74 to 1.39 times the truth for the EKF and the PF), and a census at 5 % noise with 1 % declared raises the R_t errors to 0.39 (EKF), 0.20 (PF) and 0.33 (MHE). Declared honestly, census noise of 3 to 10 % keeps every R_t error under 0.13.

TRY IT IN THE APP 5 · The Testing Cut: The Count Halves, the Census Does Not

Load the case and press run. Either side of day 70, watch the Reports 7-d avg and Census badges: the reported average falls by more than half within a week while the census hardly moves. Read the three estimates of R_t in the middle plot through the cut. Then set Census to the estimators to Off, press Reset and run again: none of the three finds the cut.

14.7.4 Where the Bound Matters

Case 8 stops the reports for three days and then publishes the held 2,505 with the day's own 1,130 as a dump of 3,635. A day of zero reports lies 4.42 standard deviations below its expected count, and the dump 15.98 above. The estimators' $\hat{\rho}$ through the outage is shown in Figure 14.4a. The EKF's falls to 0.142 through the zeros, jumps to 0.531 on the dump and is not back within 0.05 of the truth until day 91. The PF's falls to 0.152 and is back by day 74. On its own the dump would leave the PF 2.94 effective particles of 500, and the tempered update takes it in up to 6 stages, never with fewer than 119.8. The MHE's stops at its 0.1 floor on day 72, the second zero, and once the dump is in its window it reads the zeros and the dump together and is back by day 77. The same MHE with its band loosened to 0 and 5 falls to zero, where the reports no longer inform the onsets, and is not back until day 90. That loss of 13 of its 14 days of lead over the EKF is the bound's worth, and any floor from 0.05 to 0.12 keeps it, while a floor of 0.01 recovers only on day 87. With 50 or 200 Gauss–Newton iterations instead of 12 the bounded and loosened recoveries are unchanged, so the loosened MHE's slow return comes from the problem and not the solver.

What the bound protects is the estimate of ρ, which no policy reads. Through days 70 to 100 every $\hat{R}_t$ stays within 0.054 (EKF), 0.117 (PF), 0.050 (MHE) and 0.061 (loosened MHE) of the truth, and every estimate of $E + I$ between 0.936 and 1.073 times the truth, because the census carries the infections through the outage. The scene at day 80, with the EKF still at 0.47, is shown in Figure 14.3.

14.7.5 Planning Against the Beds and a Variant

The MPC on the MHE in case 6 lets the census rise to its planning ceiling and rides it while the dial eases with the immunity the epidemic builds. The census peaks at 1,421 beds on day 98, 94.7 % of the capacity, for $1,788 million over 180 days. The scenario MPC on the same case peaks at 1,126 beds for $2,198 million, the price of a hedge against a variant that does not come.

In case 7 a variant multiplies transmission by 1.5 on day 110. The nominal MPC holds the census at 1,392 beds that day, close to its 1,425-bed ceiling, and the variant takes it past the capacity on day 119.6 and to 1,542 beds, over the capacity for 33.05 days, for $2,440 million. The estimators' $\hat{\beta}$ is

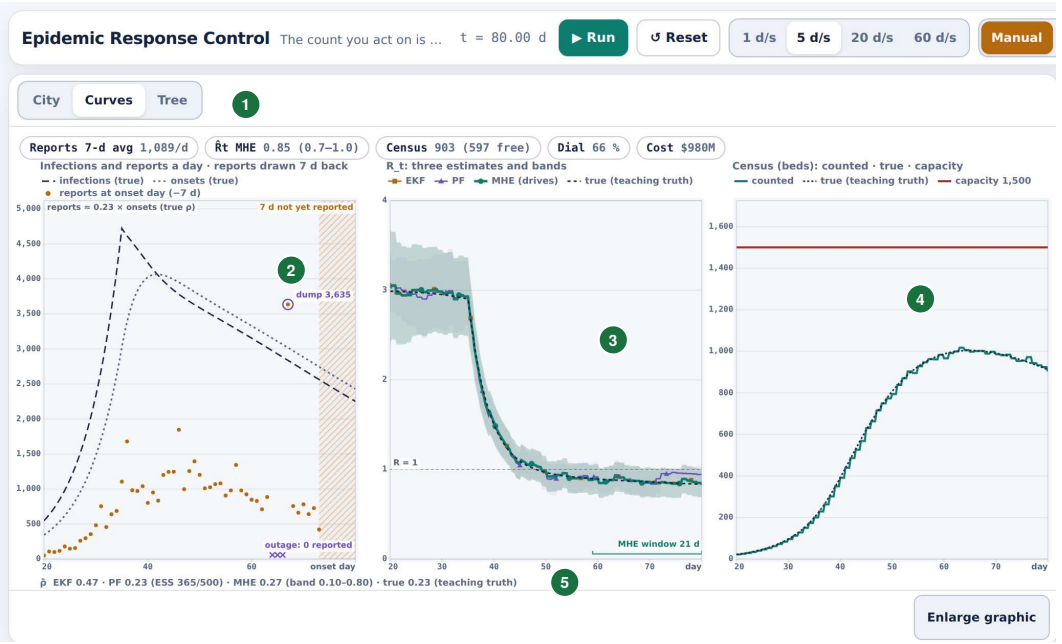


Figure 14.3 The reporting outage of case 8 at $t = 80.00$ d on the Curves view, in Manual with the dial at 66 %. The crop keeps the header and the plant view and cuts the side and lower panels. Callouts: (1) the badges, (2) the dump, (3) the estimates of R_t, (4) the census and the capacity, (5) the line of estimates of ρ. The left plot draws the true infections and onsets (dashed, teaching truth) and each day's report at the onset day it samples, with the three zeros as crosses on the axis, the dump of 3,635 ringed and the last 7 days hatched as not yet reported. The middle plot draws the three estimates of R_t with their bands against the truth, the MHE's 21-day window bracketed, and the right plot the census at 903 counted against the 1,500-bed capacity. The line under the plots shows the EKF's $\hat{\rho}$ still at 0.47 from its jump on the dump, the MHE's at 0.27 after its days on the floor and the PF's at 0.23, against a true 0.23, all made from day 79's data. The badges read a reported 7-day average of 1,089 a day, $\hat{R}_t$ 0.85 from the MHE against a true 0.84 on day 79, and a cost of $980 million.

a quarter higher only 10 days after the variant for the EKF and the MHE and 19 days after for the PF. The scenario MPC holds the census to 1,369 beds for $2,855 million because its shared week kept the census at 1,066 beds on day 110, a margin the variant used before any estimator saw it. A hand-set margin does the same. The nominal MPC with a 10 % planning margin holds at 1,485 beds for $2,599 million. What the tree adds is where its margin comes from, a declared spread of futures and the week before recourse, rather than a number chosen by hand, and at $\times 1.6$ it is the dearer insurance. The tree ends at its cap of 4 linearizations on 178 of 180 days of case 7. Re-solved from the same warm start to 30 linearizations, its first move differs by at most 5.07 points of the dial (median 0.27), and the nominal MPC's by at most 0.33.

14.7.6 The Graded 180 Days

The challenge cuts testing on day 60 and brings a variant on day 110. Its target is a true census never above 1,500 beds, a cost within $3,600 million and the selected estimator's R_t within 0.2 of the truth on at least 90 % of the days from day 21, all judged through a 28-day hand-over in which the city runs on with the dial held where it stands on day 180. The hand-over catches a dial dropped near the end: the suggested run with its dial set to zero on day 170 passes the 180 days alone at 1,375 beds

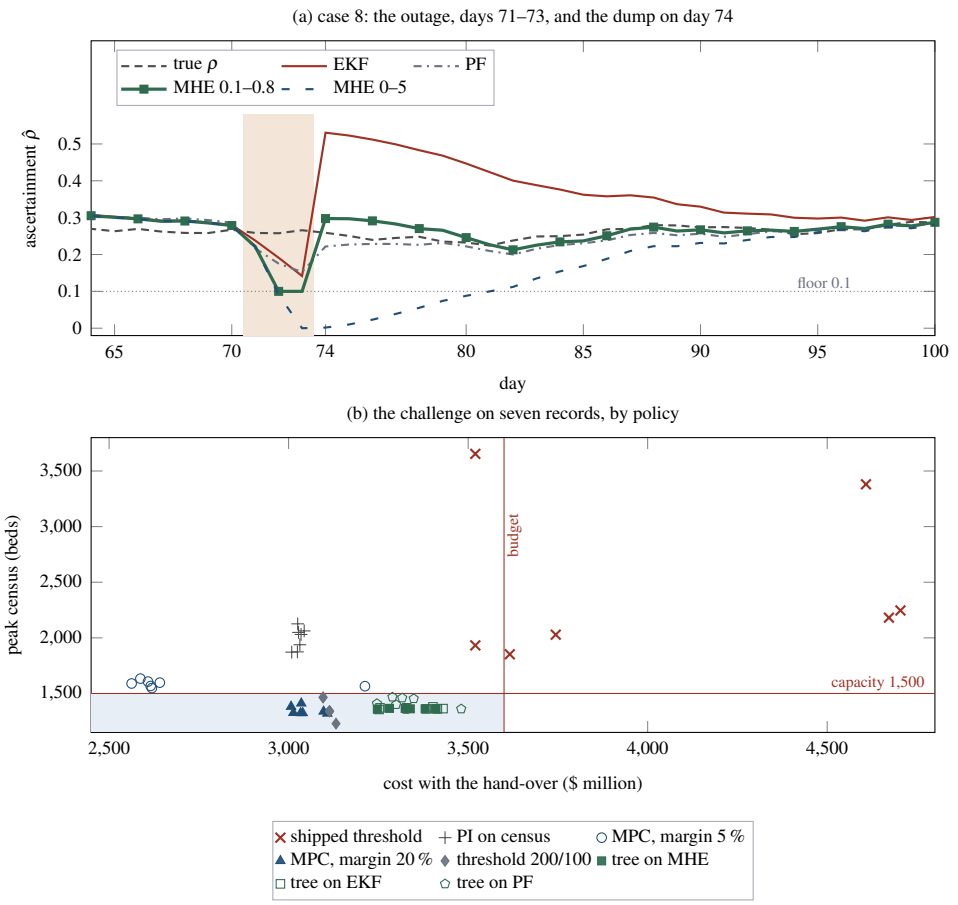


Figure 14.4 The bound and the margin. (a) Each estimator's $\hat{\rho}$ made from a day's data against that day's true ρ (densely dashed) in case 8, with the outage shaded: the EKF solid, the PF dash-dotted, the MHE with its shipped band of 0.1 to 0.8 solid with squares, and the MHE with its band loosened to 0 to 5 loosely dashed. The dotted line is the floor. (b) Peak census against cost on the challenge's seven records for each policy, one marker shape per policy, both through the 28-day hand-over, with the passing region shaded.

Table 14.5 The Challenge on Seven Paired Records (Seeds 109 and 1–6). Peak and Cost Include the Hand-Over

Policy	Pass	Peak (beds)	Cost ($M)
threshold on the EKF (shipped)	0/7	1852–3654	3520–4704
PI on the census	0/7	1872–2126	3009–3043
MPC on the MHE, 5 % margin	0/7	1549–1634	2563–3213
scenario MPC on the MHE (suggested)	7/7	1359–1366	3247–3411
scenario MPC on the EKF	7/7	1356–1380	3252–3430
scenario MPC on the PF	7/7	1357–1468	3246–3481
MPC on the MHE, 20 % margin	7/7	1320–1412	3007–3108
threshold 200/100 a day at 0.6	7/7	1229–1465	3096–3132

and $2,684 million and reaches 4,432 beds in the hand-over. The policies on seven paired records are compared in Table 14.5 and Figure 14.4b.

As shipped, the threshold switches off on day 85, when the halved count falls under 100 a day, and on again only on day 109. By day 110, 29,312 people are infectious, the variant arrives that

day and the census reaches 2,029 beds on day 128, for $3,744 million. The scenario MPC on the MHE holds the census to 1,362 beds and spends $2,868 million plus $471 million in the hand-over, $3,339 million in all, with R_t within 0.2 on 95 % of days. On the shipped EKF the same tree holds 1,364 beds for $3,430 million. The estimator does not decide this challenge. The tree passes on every estimator and record, its runs meet the R_t test on 93 to 96 % of days, and the nominal MPC at its 5 % margin breaches on every record. The margin decides it. The nominal MPC with a 20 % margin passes all seven records for $3,007 to $3,108 million, less than the tree on every record, and the threshold retuned to 200 and 100 reports a day at 0.6 passes all seven too. The R_t test is not idle: with the random walk on β raised to 20 % of its prior a day, the suggested tree holds the beds and the budget and fails on R_t alone, at 84 % of days. The held dial of the hand-over and the policy left running for 28 days give the same verdicts for the shipped and the suggested runs.

> ✦ DID YOU KNOW? **A Plain Threshold, Retuned**
>
> From the simulator: retuned to switch on at 200 reported cases a day and off at 100, with the dial at 0.6, the plain threshold passes all seven records of the challenge and peaks at 1229 to 1465 beds. The nominal MPC at its shipped 5 % margin passes none (Table 14.5).

14.7.7 A Transmissibility No Reactive Rule Holds

At $R_0 = 4$ in case 10 the infections double every 3.4 days, and the shipped threshold without an estimator lets the census reach 2,098 beds on day 51. A sweep in the app's own tests runs 350 threshold settings, 294 PI settings and 1,313 constant dials (101 levels from 13 start days), each judged on the challenge's beds and budget through the hand-over. None holds the beds within $3,600 million, and the cheapest that holds the beds costs $4,077 million, while at $R_0 = 3$ the same sweep finds 126 that pass. A policy that reads only the reports or the census either acts late and overflows or acts early and stays on too long. With an estimate of the present the MPC holds the beds on every estimator (1,436 beds and $2,783 million on the MHE). The claim is scoped. Of 17,732 two-step hand schedules, 6 pass, but each was found by rerunning this same epidemic, and at $R_0 = 4.5$ none of the six holds.

14.7.8 Checks, Published Relations and the Limits of the Evidence

The first question, code verification, is whether the engine solves the equations above. The compartments sum to N within 7.9×10^{-15} of N over 400 days. The uncontrolled run reproduces the conserved quantity's peak $E + I$ to 0.0002 % and the final size to all six digits printed, and the peak census of 7,201.1 beds on day 73.10 agrees with an independent fourth-order integration at 0.002 d to 0.000 04 %. Halving Δt moves the peak census by 0.0000 % in cases 1 and 6, by −0.023 % in case 7 under the tree and by −0.016 % in the suggested challenge (1,362.0 to 1,361.8 beds), with every verdict unchanged. Over 300,000 fresh draws the negative binomial's mean is within 0.08 % and its variance within 0.45 % of the declared values at means of 50, 500 and 2,000. Over seven seeds the bands contain the true R_t on 93.9 % (EKF), 88.7 % (PF, a 5 to 95 % band) and 93.6 % (MHE) of the challenge's days. Averaged over four filter seeds, the 500-particle filter's $\hat{\rho}$ stays within 0.015 of a 20,000-particle filter's from the dump to day 85. On a 6-day window none of 200 random feasible neighbors of the MHE's solution has a lower cost, and on a 21-day one-branch instance the MPC's plan, at $122.6 million, costs less than the best of 2,601 two-block plans on a grid that hold the ceiling, $124.7 million. All 50 of the app's checks pass, 32 written for this city and 18 shared with the other apps. In a desktop browser at 60 days a second, the suggested challenge's 180 days take 3.06 s, the day's three estimators and tree a median of 5.3 ms (the MHE 2.1 ms, the tree 2.5 ms), and frames hold at 16.7 ms.

Agreement with published relations is narrower. The final-size and peak relations are those of the compartmental literature [10, 12, 13], and the early growth rate of the uncontrolled city, fitted over days 8 to 18, is within 0.23 % of the generation-interval relation at $R_0 = 3$ and within 0.47 % at

$R_0 = 4$, the residual being the 0.3 % and 0.7 % of the city already infected by day 18. No parameter was fitted to any outbreak, so the model is verified but not validated. The app's effect on learning has not been evaluated.

14.8 The City on the Estimation Ladder

On a ladder of estimation problems the city stands one rung above fitting a model to a record and one below multirate estimation, such as the kraft digester of Volume 5, Chapter 4, whose analyzer and laboratory report on different clocks. Its estimators and planners build on the options of Section 1.6.3, and its plant is one that a student of public health already knows. A simulation can draw beside every count the truth that no surveillance system has [20], here the infections, the onsets, the ascertainment and the true R_t. Three cases carry most of the lesson. Case 1 puts a dial and a count that keeps rising for 20 days on the same screen. Case 5 turns the census on and off and shows what an anchor is worth, and case 8 shows the one surprise where a bound earns its place. A testing cut that halves the ascertainment overnight, or three days of zeros followed by an exact dump, would not happen so cleanly in a real city, but each isolates one mechanism. At its fastest speed a 180-day case plays in about 3 s, and the export carries every estimate beside the truth, each true value marked as such by its column's name so that no replay mistakes it for a measurement.

As in every app of the series (Section 1.6.4), a case counts as done only when the city responds to something the learner did, and an event the case scripts is never taken for the learner's own action. Case 5, for example, asks for a run past day 90 with the census in use, then the census switched off, a Reset and a second run past day 90, and case 7 asks for a breach under the nominal MPC, then the tree, a Reset and a run to day 180 within the capacity. Reset keeps the learner's settings (Section 1.6.1), and a change made before Run sets the initial state instead of acting as a step. The challenge is graded by the same function the scorecard shows, and changing a model parameter voids the attempt.

14.9 Technology in Practice

The city of Section 14.3 publishes two counts and is moved by one dial. A real public health system has more streams, a set of measures in place of the dial, and people where the policy block sits. Figure 14.5 draws the loop that way.

What Measures It

During the coronavirus disease 2019 (COVID-19) pandemic, surveillance data on the number of cases, hospitalizations and deaths were used to inform response policies [21]. Planning guidelines name further indicators that a state or locality might watch, among them visits to health care providers for influenza-like illness, the geographic spread and school absenteeism [2]. Each stream has its own delay and sees its own share of the infections. The chapter keeps two: the reports stand for the case stream, with a fixed 7-day delay and a hidden ascertainment near 0.30, and the census for the hospital stream.

> **TECHNOLOGY** **Sensor: Case Reports and the Hospital Census**
>
> Surveillance data on cases, hospitalizations and deaths inform response policies, but they give only a partial, time-lagged picture of transmission [21]. A case is counted after an infection, an onset, a test and a report, so the series of new infections has to be rebuilt from observations made long after transmission. The share of infections found moves with test availability, testing criteria and interest in testing. Hospitalizations are less sensitive to testing effort and more sensitive to other factors, such as the beds available [22]. A change in testing therefore shows on a trend as a step in the reports that the census does not share, the signature of case 5. In the chapter the reports sample the onsets of 7 days before at a drifting share, and the census is counted daily within 1 %.

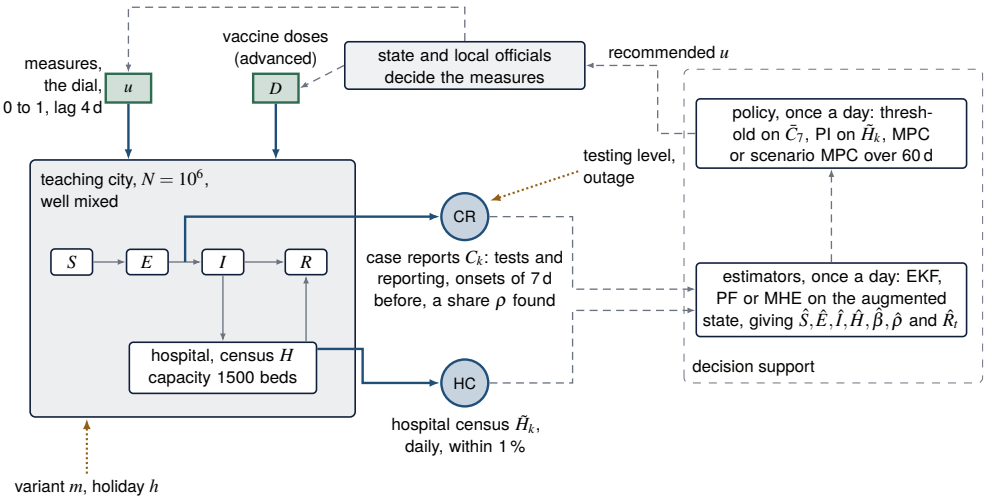


Figure 14.5 The loop of the chapter as a public health system runs it. The city publishes two surveillance streams, case reports (CR) a week late and only a share of them, and the hospital census (HC), counted daily. Estimators and a policy run once a day in a decision-support layer (dashed frame), and people, not the layer, decide the measures. Notice that the testing level and an outage act on the case reports alone, while the census reaches the estimators untouched by either.

What Moves It

The measures are phased in, or layered, on the basis of the pandemic's severity and the local transmission over time, and the more disruptive community measures are decided by state and local officials on the conditions of their jurisdictions. Their timely and effective use depends on community acceptance and active participation [2], which no actuator model carries. The dial of Equation (14.1) compresses all of this into one number between 0 and 1, with the lag τ_u and a quadratic cost.

> **TECHNOLOGY** **Final Element: Community Mitigation Measures**
>
> The handle a city moves is a set of nonpharmaceutical interventions. The CDC guidelines of 2017 group them as personal protective measures, such as home isolation of the ill and hand hygiene, community measures that increase social distancing, such as closing or dismissing schools, distancing in workplaces and postponing or cancelling mass gatherings, and environmental measures, such as routine cleaning of surfaces. Each is only partially effective, so plans call for several at once. Early in a pandemic they may be the only prevention tools at hand, because a pandemic vaccine can take up to six months to produce [2]. The chapter's dial stands for every measure at once, with one lag of four days and a quadratic cost.

Where the Controller Runs

No programmable controller closes this loop. The estimators and planners of this chapter belong to a decision-support layer under the officials who decide. Forecasting models already serve that layer: their predictions informed short-term decisions such as health care staffing, school closures and the allocation of medical supplies [21]. A model-based study of the same problem estimated the hidden states every day but revised its policy only every 25 days, because planned suppression strategies likely cannot be altered quickly [6]. The chapter's policy executes once a day, the period at which its counts arrive. Whether, and how far, any such rule should steer real measures is a policy question that the simulator does not answer.

What Is Modernizing It

The first development is wastewater surveillance. It measures community infection trends independently of health care seeking, access to care and testing capacity, and it has shown trends before case reports and hospital admissions. One statewide network sampled 65 treatment plants twice a week, analyzed the samples within 3 to 4 days and posted results daily. In another state, case rates fell while the number of people tested also fell, and steadily falling levels in wastewater showed that the decline was real [23], the question of case 5 answered by a third stream. The second is collaborative forecasting, in which a hub combined the predictions of many groups into an ensemble [21]. The third is model-based planning with state estimation and periodic revision of the policy [6], the structure this chapter teaches. None of them removes the delay between an infection and its count. Wastewater shortens it, and the other two plan around it.

> **IN PRACTICE** **Triggers Written Before the Pandemic**
>
> Community measures are planned before a pandemic arrives. The guidelines ask planners to identify activation triggers so that measures start before explosive growth, to plan for long school dismissals, and to weigh the social and economic costs of the measures against their public health benefits. Possible indicators include more visits to health care providers for influenza-like illness and a wider geographic spread, and the choice differs among states and localities with their resources. Rising influenza-associated hospitalizations are among the indicators that might confirm that measures should continue [2]. The threshold policy of case 2 is a trigger of this kind, written as a rule on the reported count.

14.10 Limitations

The city is a declared teaching model, well mixed, without age structure, waning immunity or deaths, with exponential stays and a fixed reporting delay, so it says nothing about where an epidemic spreads or who is admitted. The estimation results rest on the estimators knowing p_h, the census noise and the weekday pattern exactly, and Section 14.7 shows how a wrong p_h scales the ascertainment estimates and how an undeclared noisy census spoils R_t. The particle filter's tempering is a regularized approximation, checked against a 20,000-particle filter on one case only. The tree's protection is a margin. A nominal MPC with a hand-set margin passes the challenge for less, and the tree is preferred here only because its margin follows from a declared spread of futures. The MPC stops at 4 linearizations on most days of the tree, with a first move within about 5 points of a converged plan, its plan eases at the end of its horizon, and a policy switched on in mid run starts from a cold plan that can be far off for its first day. One recorded epidemic per case is graded, and seven records are checked for the challenge. The engine logs each whole-day row before the controller's execution at that instant, so an exported estimate must be read against its `est_day`. A simulated pass is not a public-health recommendation.

> **OPEN CHALLENGE** **Estimating Transmission in Real Time**
>
> Policy makers and public health officials use the effective reproduction number to judge whether interventions work. Two challenges are common to every method of estimating it: specifying the generation interval, and rebuilding the series of new infections from observations made long after transmission. Naive corrections for the delay, such as subtracting delays drawn from a distribution, can bias the estimate [22]. Methods that assume cases, hospitalizations and deaths all arise from one infection process might infer the infections more accurately while estimating the delays and the changes in ascertainment of each source [22]. Progress would be such joint estimates, tested where the truth is known, as it is in this chapter's city.

14.11 Conclusion

A city that acts on week-old, partial counts is an estimation problem before it is a control problem, and this chapter puts a declared compartmental city, its reports and its census, three estimators and four policies on one screen with the truth beside every count. A dial turned the day after the reports

pass 300 a day still sees them rise for 20 days, and feedback on the census peaks at 1.94 times its target. With the census to anchor them the EKF, the PF and the MHE all read a testing cut as testing, within 0.10 in R_t, and the MHE's bound earns its place only on a reporting outage, where it brings the ascertainment estimate back on day 77 against the EKF's 91 while every R_t stays within 0.12. A scenario tree holds a variant that breaches the nominal MPC, and so does a 10 or 20 percent margin, for less. At $R_0 = 4$ no threshold, PI or constant dial in a sweep of 1,957 holds the beds within the budget, while the MPC on any estimate does. The model is well mixed and every value is a teaching value. A fuller model would add age groups and waning immunity, delays as distributions, a census whose noise the estimators must learn, and a tree whose branches are revealed when the data can tell them apart. The simulator shows that what a constrained estimator or a scenario tree is worth depends on what else the measurements already pin down.

KEY POINTS

- A city that acts on its case reports acts on the epidemic of about two weeks before, and a dial turned the day after the reports pass 300 a day still sees them rise for 20 days.
- Feedback on the census peaks at 1.94 times its target, because the lag is built into the count.
- With the census to anchor them, all three estimators read a testing cut as testing. The MHE's bound earns its place only on a reporting outage.
- A scenario tree holds a variant that breaches the nominal MPC, and so does a 10 or 20 percent margin, for less.

Where else this applies. Volume 5, Chapter 4 estimates a plant from a late laboratory result, and Volume 9, Chapter 7 plans on a scenario tree of inflow forecasts.

References

[1] R. J. Hatchett, C. E. Mecher, and M. Lipsitch, "Public health interventions and epidemic intensity during the 1918 influenza pandemic," *Proceedings of the National Academy of Sciences*, vol. 104, no. 18, pp. 7582–7587, 2007.

[2] N. Qualls, A. Levitt, N. Kanade, N. Wright-Jegede, S. Dopson, M. Biggerstaff, C. Reed, and A. Uzicanin, "Community mitigation guidelines to prevent pandemic influenza—United States, 2017," *MMWR Recommendations and Reports*, vol. 66, no. 1, pp. 1–34, 2017.

[3] J. F. Lawless, "Adjustments for reporting delays and the prediction of occurred but not reported events," *Canadian Journal of Statistics*, vol. 22, no. 1, pp. 15–31, 1994.

[4] E. L. Haseltine and J. B. Rawlings, "Critical evaluation of extended Kalman filtering and moving-horizon estimation," *Industrial & Engineering Chemistry Research*, vol. 44, no. 8, pp. 2451–2460, 2005.

[5] C. V. Rao, J. B. Rawlings, and D. Q. Mayne, "Constrained state estimation for nonlinear discrete-time systems: Stability and moving horizon approximations," *IEEE Transactions on Automatic Control*, vol. 48, no. 2, pp. 246–258, 2003.

[6] C. Tsay, F. Lejarza, M. A. Stadtherr, and M. Baldea, "Modeling, state estimation, and optimal control for the US COVID-19 outbreak," *Scientific Reports*, vol. 10, no. 1, p. 10711, 2020.

[7] J. Köhler, L. Schwenkel, A. Koch, J. Berberich, P. Pauli, and F. Allgöwer, "Robust and optimal predictive control of the COVID-19 outbreak," *Annual Reviews in Control*, vol. 51, pp. 525–539, 2021.

[8] M. M. Morato, S. B. Bastos, D. O. Cajueiro, and J. E. Normey-Rico, "An optimal predictive control strategy for COVID-19 (SARS-CoV-2) social distancing policies in Brazil," *Annual Reviews in Control*, vol. 50, pp. 417–431, 2020.

[9] A. Cori, N. M. Ferguson, C. Fraser, and S. Cauchemez, "A new framework and software to estimate time-varying reproduction numbers during epidemics," *American Journal of Epidemiology*, vol. 178, no. 9, pp. 1505–1512, 2013.

[10] H. W. Hethcote, "The mathematics of infectious diseases," *SIAM Review*, vol. 42, no. 4, pp. 599–653, 2000.

[11] J. Wallinga and M. Lipsitch, "How generation intervals shape the relationship between growth rates and reproductive numbers," *Proceedings of the Royal Society B: Biological Sciences*, vol. 274, no. 1609, pp. 599–604, 2007.

[12] W. O. Kermack and A. G. McKendrick, "A contribution to the mathematical theory of epidemics," *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, vol. 115, no. 772, pp. 700–721, 1927.

[13] J. Ma and D. J. D. Earn, "Generality of the final size formula for an epidemic of a newly invading infectious disease," *Bulletin of Mathematical Biology*, vol. 68, no. 3, pp. 679–702, 2006.

[14] K. M. Gamado, G. Streftaris, and S. Zachary, "Modelling under-reporting in epidemics," *Journal of Mathematical Biology*, vol. 69, no. 3, pp. 737–765, 2014.

[15] N. J. Gordon, D. J. Salmond, and A. F. M. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," *IEE Proceedings F Radar and Signal Processing*, vol. 140, no. 2, p. 107, 1993.

[16] N. Oudjane and C. Musso, "Progressive correction for regularized particle filters," in *Proceedings of the Third International Conference on Information Fusion*, vol. 2. IEEE, 2000, pp. THB2/10–THB2/17.

[17] E. L. Ionides, C. Bretó, and A. A. King, "Inference for nonlinear dynamical systems," *Proceedings of the National Academy of Sciences*, vol. 103, no. 49, pp. 18 438–18 443, 2006.

[18] M. Ellis, H. Durand, and P. D. Christofides, "A tutorial review of economic model predictive control methods," *Journal of Process Control*, vol. 24, no. 8, pp. 1156–1178, 2014.

[19] S. Lucia, T. Finkler, and S. Engell, "Multi-stage nonlinear model predictive control applied to a semi-batch polymerization reactor under uncertainty," *Journal of Process Control*, vol. 23, no. 9, pp. 1306–1319, 2013.

[20] J. Park, R. A. Martin, J. D. Kelly, and J. D. Hedengren, "Benchmark temperature microcontroller for process dynamics and control," *Computers & Chemical Engineering*, vol. 135, p. 106736, 2020.

[21] E. Y. Cramer, E. L. Ray, V. K. Lopez, J. Bracher, A. Brennen *et al.*, "Evaluation of individual and ensemble probabilistic forecasts of COVID-19 mortality in the United States," *Proceedings of the National Academy of Sciences*, vol. 119, no. 15, p. e2113561119, 2022.

[22] K. M. Gostic, L. McGough, E. B. Baskerville, S. Abbott, K. Joshi, C. Tedijanto, R. Kahn, R. Niehus, J. A. Hay, P. M. De Salazar *et al.*, "Practical considerations for measuring the effective reproductive number, R_t," *PLOS Computational Biology*, vol. 16, no. 12, p. e1008409, 2020.

[23] A. E. Kirby, M. S. Walters, W. C. Jennings, R. Fugitt, N. LaCross, M. Mattioli, Z. A. Marsh, V. A. Roberts, J. W. Mercante, J. Yoder, and V. R. Hill, "Using wastewater surveillance data to support the COVID-19 response—United States, 2020–2021," *MMWR Morbidity and Mortality Weekly Report*, vol. 70, no. 36, pp. 1242–1244, 2021.

14.12 Exercises

14.1 ○ Concept check · 2 min

In case 1 the Intervention dial is set to 70 % on day 39, the day after the reported 7-day average passes 300 a day, and the true infections turn that same day. Why does the reported 7-day average keep rising until day 59?

(a) The dial's effect builds over four days, and the infections keep rising until it is complete
(b) The reports sample onsets a week old, from people infected before the dial moved
(c) The ascertainment drifts upward as testing catches up with a larger epidemic
(d) The reports count admissions, which follow the infections by about ten days

14.2 ○ Concept check · 2 min

In case 5 testing is cut in half on day 70, and the reported 7-day average falls from 1,123 to 490 a day within a week. With the census in use, all three estimators read the change as testing and not as a turn in the epidemic. What lets them do so?

(a) The MHE's floor of 0.1 on ρ, which the other two estimators share through the prior
(b) The weekday pattern of the reports, which the estimators know and the cut disturbs
(c) The census, which pins the number infected, so only ρ is left to explain the drop
(d) The 21-day window, long enough to hold the reports from before and after the cut

14.3 ○ Concept check · 1 min

The PI policy of case 3 moves the dial on the census count against a target of 900 beds. In the convention of Section 1.6.2, what is its action, and why?

(a) Reverse acting, $a = +1$, because the census rises when the dial is raised
(b) Reverse acting, $a = +1$, because the error is the target minus the census
(c) Direct acting, $a = -1$, because the dial is bounded below at zero
(d) Direct acting, $a = -1$, because a higher dial lowers the census

14.4 Theory and calculation · 5 min

The city has a latent period $1/\sigma = 4\,\mathrm{d}$ and an infectious period $1/\gamma = 6\,\mathrm{d}$. Early in the epidemic the infections grow at a rate r that satisfies $(1 + r/\sigma)(1 + r/\gamma) = R_0$.

(a) For $R_0 = 3$, find the transmission β_0, the growth rate and the doubling time. *Checkpoint:* $0.5\,\mathrm{d}^{-1}$, $0.1477\,\mathrm{d}^{-1}$ and 4.69 d.

(b) Repeat for $R_0 = 4$, the transmissibility of case 10. *Checkpoint:* $0.667\,\mathrm{d}^{-1}$, $0.2020\,\mathrm{d}^{-1}$ and 3.43 d.

(c) A student takes $r = \beta_0 - \gamma$, the rate of a model with no latent class. Find the doubling time it predicts at $R_0 = 3$, and say why it is wrong for this city. *Checkpoint:* 2.08 d.

14.5 Theory and calculation · 6 min

With the dial at zero, the quantity $s + e + i - \ln(s)/R_0$ is conserved, where s, e and i are the fractions of $N = 10^6$ in S, E and I. Take $R_0 = 3$ and a vanishing seed, so that $s = 1$ and $e = i = 0$ at the start.

(a) Find the largest number of people in E and I together. *Checkpoint:* About 300,500 people, $0.3005N$.

(b) Find the final size z from $z = 1 - \mathrm{e}^{-R_0 z}$ by fixed-point iteration from $z = 0.9$, to four digits. *Checkpoint:* 0.9405.

(c) Find the share still susceptible at the end, compare it with $1/R_0$, and explain the difference. *Checkpoint:* 0.0595 against 0.333.

14.6 Theory and calculation · 5 min

The daily report is negative binomial with mean μ and variance $\mu + \mu^2/\kappa$, with $\kappa = 20$.

(a) For $\mu = 1000$, find the standard deviation and the number of standard deviations by which a report of zero lies below the mean. *Checkpoint:* 226 and 4.43.

(b) Show that a report of zero can never lie more than $\sqrt{\kappa}$ standard deviations below its mean, and evaluate the bound. *Checkpoint:* 4.47.

(c) Case 8 finds its three zeros 4.42 standard deviations low and the dump 15.98 high. What does the bound mean for an outage alarm that flags any report more than 5 standard deviations from its mean?

14.7 App challenge · 8 min

Load scenario 1, The uncontrolled wave and its two lags. Use the default seed and set the speed to 1 d/s.

(a) Run until the Reports 7-d avg badge first shows 300 a day or more, and pause. Read the day. *Checkpoint:* Day 38.

(b) Before day 39 is reached, set the Intervention dial to 70 % and run to day 120. Read the Peak census tile, and find in the exported file the largest value of `reported_7d` and its day. *Checkpoint:* 1,435 beds on day 62; 1,794 a day on day 59.

(c) Press the reload button, run past day 39, pause before day 40, set the dial to 70 % and run to day 120. Read the Peak census tile. *Checkpoint:* 1,610 beds on day 62, 15 days over the capacity.

Accept a peak within 2 beds. State what one day of waiting costs and why the reports peak on the same day in both runs.

14.8 App challenge · 8 min

Load scenario 3, PI on the census: too late by construction. Use the default seed and run the case to its end at day 150, then export the file.

(a) Find the first row with `dial_pct` above zero, and the first row in which it reaches 100. Read `census_beds` in that second row. *Checkpoint:* Day 37; day 47 with 1,019 beds.

(b) Read the Peak census tile, and divide the peak by the target. *Checkpoint:* 1,749 beds on day 57, 23 days over; 1.94.

(c) Set the Census target (PI) to 300 beds, press Reset and run to day 150. Read the Peak census tile and divide by the new target. *Checkpoint:* 1,051 beds; 3.50.

Accept a peak within 2 beds. Explain why the lower target does not cure the overshoot.

14.9 App challenge · 9 min

Load scenario 7, The variant: nominal MPC breaches, the tree pre-empts. Use the default seed and run each case for 180 days. Read the Peak census and Intervention cost tiles each time.

(a) Run the case as shipped, with the nominal MPC and a 5 % planning margin. *Checkpoint:* 1,542 beds, 33 days over; $2,440 million.

(b) Set the Policy to Scenario MPC, press Reset and run again. *Checkpoint:* 1,369 beds; $2,855 million.

(c) Set the Policy back to MPC and the Planning margin to 10 %, press Reset and run again. *Checkpoint:* 1,485 beds; $2,599 million.

Accept a peak within 2 beds and a cost within $2 million. Which protection is cheaper, and what does the tree give that the margin does not?

14.10 Application area · 3 min

The estimators are told the hospitalized share $p_h = 0.03$ exactly. Suppose the true share is 0.03 but the estimators are told 0.021, 30 % low.

(a) By what factor does the number infected that the census implies change, and what does that do to the estimate $\hat{\rho}$? *Checkpoint:* About 1.43 times; $\hat{\rho}$ about 0.7 times the truth.

(b) Which of the chapter's estimates can a health department still publish, and which needs an independent check of p_h first?

14.11 Application area · 4 min

The dial costs $c_u u^2$ a day with $c_u = \$50$ million.

(a) The shipped threshold of case 2 sets the dial to 0.8 while its measures are on. Find the daily cost while on, and the number of days on that make up the $4,768 million of its 240-day run. *Checkpoint:* $32 million a day; 149 days.

(b) Find the constant dial that, held for the 180 days of the challenge and its 28-day hand-over, spends exactly the $3,600 million budget. *Checkpoint:* 0.588.

(c) Compare the cost of a dial of 0.5 for 60 days with a dial of 1.0 for 30 days. *Checkpoint:* $750 million against $1,500 million.

14.12 Review the agent · 5 min

An AI assistant is asked how to pass the challenge and replies as follows. "Testing is cut in half on day 60. The EKF carries ρ with no bound, so it reads the halved count as a collapse in transmission, its $\hat{R}_t$ drops and any planner fed by it relaxes too early. The MHE holds ρ in a band of 0.1 to 0.8 and cannot be fooled this way. Use the scenario MPC, and it is essential to switch the estimator to the MHE, because on the EKF the tree will fail." Find the flaw and say how you would check it.