

Optimal Disease Outbreak Decisions using Stochastic Simulation

Mike Ludkovski
Jarad Niemi

Department of Statistics and Applied Probability
University of California
Santa Barbara, CA 93106-3110, USA

ABSTRACT

Management policies for disease outbreaks balance the expected morbidity and mortality costs versus the cost of intervention policies. We present a methodology for dynamic determination of optimal policies in a stochastic compartmental model with parameter uncertainty. Our approach is to first carry out sequential Bayesian estimation of outbreak parameters and then solve the dynamic programming equations. The latter step is simulation-based and relies on regression Monte Carlo techniques. To improve performance we investigate lasso regression and global policy iteration. Comparisons demonstrate the realized cost savings of choosing interventions based on the computed dynamic policy over simpler decision rules.

1 INTRODUCTION

Management and design of cost-effective intervention policies to mitigate disease outbreaks is one of the central goals of public health biosurveillance. Random interactions between individuals make the overall outbreak progression stochastic. In addition, health surveys cover only a fraction of total population, and parameters governing outbreak severity are only known in retrospect. As a result, inference of outbreak state, such as the current number of infecteds, is a significant statistical problem. For the aim of designing intervention policies, sequential inference is necessary and has been studied among others in Dukic, Lopes, and Polson (2010), Merl, Johnson, Gramacy, and Mangel (2009).

Management of disease outbreaks constitutes a multi-period control problem. Given the multiple levels of uncertainty and stochasticity, adaptive policies are desired. We propose to adopt a Bayesian stochastic control setup which takes the full posterior of outbreak parameters and population types as the system hyper-state. This leads to a high-dimensional problem for which standard methods of Dynamic Programming and Stochastic Dynamic Programming are infeasible. To overcome this challenge we construct a simulation-based algorithm that relies on techniques from high-dimensional regression.

For concreteness we focus on a popular stochastic SIRD model which is appropriate for diseases with short incubation periods and relatively high mortality, such as pandemic influenza (Ludkovski and Niemi 2010, Patel, Longini, and Halloran 2005) or measles (Ludkovski and Niemi 2011, World Health Organization 2010). In terms of interventions we focus on the interplay between information collection and direct action (especially vaccination). Indeed, parameter uncertainty obscures true outbreak dynamics but can be actively managed by “probing” the system to gather more information through increased survey

proportions. Such actions can reduce uncertainty and expected costs and allow to quantify precisely the *value of information*.

One of our aims in this paper is to assess a number of current data gathering practices and determine whether it may be worthwhile to invest in biosurveillance systems that provide more rapid or accurate assessment of the current outbreak state. For example, most surveillance methods provide little information on recovered. However, the recovery rate is key for determining the basic reproductive ratio \mathcal{R} of the outbreak which controls the proportion of the susceptible population that will become infected. But if there is already enough *a priori* information about the recovery rate, then this data collection is unnecessary. In addition to the importance of recovery rate information, another key factor is timeliness of information. Current biosurveillance systems have data delays around a couple of weeks. For outbreaks such as measles that occur on the scale of weeks, this could have devastating impacts from morbidity and mortality incurred due to delayed interventions, which must be weighed versus the cost of speeding up data delivery.

To answer these questions, we extend and improve the Monte Carlo framework of Ludkovski (2009) and Ludkovski and Niemi (2010). In particular, we design (i) a new global policy iteration mechanism to facilitate convergence and (ii) propose the use of shrinkage regression, such as lasso, during the prediction step. These ideas allows us to significantly speed-up the performance of the algorithm and consider more sophisticated model specifications than in the past. Moreover, our approach can be straightforwardly adjusted for many other outbreak specifications.

2 MODEL

In this section, we describe our model that is based on discrete-time surveillance observations of the numbers of individuals who have become infected or recovered since the last surveillance data. The corresponding inferential procedure incorporates learning of outbreak parameters at a rate that is proportional to the population sampling rate and the variability of this sampling procedure.

We assume a stochastic susceptible-infected-recovered-deceased (SIRD) model for disease outbreak (Andersson and Britton 2000). Thus, individuals transition between four possible states $X_t = (S_t, I_t, R_t, D_t)$ according to a continuous-time jump-Markov process within a population of constant size $N = S_t + I_t + R_t + D_t$. The transition rates are summarized by four reaction channels with intensity functions

$$\left\{ \begin{array}{ll} S + I \xrightarrow{\theta_1 h_1} I & h_1 = k_I^{\phi_t} S_t I_t \\ I \xrightarrow{\theta_2 h_2} R & h_2 = I_t \\ S \xrightarrow{\theta_3 h_3} R & h_3 = \phi_t^V S_t \\ I \xrightarrow{\theta_4 h_4} D & h_4 = I_t \end{array} \right\}, \quad (1)$$

which respectively correspond to infection, recovery, vaccination and death. The *a priori* unknown outbreak parameters Θ can be interpreted as: infectiousness (θ_1), mean recovery time ($1/\theta_2$), vaccination rate (θ_3), and death rate (θ_4). A key quantity is the basic reproduction number $\mathcal{R} := \theta_1/(\theta_2 + \theta_4)$. The ϕ_t 's are the interventions described in Section 3 ($\phi = 0$ indicates no intervention).

Simulation of the SIRD model can be accomplished using Gillespie's SSA algorithm. Since in our applications N is moderately large (on the order of 10^4) and our method is simulation-intensive, we employ the tau-leaping approximation (Gillespie 2001). Accordingly, total transitions over an interval of length tau, taken to be 1 for simplicity, are given by the Poisson rates

$$\tilde{I}_t \sim Po(\theta_1 h_1), \quad \tilde{R}_t \sim Po(\theta_2 h_2), \quad (2a)$$

$$\tilde{V}_t \sim Po(\theta_3 h_3), \quad \tilde{D}_t \sim Po(\theta_4 h_4), \quad (2b)$$

where $Po(\lambda)$ represents the Poisson distribution with mean λ . The compartmental populations are then updated as

$$S_{t+1} = S_t - \Delta I_t - \Delta V_t \quad (3a)$$

$$I_{t+1} = I_t + \Delta I_t - \Delta R_t - \Delta D_t \quad (3b)$$

$$R_{t+1} = R_t + \Delta R_t + \Delta V_t \quad (3c)$$

$$D_{t+1} = D_t + \Delta D_t, \quad (3d)$$

where to ensure that S, I, R, D remain non-negative we place the constraints

$$\Delta I_t = \min(\tilde{I}_t, S_t) \quad (4a)$$

$$\Delta R_t \sim \text{Bin}(I_t + S_t, \theta_2 / [\theta_2 + \theta_4]) \quad \text{if } \tilde{R}_t + \tilde{D}_t > I_t + \Delta I_t \quad (4b)$$

$$\Delta V_t = \min(\tilde{V}_t, S_t - \Delta I_t) \quad (4c)$$

$$\Delta D_t = \tilde{R}_t + \tilde{D}_t - \Delta R_t, \quad (4d)$$

with $\text{Bin}(n, q)$ representing the binomial distribution with n trials and success probability q . These transitions ensure that 1) only available susceptibles become infected, 2) recovered and deceased are not greater than the number of infecteds plus susceptibles (which implies that in the temporal dynamics an individual can become infected and either recover or die in the same time period), 3) the number of recovered and deceased individuals are proportioned appropriately, and 4) if an individual is vaccinated and infected in the same period, they still become infected.

Biosurveillance programs do not observe S, I, R, D but instead collect information through sampling models, assumed to be independent of the system. Given sampling proportions \tilde{p}_i , the observations are $\vec{Y}_t = (Y_{1,t}, \dots, Y_{4,t})$ with

$$Y_{1,t} \sim \text{Bin}(\Delta I_t, p_1) \quad p_1 = \tilde{p}_1 + k_{\Delta I} \phi_t^{\Delta I} \quad (5a)$$

$$Y_{2,t} \sim \text{Bin}(\Delta R_t, p_2) \quad p_2 = \tilde{p}_2 + k_{\Delta R} \phi_t^{\Delta R} \quad (5b)$$

$$Y_{3,t} \sim \text{Bin}(\Delta V_t, p_3) \quad p_3 = \tilde{p}_3 \quad (5c)$$

$$Y_{4,t} \sim \text{Bin}(\Delta D_t, p_4) \quad p_4 = \tilde{p}_4, \quad (5d)$$

where $k_{\Delta I}, k_{\Delta R}$ are known constants and $\phi^{\Delta I}, \phi^{\Delta R}$ are associated with the controls discussed in Section 3.

2.1 Inference

Ideally, inference would be performed accounting for uncertainty in parameters, as well as the underlying SIRD states. Unfortunately, the highly computational nature of the stochastic control procedure introduced in Section 3 precludes the additional computational overhead involved with Monte Carlo techniques. Therefore, we strive to retain as much uncertainty as possible while still maintaining computational tractability in construction of an optimal control policy.

We assume the state of the system is fully observed, but parameter posteriors are only updated utilizing sampled transition data. This allows us to maintain conjugacy relations but still capture sampling noise. Assuming independent gamma priors for the model parameters, the posterior distributions also have the gamma distribution provided by the approximate conjugate updating in equation (6):

$$p(\theta) = \prod_{i=1}^4 p(\theta_i; \alpha_{i,0}, \beta_{i,0}) \quad \implies \quad p(\theta | y_{1:t+1}) = \prod_{i=1}^4 p(\theta_i; \alpha_{i,t} + y_{i,t+1}, \beta_{i,0} + p_i h_i(X_t)). \quad (6)$$

With these assumptions, the current information available at time t by which a decision can be made is summarized by the Markov hyper-state

$$\mathcal{X}_t := (S_t, I_t, D_t, \alpha_{1,t}, \beta_{1,t}, \alpha_{2,t}, \beta_{2,t}, \alpha_{3,t}, \beta_{3,t}, \alpha_{4,t}, \beta_{4,t}). \quad (7)$$

Since the number of recovered is simply $R_t = N - S_t - I_t - D_t$, it is omitted for the remainder of the paper. The state space of \mathcal{X} is of mixed type: $\Delta_N^3 \times (\mathbb{Z}_+ \times \mathbb{R}_+)^4$, where $\Delta_N^3 = \{(x_1, x_2, x_3) \in \mathbb{Z}^3 : x_i \geq 0, \sum_i x_i \leq N\}$.

3 STOCHASTIC CONTROL

We assume the available public health interventions are 1) increasing the sampling rate for new infecteds, 2) sampling newly recovered, 3) isolation of infected individuals, and 4) vaccination. Each of these interventions is controlled by a parameter ϕ which is 0 when the intervention is not in effect and 1 when the intervention is. Table 1 provides a summary of the interventions, their control parameter, and associated costs.

Table 1: Table of interventions and costs .

Intervention	ϕ	Effect	Costs	
			Fixed	Running
Raise new infecteds' sampling	$\phi^{\Delta I} \in \{0, 1\}$	$k_{\Delta I} = 0.9$	$K_{\Delta I} = 100$	$c_{\Delta I} = 100$
Sample new recovered	$\phi^{\Delta R} \in \{0, 1, 2\}$	$k_{\Delta R} = 0.4$	$K_{\Delta R} = 100$	$c_{\Delta R} = 100\phi^{\Delta R}$
Isolation and quarantine	$\phi^I \in \{0, 1\}$	$k_I = 0.6$	$K_I = 1000$	$c_I = 5I_t + 400$
Susceptibles vaccination	$\phi^V \in \{0, 1\}$		$K_V = 50,000e^{-t/7} + 5000$	$c_V = 500$
Outbreak Morbidity			$d_t^1 = I_t + 0.01[\max(I_t - I^-, 0)]^2$	
Outbreak Mortality			$d_t^2 = 500(D_{t+1} - D_t)$	

The costs are broken up into two general categories: fixed and running. The direct costs of the outbreak are provided in the last two lines of the table where $I^- = 400$ is a threshold that converts the cost from linear to quadratic. Fixed costs K_j are incurred immediately when an intervention is initiated, while running costs c_j are incurred the entire time an intervention is active. The cost over one period $[t, t+1]$ is, therefore,

$$C_t^{t+1}(\phi_{t+1}) = \sum_j K_j 1_{\{\phi_{t+1}^j \neq \phi_t^j\}} + \sum_j c_j \phi_t^j + d_t^1 + d_t^2, \quad (8)$$

where the sums are over all possible control policies $j \in \{\Delta I, \Delta R, I, V\}$. The total cost of the outbreak and interventions, denoted as C_0^τ , is the sum over all time points of these costs, where τ is the horizon of the outbreak (i.e., first time when $I_t = 0$; due to finite population every outbreak eventually stops). Additional constraints on possible intervention sequences can be incorporated via the sets $F(\phi)$ which denote possible policies given that previous policy was ϕ .

Since data are gathered sequentially in time, any decision policy is really a *policy map* $\Phi(t, \mathcal{X}_t, \phi_t)$ that describes the decision that should be made at time t based on the latest hyper-state \mathcal{X}_t and currently implemented policy ϕ_t , representing the vector of all controls at time t . Due to the Markov property, \mathcal{X}_t in fact summarizes *all* the relevant information up to t . If information is collected with a delay δ , then Φ is based on $\mathcal{X}_{t-\delta}$. While controls ϕ^I and ϕ^V influence the transitions of (S_t, I_t, R_t, D_t) , the controls $\phi^{\Delta I}$ and $\phi^{\Delta R}$ influence the hyper-parameters $\alpha_{i,t}$ and $\beta_{i,t}$, so the entire \mathcal{X} is effectively controlled.

3.1 Dynamic Programming

The policy objective can be stated as the impulse control problem for \mathcal{X} ,

$$\inf_{\phi_{0:T}} \mathbb{E} [C_0^T(\phi) \mid \mathcal{X}_0 = \mathbf{x}_0], \quad \phi_0 = \vec{0}, \quad (9)$$

where the sub- and super-scripts on C correspond to outbreak realizations from time 0 to cut-off horizon T and the expectation is both over future observations \tilde{Y}_t and future outbreak realizations $\mathcal{X}_t, t \in \{1, 2, \dots, T\}$. Since the total control space is finite (discrete in time and discrete at each time-point), the infimum in

equation (9) is achieved and an optimal management policy exists. To solve (9) we take advantage of the Markov nature of \mathcal{X} to re-formulate it using the dynamic programming principle. Namely, the continuation cost \mathfrak{C} satisfies (cf. Bertsekas (2005), Fleming and Soner (1993))

$$\mathfrak{C}(t, \mathcal{X}_t, \phi_t) := \min_{\phi_{t+1:T}} \mathbb{E} \left[C_t^T(\phi) \mid \mathcal{X}_t \right] = \min_{\phi_{t+1} \in F(\phi_t)} \left\{ \mathbb{E} [C_{t+1}^{t+1}(\phi_{t+1}) \mid \mathcal{X}_t] + \mathbb{E} [\mathfrak{C}_{t+1}(\phi_{t+1}) \mid \mathcal{X}_t] \right\}. \quad (10)$$

To lighten notation and highlight the role of the current policy, we simply write $\mathfrak{C}_t(\phi_t) = \mathfrak{C}(t, \mathcal{X}_t, \phi_t)$. We interpret $\mathfrak{C}_t(\phi_t)$ as the minimal expected costs on $[t, T]$ given starting state \mathcal{X}_t and initial policy ϕ_t . The original objective starting with no interventions is thus to compute $\mathfrak{C}_0(\vec{0})$.

Standard dynamic programming consists of solving (10) via backward recursion starting with the terminal condition $\mathfrak{C}_T(\phi) \equiv 0$. However, given that \mathcal{X} is 11-dimensional, standard space-discretization methods are not computationally feasible. Instead, the idea of regression Monte Carlo methods is to compute the solution on a stochastic mesh corresponding to a collection of simulated paths $x_{0:T}^{(k)}$ of \mathcal{X} , $k = 1, \dots, K$.

Recall that in a basic Monte Carlo method, the one-step-ahead expectations in equation (10) are approximated by the empirical average $\mathbb{E} \left[\mathfrak{C}_{t+1}(\phi_{t+1}) \mid \mathcal{X}_t = x_t^{(k)} \right] \approx \frac{1}{n} \sum_{j=1}^n \tilde{c}_{t+1}^{[j]}$, where $\tilde{c}_{t+1}^{[j]}$, $j = 1, \dots, n$ are i.i.d. realizations of $\mathfrak{C}_{t+1}(\phi_{t+1}) \mid x_t^{(k)}$. To avoid having to generate such a Monte Carlo *forest* for each node $x_t^{(k)}$ in the stochastic mesh, we rather generate just a single realization $c_{t+1}^{(k)} \sim \mathfrak{C}_{t+1}(\phi_{t+1}) \mid x_t^{(k)}$. We then replace empirical averaging by a cross-sectional regression of the $c_{t+1}^{(k)}$'s against chosen outbreak statistics $B_i(x_t^{(k)})$, $i \in I$ at the mesh points $x_t^{(k)}$. In other words, we take the projection of $\mathfrak{C}_{t+1}(\phi_{t+1}) \mid \mathcal{X}_t$ onto $\text{span}(B_i(x), i \in I)$ and approximate this projection through empirical regression. The obtained regression coefficients $\vec{\alpha}_{t+1}^\phi$ provide a prediction for expected future costs based on today's intervention ϕ_{t+1} . Minimizing over all possible $\phi_{t+1} \in F(\phi_t)$ gives the new policy map Φ_t at date t . Iterating this procedure backwards using (10) we eventually obtain \mathfrak{C}_0 . Note that the policy maps Φ_t are pre-computed, and the policy maker only needs access to the coefficients $\vec{\alpha}$ to implement the computed strategy based on the current \mathcal{X}_t .

The overall computational complexity of our method is $O(T^2 \cdot K \cdot D)$ where T is the number of decision epochs, K is the number of Monte Carlo simulations used, and D is the number of maximal control combinations (24 in our example). Importantly, in the algorithm the inference step must be performed T^2KD times, which requires fast inference techniques and motivates our selection of the conjugate updating setup in (6). Theoretically, as $K \rightarrow \infty$, this approach is guaranteed to obtain the optimal policy maps (Ludkovski 2009). Practically, three crucial steps must be addressed to obtain acceptable performance. These concern how to (i) generate the stochastic mesh ($x_t^{(k)}$); (ii) select regression basis functions ($B_i(x)$) $_{i \in I}$; (iii) control accumulation of errors during backward recursion.

3.2 Scenario Re-simulation

Repeated approximations in (10) can lead to rapid error accumulation. To mitigate this problem, we focus on approximating the policy maps Φ rather than the continuation costs \mathfrak{C} . More precisely, at each step t , we re-simulate one forward outbreak path $\tilde{x}_{t+1:T}^{(k, \phi)}$ starting from $x_t^{(k)}$, using control ϕ at t , and implementing future controls based on the already constructed policy maps from time $t+1$ to T . Summing the associated costs gives the pathwise cost $c_{t+1}^{(k)}$ which is exact for this scenario modulo wrong future policy decisions. Thus, approximation of (10) only contributes an error when it leads to incorrect *ranking* of optimal future interventions. The justification for such re-simulation to reduce noise can be found in Ludkovski and Niemi (2010) and goes back to the seminal idea of Longstaff and Schwartz (2001).

3.3 Shrinkage Regression

The choice of basis functions $B_i(x)$, $i \in I$ directly influences the shape of the policy maps Φ . For instance, if B_i consist of degree-2 polynomials in the components of \mathcal{X} then policy maps will consist of parabolic

boundaries. To achieve flexibility, a large number $|I|$ (exponential in $\dim(\mathcal{X})$) of basis functions is desired. Conversely, large $|I|$ may lead to overfitting or instability during the regression step, especially in the presence of outliers that randomly arise from Monte Carlo sample noise. This effect is widespread in our experience. To have a uniform bound on the variance of the regression coefficients α , the number of paths K must be exponential in $|I|$.

We resolve this issue by implementing *shrinkage regression* during the prediction step. In particular, the lasso methods (Hastie et al. 2009) introduce an L^1 penalty on the regression coefficients $\vec{\alpha}$,

$$\vec{\alpha}_t^\phi = \operatorname{argmin}_{\vec{\alpha} \in \mathbb{R}^{|I|+1}} \sum_{k=1}^K \left[c_t^{k,\phi} - \vec{\alpha}' B \left(x_t^{(k)} \right) \right]^2 + \lambda \sum_{i=1}^{|I|} |\alpha^i|. \quad (11)$$

Intuitively, shrinkage allows to pick a large *basis dictionary* I and then adaptively select a sparse subset $I_0 \subset I$ of the most relevant basis functions for the fitting at hand. Lasso regression is furthermore computationally efficient and can be done in same time complexity as regular least-squares. We find that reducing the L^1 norm of $\vec{\alpha}$ by a factor of 3–5 (relative to pure L^2 -projection when $\lambda = 0$) works well. In (Ludkovski and Niemi 2011) we also investigated the use of multivariate adaptive regression splines (MARS) (Hastie et al. 2009, Ch. 3), which is another nonparametric method for robust regression in high dimensions.

3.4 Base-Case Policy Iteration

The goal of the regression step is to approximate the map $(x_t, \phi_t) \mapsto \mathbb{E}[\mathcal{C}_{t+1}(\phi_{t+1}) | \mathcal{X}_t = x_t, \phi_t]$. As with any regression, accuracy of predictions is highest in neighborhoods where many scenarios reside. But, to have a good approximation to the true solution, we need the regression to be accurate in the neighborhood where the optimal \mathcal{X}_t^* is likely to be found (e.g. if optimally managed outbreaks always stay below 50 infecteds, our forward-simulated scenarios should reflect this fact) —which is *a priori* unknown. Indeed, during the forward re-simulation steps we need \tilde{x}_s to lie close to the original $x_s^{(k)}$; otherwise the use of policy map would require *extrapolation* and lead to potentially large errors.

All of the above suggests that we generate the original stochastic mesh $(x_{0:T}^{(k)})$ by simulating a large outbreak scenario database starting with known initial condition \mathcal{X}_0 and a *base-case policy map* Φ^0 . This Φ^0 will influence the accuracy of the solution. Accordingly, we employ *policy iteration* coupled with a cooling schedule. We begin with a guess for Φ^0 and then re-run the entire algorithm using the computed policy Φ as the next stage base-case policy. To be able to judge non-optimal control sequences that are considered during backward recursion, we allow deviations from $\Phi^{(j-1)}$ when constructing $x_t^{(k,j)}$. Namely, to generate $x_{t+1}^{(k,j)}$ we use the recommended intervention $\Phi^{(j-1)}(t, x_t^{(k,j)}, \cdot)$ with probability $1 - q_j$, and an arbitrary intervention with probability q_j . In analogue to other cooling schedules, q_j decreases over cycles of policy iterations j . Practically, we find that three or four such global policy iterations with, e.g., $q_j = 0.2 - 0.05j$ suffice to achieve convergence. The resulting Algorithms are summarized in the Appendix.

4 RESULTS

As an example we took costs as in Table 1 and priors $\vec{\alpha}_{\cdot,0} = (22.5, 15, 4, 200)$, $\vec{\beta}_{\cdot,0} = (30, 30, 80, 1000)$ with $N = 20,000$ and $I_0 = 20, R_0 = D_0 = 0$. This roughly corresponds to a recent outbreak of measles in Harare, Zimbabwe (World Health Organization 2010). While Harare is a multi-million metropolis, the unvaccinated (mostly young) population susceptible to measles was much smaller. Using $K = 4 \times 10^4$ simulations, and $|I| = 17$ basis functions, we computed the optimal solution.

The resulting frequency of intervention actions over time is shown in Figure 1. To generate that figure, we employed lasso regression with shrinkage factor of 4. We found that use of shrinkage regression reduces temporal fluctuations in $\vec{\alpha}(t)$ that are observed with ordinary least-squares regression (some residual fluctuations can still be observed in the sampling policy frequency in Figure 1). This corresponds to more robust approximations of the continuation costs \mathcal{C} and less sensitivity to Monte Carlo simulation noise.

The minimized expected costs were about $2.62 \cdot 10^4$. To better judge the contribution of different aspects of the model, Table 2 shows the impact of several possible policy constraints on optimal costs. In

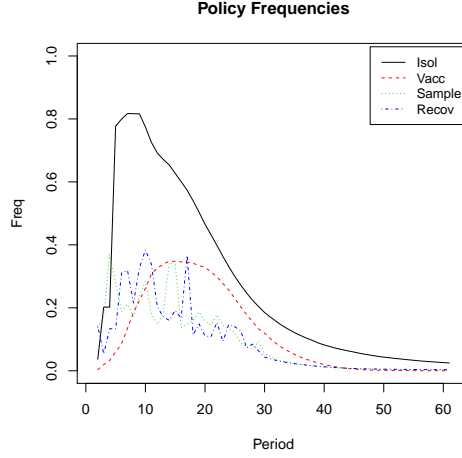


Figure 1: Frequency of different intervention actions over time. Average over 40000 outbreak paths generated based on computed policy map $\Phi^{(4)}$. Since more than one intervention can be undertaken simultaneously, the frequencies sum to more than 1.

Table 2: Impact of policy constraints on optimal costs. Last column compares the effect of constraints on expected costs relative to the base case.

Constraint	Optimal costs ($\times 10^4$)	Relative costs
Base case	2.62	—
$\phi^{\Delta R} \equiv 0$	2.76	105%
$\phi^{\Delta I} = \phi^{\Delta R} \equiv 0$	3.52	134%
Delay $\delta = 1$	2.78	106%
Delay $\delta = 2$	2.98	114%

particular, we observe the importance of timely information collection but also that with above priors, the value of information about recovered is quite low.

To understand the performance of our dynamic policy produced by the scheme in Section 3, we examine it relative to other possible policy alternatives. Beyond the disastrous do-nothing choice that had expected costs of over $55 \cdot 10^4$, we have also tried $\phi^{(I)}$, which corresponds to immediate and permanent isolation measures; $\phi^{(II)}$, that makes a single intervention based on information available after 7 periods; $\phi^{(III)}$, which does a single intervention but at arbitrary date t , and, finally, an (optimized) threshold strategy $\phi^{(IV)}$ which starts isolation as soon as $I_t > 28$ and begins a vaccination campaign as soon as $I_t > 45$. All of the above cases can be seen as simplifications of our full dynamic setup and have been computed using the same numerical procedure. Table 3 shows that our policy essentially matches the threshold strategy and vastly outperforms all other choices. Note that the threshold policy is based on heuristic arguments and the threshold levels were optimized by hand; in contrast our dynamic approach is a ‘black box’ that in this case can be seen as validating such heuristics. Without this black box we would not know how to validate such choices; conversely this example shows that there is room for improvement of our method since the returned result is clearly not yet globally optimal.

Figure 2 further visualizes the performance of our dynamic policy vis-a-vis several alternatives. Besides comparing the corresponding cost distribution $C_0^T(\phi)$, we also look at total deaths D_τ , number of infected days $\sum_t I_t$, and duration of outbreak τ . The isolation-only strategy $\phi^{(I)}$ is often highly efficient but has an extreme right tail that arises when the reproduction ratio is high. In the latter case, all the outbreak metrics are very large and the resulting costs are enormous. In other words, $\phi^{(I)}$ does not manage the risk of extreme outbreaks. The other three strategies all exhibit *bimodal* cost distributions which correspond to

Table 3: Comparison of optimal policy to simpler alternative strategies.

Strategy	Costs ($\times 10^4$)
Isolation forever	11.31
Single Decision at $t = 7$	3.64
Single Adaptive Decision	3.59
Threshold strategy	2.61
Optimal Dynamic Policy	2.62

cases where isolation alone/isolation and vaccination are required. We observe that the optimal dynamic strategy is best able to mitigate worst-case outbreaks (thinnest right tail) and ends the outbreak fastest. Counter-intuitively we find that it does not minimize the average number of infected days, which highlights the ongoing trade-off between intervention and morbidity costs. Overall, we see that the complexity of the problem makes it difficult to directly predict properties of optimal policies, showcasing the need for automated software to provide guidance.

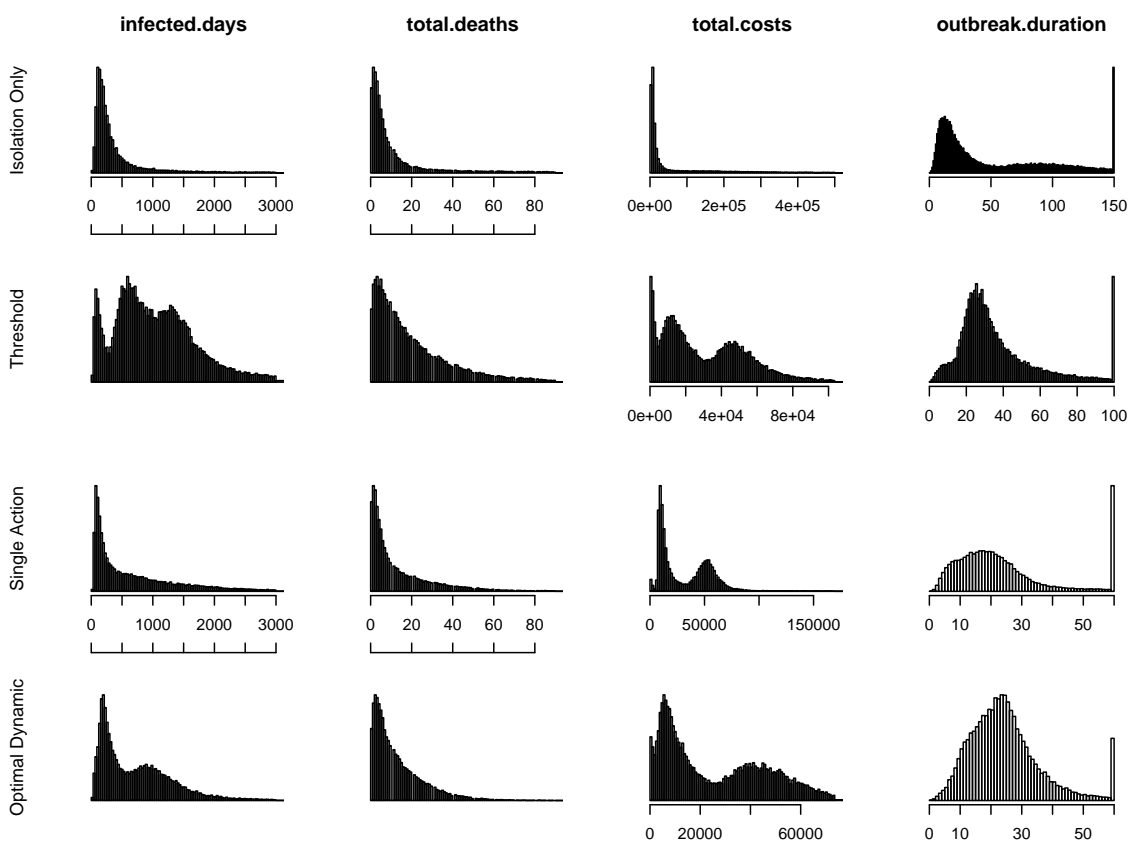


Figure 2: Cost distributions for different strategies. The rows correspond to four alternative policies $\phi^{(I)}$, $\phi^{(III)}$, $\phi^{(IV)}$, and the dynamic ϕ^* , respectively. The columns show histograms of four different outbreak metrics over 40,000 simulations. Note the varying scales in the third and fourth columns.

ACKNOWLEDGEMENTS

Ludkovski's research was partially supported by the Hellman Family Foundation Grant.

A APPENDICES

A.1 Sequential flu management algorithm

Algorithm 1 creates the (approximately) optimal policy map $\Phi(t, x_t, \phi_{t-1})$ for all time points $t \in \{0, 1, \dots, T-1\}$. As input, the algorithm requires 1) the number of simulations K to perform at all time points, 2) a vector of outbreak statistics $B(x_t)$ to use in the regression step, and 3) a default policy map Φ^0 for initial simulation. The algorithm returns the regression coefficients α_t^ϕ .

Algorithm 1 Optimal policy map creation

for $k \in \{1, 2, \dots, K\}$ **do**

 Simulate $x_{1:T}^{(k)}$ using default policy map Φ^0 by repeated calls to Algorithm 2

end for

for $t = (T-1), \dots, 1, 0$ **do**

for each policy ϕ **do**

for $k \in \{1, 2, \dots, K\}$ **do**

 Simulate $\tilde{x}_{t+1}^{(k, \phi)}$ and calculate costs $c_t^{k, \phi}$ starting from $\tilde{x}_t^{(k)} = x_t^{(k)}$ using Algorithm 2 and implementing ϕ .

for $s = t+1, \dots, T-1$ **do**

 Simulate $\tilde{x}_{s+1}^{k, \phi}$ and record costs $\tilde{c}_t^{(k, \phi)}$ starting from $\tilde{x}_s^{(k)}$ using Algorithm 2 and implementing $\Phi(s, \tilde{x}_s^{(k)}, \phi_s - 1)$.

 Calculate cumulative costs $c_t^{(k, \phi)} = c_t^{(k, \phi)} + \tilde{c}_t^{(k, \phi)}$.

end for

end for

 Compute the regression coefficients according to (11). This provides the policy map

$$\Phi(t, x_t, \phi_{t-1}) = \underset{\phi \in F(\phi_{t-1})}{\operatorname{argmin}} (\alpha_t^\phi)' B(x_t)$$

end for

end for

return Return regression coefficients $\alpha_{0:T}^\phi$ for all ϕ .

A.2 Forward simulation and associated costs

Algorithm 2 consists of simulating one-step of a future controlled outbreak scenario and calculating the realized cost. The required inputs are a control ϕ and initial information x_t . The algorithm returns updated information x_{t+1} and one-step costs c_{t+1} .

Algorithm 2 Simulate an outbreak scenario from time t to time $t+1$ using policy map $\Phi(t, x_t, \phi_{t-1})$ and calculate costs c_{t+1} .

 Sample $\theta \sim p(\theta|x_t)$, i.e. equation (6).

 Simulate one step of an outbreak scenario using equations (2)-(5) and control ϕ .

 Update hyperparameters using equation (6).

 Calculate costs c_{t+1}

return Simulated outbreak information \mathcal{X}_{t+1} and cost c_{t+1} .

REFERENCES

- Andersson, H., and T. Britton. 2000. *Stochastic epidemic models and their statistical analysis*, Volume 151 of *Lecture Notes in Statistics*. New York: Springer-Verlag.
- Bertsekas, D. P. 2005. *Dynamic programming and optimal control. Vol. I*. Third ed. Athena Scientific, Belmont, MA.
- V. Dukic and H.F. Lopes and N.G. Polson 2010. “Tracking flu epidemic using Google Flu Trends and particle learning”. working paper.
- Fleming, W. H., and H. M. Soner. 1993. *Controlled Markov processes and viscosity solutions*, Volume 25 of *Applications of Mathematics (New York)*. New York: Springer-Verlag.
- Gillespie, D. 2001. “Approximate accelerated stochastic simulation of chemically reacting systems”. *The Journal of Chemical Physics* 115:1716.
- Hastie, T., R. Tibshirani, and J. Friedman. 2009. *The elements of statistical learning: data mining, inference, and prediction*. 2 ed. Springer Verlag.
- Longstaff, F., and E. Schwartz. 2001. “Valuing American options by simulations: a simple least squares approach”. *Rev. Finan. Studies* 14:113–148.
- Ludkovski, M. 2009. “A Simulation Approach to Optimal Stopping Under Partial Observations”. *Stochastic Processes and Applications* 119(12):4061–4087.
- Ludkovski, M., and J. B. Niemi. 2010. “Optimal dynamic policies for influenza management”. *Statistical Communications in Infectious Diseases* 2 (electronic).
- Ludkovski, M., and J. B. Niemi. 2011. “Optimal management of disease outbreaks”. Technical report, in preparation.
- Merl, D., L. Johnson, R. Gramacy, and M. Mangel. 2009. “A statistical framework for the adaptive management of epidemiological interventions”. *PLoS ONE* 4(6):e5087.
- Patel, R., I. Longini, and M. Halloran. 2005. “Finding optimal vaccination strategies for pandemic influenza using genetic algorithms”. *Journal of Theoretical Biology* 234:201212.
- World Health Organization 2010, December. “Zimbabwe cholera epidemiological bulletin number 89”. http://www.who.int/entity/hac/crises/zwe/sitreps/zwe_epi_12december2010.pdf.

AUTHOR BIOGRAPHIES

MIKE LUDKOVSKI is an Assistant Professor in the Department of Statistics and Applied Probability, at University of California Santa Barbara. He received a Bachelors of Science from Simon Fraser University in British Columbia, Canada, a Ph.D. in Operations Research and Financial Engineering from Princeton University and was previously a term assistant professor at University of Michigan. His research interests are in stochastic control and applied probability with applications ranging from stochastic games in resource management to sequential change detection and control in biosurveillance models. His email address is ludkovski@pstat.ucsb.edu.

JARAD NIEMI is an Assistant Professor in the Department of Statistics and Applied Probability at the University of California, Santa Barbara. He received a Bachelors of Chemical Engineering and a Masters of Science in Biostatistics from the University of Minnesota in Minneapolis, MN, USA and a Ph.D. in Statistics from Duke University in Durham, NC, USA. His research interests are in building dynamic models for disease outbreaks and biochemical systems and analysis through Bayesian methods including Markov chain Monte Carlo and sequential Monte Carlo. His email address is niemi@pstat.ucsb.edu.