amei: an R package for the Adaptive Management of Epidemiological Interventions

Daniel Merl
Department of Statistical Science
Duke University, Durham NC
dan@stat.duke.edu

Leah R. Johnson
Department of Ecology, Evolution, and Marine Biology
University of California, Santa Barbara
lrjohnson@lifesci.ucsb.edu

Robert B. Gramacy
Statistical Laboratory
University of Cambridge, United Kingdom
bobby@statslab.cam.ac.uk

Marc S. Mangel
Center for Biomolecular Science and Engineering, and
Department of Applied Mathematics and Statistics
University of California, Santa Cruz
msmangel@ams.ucsc.edu

April 1, 2009

Abstract

The amei package for R [26] is a tool that provides a flexible statistical framework for generating optimal epidemiological interventions that are designed to minimize the total expected cost of an emerging epidemic. Uncertainty regarding the underlying disease parameters is propagated through to the decision process via Bayesian posterior inference. The strategies produced through this framework are adaptive: vaccination schedules are iteratively adjusted to reflect the anticipated trajectory of the epidemic given the current population state and updated parameter estimates. This document briefly covers the background and methodology underpinning the implementation provided by the package and contains an extensive example showing the functions and methods in action.

Intended audience

This document is intended to familiarize a (potential) user of `amei` with the models implemented and analyses available in the package. After a brief overview, the bulk of this document consists of a detailed example illustrating the various functions and methodologies implemented. This document has been authored in `Sweave` (try `help(Sweave)`). This means that the code quoted throughout is certified by R, and the `Stangle` command can be used to extract it. The demo available in this package will run the same code via `demo("amei")`.

Note that this tutorial was not meant to serve as an instruction manual. For more detailed documentation of the functions contained in the package, see the package help–manuals. At an R prompt, type `help(package=amei)`. PDF documentation is also available on the world-wide-web.

`http://www.cran.r-project.org/web/packages/amei/index.html`

This tutorial is comprised of four main sections. Section 1 provides some background and motivation. The mathematical specification of the Bayesian models used for inference and the Monte Carlo methods for constructing optimal vaccination strategies (both static or adaptive and on–line) are contained in Section 2. In Section 3, the functions and methods implemented in the package are illustrated by following a single, detailed, example, whose results were first reported by Merl et al., [20]. The paper concludes in Section 5 with a discussion of the methodology and related work highlighting other freely available software with comparisons and contrasts. Miscellaneous details on implementation, etc., are provided in an appendix.

1 Motivation

The goal of the study of infectious diseases is to better understand how infections are spread and maintained, and ultimately to find ways to control the spread of a disease. The most common methods for intervening in the spread of an infectious disease either remove susceptible individuals or apply treatment to infected individuals. For instance, the susceptible population may be culled, as in the case of foot-and-mouth disease [29, 9], or the infected population may be quarantined, as in the case of SARS [18]. Most commonly, susceptibles may be vaccinated, as in the case with influenza or smallpox [10, 13].

Each of these actions incurs a quantifiable epidemiological cost. For culling, the cost is an additional number of deaths; for quarantine, the cost is likely to be measured in monetary units rather than lost lives; for vaccination, the cost may be measured in both monetary units as well the number of additional vaccine–induced infections; for medical treatment, the cost is again monetary. Furthermore, the costs associated with each action can depend upon the state of the disease within the population of interest. This raises the question of how to find optimal epidemiological interventions in a manner that adaptively depends on the state of the epidemic.

Most existing methods for finding optimal intervention strategies are concerned with the situation of pre-emptive intervention which is assumed to be completed before the onset of the epidemic (for instance see [3, 25, 29]). In this kind of situation, there is no reason to consider sequentially updated (i.e. adaptive) interventions: as soon as the intervention policy is triggered, the epidemic threat will be eradicated. However, in most scenarios, total and instantaneous intervention will not be an implementable strategy. Moreover, these methods usually involve calculations that assume no uncertainty in key model parameters, including transmission rate, recovery rate, and others (however, see Elderd, et al. [8] and our brief discussion in Section 5).

Here we introduce `amei`, a software package that implements a statistical framework introduced by Merl, et al. [20], that allows one to respond to an emerging epidemic while simultaneously learning about it. We consider vaccination strategies defined by a fraction of the current susceptible population to be targeted for vaccination, and a threshold number of susceptibles below which the vaccination campaign is called off. We couple the evaluation of optimal, adaptive, intervention strategies with Bayesian procedures for performing on-line estimation of the parameters of the underlying epidemic model, thereby propagating parameter uncertainty through to policy decisions. We demonstrate the advantages of adaptive intervention via the functions provided by the package using simulations modeled after an influenza outbreak at a British boarding school described by Murray [21]. We compare the distribution of costs arising from epidemiological intervention under the adaptive policies to those arising from non-adaptive policies (i.e. policies not dependent on the state of the epidemic and/or not reflecting parameter uncertainty), and find that the adaptive policies result in low total costs, efficient use of available resources, and are robust to model misspecification.

2 Methods

2.1 SIR Model

In `amei`, we consider a standard Susceptible–Infected–Removed (SIR) model [1, 14] with permanent immunity and with mortality. In this model, the dynamic variables at time t are the number of susceptible individuals, $S(t)$; the number of infected individuals, $I(t)$; the number of recovered individuals, $R(t)$; and the number of removed/dead individuals, $D(t)$. We assume the population is closed to immigration or emigration such that $S(t) + I(t) + R(t) + D(t) = N$, where N is constant.

Models of this type can include transmission dynamics ranging from the very simple to the very complex. We adopt a flexible negative binomial form for the transmission function [19]. Under this assumption, the SIR model is described

by the following system of differential equations [14, 19]:

$$\frac{\mathrm{d}S}{\mathrm{d}t} = -kS\ln\left(1+\frac{bI}{k}\right) \tag{1}$$

$$\frac{\mathrm{d}I}{\mathrm{d}t} = kS\ln\left(1+\frac{bI}{k}\right) - (\nu+\mu)I$$

$$\frac{\mathrm{d}R}{\mathrm{d}t} = \nu I \tag{2}$$

$$\frac{\mathrm{d}D}{\mathrm{d}t} = \mu I \tag{3}$$

The model parameters are: the transmission rate b; the overdispersion (or "clumpiness") parameter k; the death rate μ; and the rate of recovery to the immune class ν. The negative binomial distribution can be interpreted as a compound stochastic process in which encounters between infected and susceptible individuals occur randomly (i.e., according to a Poisson process) such that the encounter rate varies according to a gamma distribution with coefficient of variation $k^{-1/2}$. Thus, via k, the negative binomial transmission can account for social interactions and/or network factors in disease transmission, without requiring explicit characterization of the population structure.

This SIR formulation leads to a natural discrete time approximation for the numbers of infections ($\tilde{I}$), recoveries ($\tilde{R}$), and deaths ($\tilde{D}$) arising in the unit time interval from t to $t+1$. Holding the total number of infected individuals, I, constant and integrating Eq. (1) over a unit time interval gives

$$S(t+1) = S(t)\left[\frac{k}{k+bI(t)}\right]^k,$$

so that the fraction of susceptible individuals surviving a unit time interval is $\left[\frac{k}{k+bI(t)}\right]^k$. Viewed as a discrete time stochastic process, the number of new infections occurring between time t and $t+1$ when $S(t)=s$ and $I(t)=i$ can be therefore described by

$$\tilde{I}|s,i \sim \mathrm{Bin}(s, p_i(i,b,k)), \tag{4}$$

where

$$p_i(i,b,k) = 1-\left(\frac{k}{k+bi}\right)^k$$

and $\mathrm{Bin}(n,\pi)$ is the standard binomial distribution. Similarly, by integrating Eqs. (2–3), the numbers of recoveries and deaths occurring between time t and $t+1$ can be described by

$$\tilde{R}|i \sim \mathrm{Bin}(i, p_r) \tag{5}$$

$$\tilde{D}|i,\tilde{r} \sim \mathrm{Bin}(i-\tilde{r}, p_d) \tag{6}$$

where $p_r = 1 - e^{-\nu}$ and $p_d = 1 - e^{-\mu}$. The forward dynamics for the total numbers of susceptible and infected individuals are therefore

$$\begin{aligned} S(t+1) =& S(t) - \tilde{I}|s,i \\ I(t+1) =& I(t) + (\tilde{I}|s,i) - (\tilde{R}|i) - (\tilde{D}|i,\tilde{r}+\tilde{i}). \end{aligned}$$

Here the lower case symbols $\{s, i, r\}$ denote the realized value of the associated capital letter random variable. In this discrete time approximation we have assumed a particular ordering of events, namely that recoveries occur first, followed by deaths from among those infected individuals who did not recover, followed by new infections. Simulation studies indicated that these assumptions, as well as other possible orderings, resulted in system dynamics that were equal in expectation to the deterministic solutions to the continuous time SIR model.

2.2 Online Parameter Estimation

An important task of `amei` is to be able to estimate the SIR model parameters. Given the discrete time approximation given in the previous section, it is possible to do this online (i.e., as the epidemic progresses) via straightforward parametric Bayesian methods. In particular, we use Markov Chain Monte Carlo (MCMC) [11] to learn about the posterior distributions of b, k, ν, and μ conditioned on the evolution of the epidemic observed so far. The likelihood is given recursively in Eq. (4–6). Let $\tilde{i}_t = S(t-1) - S(t)$ be the number of new infecteds at time T, and similarly for the newly recovered and dead individuals $\tilde{r}_t$ and $\tilde{d}_t$ so that $\tilde{r}_t + \tilde{d}_t \leq I(t-1)$. Then, the likelihood up to time T is given by

$$\begin{aligned} &\prod_{t=1}^{T} \mathrm{Bin}(\tilde{i}_t \,|S(t-1), p_i(I(t-1), b, k)) \\ &\times \prod_{t=1}^{T} \mathrm{Bin}(\tilde{r}_t|p_r) \times \prod_{t=1}^{T} \mathrm{Bin}(\tilde{d}_t|I(t) - \tilde{r}_t, p_d) \end{aligned}$$

and we can see that it consists of three mutually independent components.

Conditional conjugacy can be exploited for ν and μ via beta priors for p_r and p_d. A $\mathrm{Beta}(\alpha_r, \beta_r)$ prior for p_r implies that

$$p(\nu) = (1 - e^{-v})^{\alpha_r - 1} e^{-v\beta_r}. \tag{7}$$

Conjugate updating leads to the posterior conditional

$$p_r|\cdots \sim \mathrm{Beta}(\alpha_r + \textstyle\sum_{t=1}^{T} \tilde{r}_t, \beta_r + \sum_{t=1}^{T} I(t) - \tilde{r}_t). \tag{8}$$

The form of the conditional posterior for ν is similar to Eq. (7) and can be simulated by first drawing p_r via Eq. (8) and then applying the inverse transformation $\nu = -\log(1 - p_d)$. Sampling for μ proceeds similarly with

$$p_d|\cdots \sim \mathrm{Beta}(\alpha_d + \textstyle\sum_{t=1}^{T} \tilde{d}_t, \beta_d + \sum_{t=1}^{T} I(t) - \tilde{r}_t - \tilde{d}_t). \tag{9}$$

Thus it is possible to take Gibbs samples for ν and μ so long as appropriate hyperparameters $\alpha_r, \beta_r, \alpha_d, \beta_d$ can be found to represent our prior beliefs. In ignorance we simply set these to unity by default, leading to a uniform prior on p_r and p_d. The user is free to specify his/her own prior parameterization in the package.

Obtaining samples for b and k requires the Metropolis–Hastings algorithm. Our prior beliefs can be encoded with gamma distributions, and conditional on a previous sample (b, k) the next sample (b', k') can be obtained by Metropolis–within–Gibbs steps using:

$$p(b'|k, \dots) \tag{10}$$
$$\propto \Gamma[b'|\alpha_b, \beta_b] \prod_{t=1}^{T} \text{Bin}(\tilde{i}_t \,|\, S(t-1), p_i(I(t-1), b', k))$$
$$p(k'|b', \dots) \tag{11}$$
$$\propto \Gamma[k'|\alpha_k, \beta_k] \prod_{t=1}^{T} \text{Bin}(\tilde{i}_t \,|\, S(t-1), p_i(I(t-1), b', k')).$$

For the prior settings, the default values set in the package are $(\alpha_b, \beta_b) = (\alpha_k, \beta_k) = (1, 3)$ which (though seemingly informative at first glance) turns out to be uninformative on the scale of the support of the posterior. As before, these can easily be changed by the user. We use random walk uniform proposals on the positive real line, i.e., $b' \sim U[3b/4, 4b/3]$, which gives reasonably good mixing from the Markov chain.

In the presence of a vaccination strategy (described in the next section) necessitates a simple change to the above equations. If $0 \leq v_t \leq S(t-1)$ is the number of susceptibles which have been vaccinated, then we simply replace $S(t-1)$ with $S(t-1) - v_t$ so that $\tilde{i}_t = S(t-1) - v_t - S(t)$.

2.3 Optimal Vaccination Strategies: Fixed and Adaptive

Once we know, or have estimated, the SIR model parameters at some time during the epidemic, we next want find the best way to intervene in the spread of the epidemic. The first step is to define what we mean by the "best" strategy. This requires a specification of the costs of various actions, such as vaccination, verses the cost of allowing the epidemic to spread in an uncontrolled fashion. In `amei`, we formulate the total expected cost of the epidemic in terms of the underlying costs associated with maintaining infected individuals until recovery, suffering death, and administering vaccinations. As it is currently formulated, these costs must all be in some common currency (such as monetary cost, or simply numbers of deaths).

We have formulated our costs and vaccination strategies in terms of a policy where a fraction, α, of susceptibles are prevented from risk of infection by moving them directly into an immune/recovered class, such as by perfect vaccination, until the number of individuals that are still susceptible drop below a threshold,

γ, and vaccination is discontinued. We let $c_1(\alpha, \gamma, s)$ denote the cost associated with the vaccination strategy (α, γ) when $S(t) = s$. Letting c_v denote the cost per vaccine unit, then

$$c_1(\alpha, \gamma, s) = \begin{cases} c_v \alpha s & \text{if } s > \gamma \\ 0 & \text{if } s \leq \gamma \end{cases}$$

We let $c_2(i)$ denote the cost component that depends on the number of infections in the population, $I(t) = i$. This component includes the costs associated with maintaining the non-recovered infected individuals and costs associated with deaths, as in

$$c_2(i) = c_t i + c_d \tilde{d},$$

where c_t is the cost per treatment/maintenance of a non-removed infected individual, and c_d is the cost per death.

Assuming the initial epidemiological state is $S(0) = s_0$, $I(0) = i_0$, the expected total cost of the epidemic under intervention strategy (α, γ) can be expressed recursively as

$$E\{C_0\} = c_1(\alpha, \gamma, s_0) + c_2(i_0) + E\{C_1\}, \tag{12}$$

where $E\{C_t\}$ denotes the expected cost accumulated from time t onwards. The optimal intervention strategy (α, γ) is the one that minimizes the total accumulated cost over the course of the epidemic. Two methods for calculating such strategies are as follows.

The first case we are interested in is when the parameters of the SIR model are exactly known, and we wish to calculate the single best intervention strategy (α, γ) to use over the whole epidemic. The total expected cost depends on the parameter values and the initial epidemiological state (s_0, i_0). Thus, conditional on a set of parameter values, Monte Carlo simulation can be used to search over values of α and γ in order to find the combination that minimizes $E\{C_0\}$. For each combination of α and γ considered, we conduct n stochastic simulations of the outbreak in order to estimate the mean cost associated with the intervention (α, γ). The strategy producing the lowest mean cost is defined to be the optimal intervention. Typically we discretize and create a grid of admissible α and γ settings. In the examples in Section 3 we allow α to range from 0 to 1 in increments of 0.1, and γ to range from from 2 to $s_0 - s_0/10$ in increments of $s_0/10$, i.e., taking 10 steps.[1]

In the second case, we want to calculate an adaptive strategy that updates the best strategy (α, γ) as we learn more about the epidemic over time. As above, the expected cost surface associated with a given set of parameter values (as obtained by MCMC, described above), can be explored using standard Monte Carlo methods. At each time step, MCMC is used to produce samples from the current posterior distribution on model parameters. These samples are used to calculate the optimal vaccination strategy as outlined as above, treating the

[1] We generally do not include $\gamma = s_0$ in the grid since this policy (coupled with any α) is equivalent to $\alpha = 0$ for any γ.

current time step as time zero. The adaptive strategy to be implemented at that time step is defined to be the strategy that most frequently minimizes the cost over the samples from the posterior distribution of the parameters.

3 An illustrative example

In this section we demonstrate the advantages of adaptive intervention using simulations modeled after an influenza outbreak at a British boarding school described by Murray [21]. We shall compare the distribution of costs arising from epidemiological intervention under the adaptive policies to those arising from non-adaptive policies. The epidemic conforms to many standard assumptions of SIR models: a population essentially closed to immigration and emigration; includes recovery and immunity; and has near homogeneous mixing of susceptibles and infectives. The epidemic was traced back to a single infected student out of a population of 763 individuals.

For reproducibility of the results in this section we have set the random seed as follows.

```
> seed <- 12345
> set.seed(seed)
```

We begin by exploring the behavior of the SIR model, described in Section 2.1, without any intervention. The first step is to set the relevant parameters that are necessary for simulating the epidemic. These consist of: the "true" underlying parameters for the SIR model; the initial condition of the population at the beginning of the epidemic; and the relative costs of infections, deaths, and vaccinations.

```
> true <- list(b = 0.00218, k = 10, nu = 0.4, mu = 0)
> init <- list(S0 = 762, I0 = 1, R0 = 0, D0 = 0)
> costs <- list(vac = 2, death = 4, infect = 1)
```

Murray [21] provides estimates of the transmission rate (b) and recovery rate (ν), which are what we use in the `true` parameterization set above. We also set the death rate (μ) to zero since there are no deaths in this epidemic. Finally, we set the "clumpiness" parameter (k) to be large to reflect the homogeneous mixing of the population. The `costs` above describe the unit cost (or loss) for a single vaccination or death, and the daily cost of maintaining an infected individual.

We are interested in the costs of a no–vaccination policy on epidemic trajectories with the above parameters. To explore this, we set the vaccination policy used in the simulation to zero.

```
> vac <- list(frac = 0, stop = 0)
```

We are now ready to run the Monte Carlo experiment. The function `MCepi` can be used to simulate the population and cost trajectories for the experiments.

First, it simulates the stochastic SIR model under a given vaccination strategy. Second, for each realization of the epidemic progression the function calculates the cost over time for the epidemic. This is repeated many times, and the mean trajectories for both the populations and the costs, as well as the 5th and 95th quantiles, are recorded.

```
> init.MCepi <- MCepi(init, true, vac, costs)
```

Now we can simply plot the results to look at the distribution of the susceptible, infected, and recovered individuals in the population as the epidemic progresses (top panel of Figure 1), as well as the distribution of costs over time (bottom panel of Figure 1).

Next we examine the dynamics of the system when we have perfect information and use a fixed vaccination policy. First we need to set up the grid of admissible policies. In this case, we are looking at strategies that: vaccinate a fixed proportion of the susceptible population, from 0 to 100%, in steps of 10%; and that stop vaccinating susceptibles when the remaining susceptible population falls below some threshold between 2 and the initial susceptible population minus 75, in increments of 75, as explained in Section 2.3.

```
> vacgrid <- list(fracs = seq(0, 1, 0.1), stops = seq(2,
+     init$S0 - 75, 75))
```

Once this grid has been initialized, we can run the Monte Carlo experiment, using the function `optvac`, which finds the vaccination policy that minimizes the total cost of the (stochastic) epidemic. This is done by simulating the epidemic forward under the known, `true`, parameterization and calculating the cost of each vaccination strategy.

```
> out.optvac <- optvac(init, true, vacgrid, costs)
```

This function outputs the costs for each of the possible vaccination strategies. The best and worst policies can be obtained as follows:

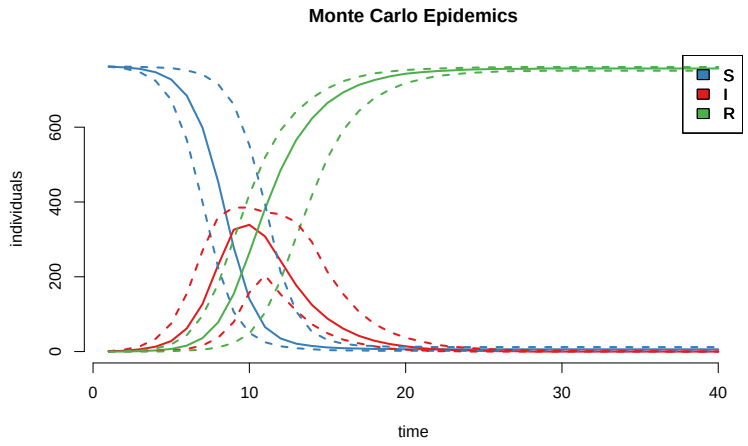
```
> best <- getpolicy(out.optvac)
> worst <- getpolicy(out.optvac, which = "worst")
> rbind(best, worst)
```

```
      row col frac stop cost
best   10   3  0.9  152 1657
worst   1   1  0.0    2 2297
```

The same information can be obtained via the generic `print` and `summary` commands, which will be shown later. We can also plot the cost surface over the space of possible vaccination strategies. This takes the form of a heat plot, where lower cost areas are in deep red, and high cost areas are light yellow (Figure 2).

Given the calculated optimal policy, we can explore effects of the vaccination strategy on the progression of the epidemic together with the trajectory of costs under this strategy. We do this by again simulating the epidemic dynamics using the function `MCepi`, however this time we include the best fixed vaccination policy.

```
> plot(init.MCepi)
```




```
> plot(init.MCepi, type = "costs")
```

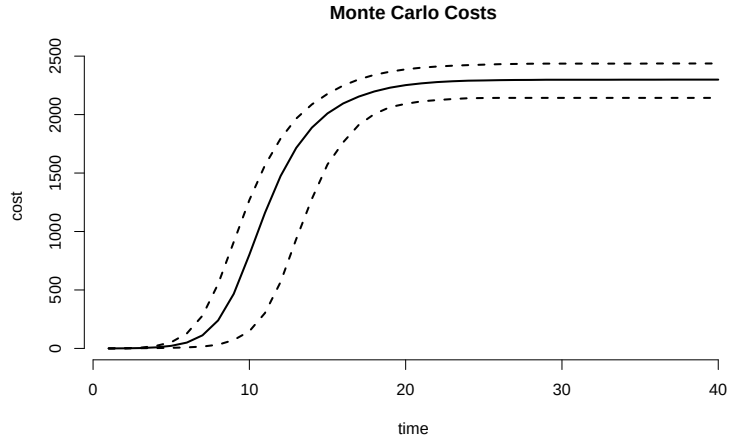


Figure 1: Monte Carlo simulated epidemic trajectories (top) for the numbers of susceptible, infected, and recovered individuals, and the associated cost(s) (bottom) with a *null* vaccination strategy. (2.5,50,97.5%) quantiles are shown.

```
> vac.opt <- best[3:4]
> opt.MCepi <- MCepi(init, true, vac.opt, costs)
```

Figure 3 summarizes the trajectories (top) of the epidemic under the optimal vaccination strategy over time, and the corresponding costs (bottom). By default we assume a fixed lag of 7 time steps from when the first infection appears

```
> plot(out.optvac)
```

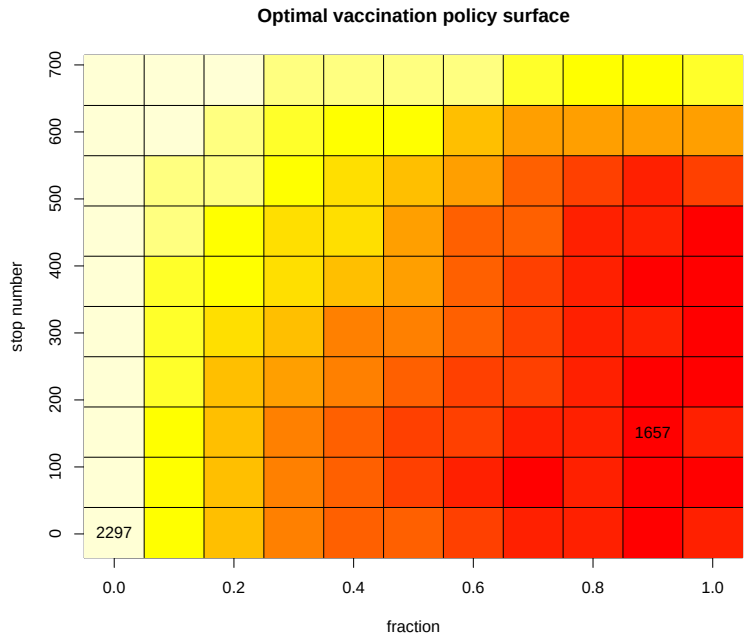


Figure 2: Heatmap depicting the expected cost surface associated with variable stop time vaccination strategies based on the true parameter values. The minimum expected cost (1657) is achieved by a strategy of vaccinating 90% of susceptibles at each time step, until the number of susceptibles falls below 152. As expected, the maximum expected cost (2297 cost units) is realized through inaction (top row and left column policies are never implemented)

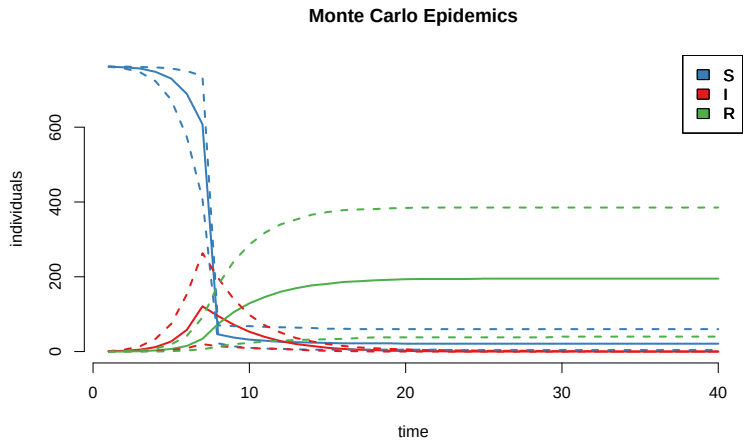
to when the first intervention can take place[2]. This is apparent in Figure 3, where there is a sharp transition from day 7 when the vaccinations begin, and the susceptible population drops dramatically. The number of vaccinated individuals can be added to the plot by specifying the argument `showv = TRUE`, however we omit this here to reduce clutter in the figure. Information on the distribution of the number of vaccine units dispensed can be extracted as follows.

```
> getvac(opt.MCepi)

  q0.025   mean median q0.975
1    368 538.38    548    664
```

We can compare these results to the case without vaccination (Figure 1, top), and see that the optimal vaccination strategy effectively suppresses the spread of the infection. The costs also spike around time 7 as the vaccination policy

[2]This can be varied with the `start` argument to `MCepi`.

```
> plot(opt.MCepi)
```




```
> plot(opt.MCepi, type = "costs")
```

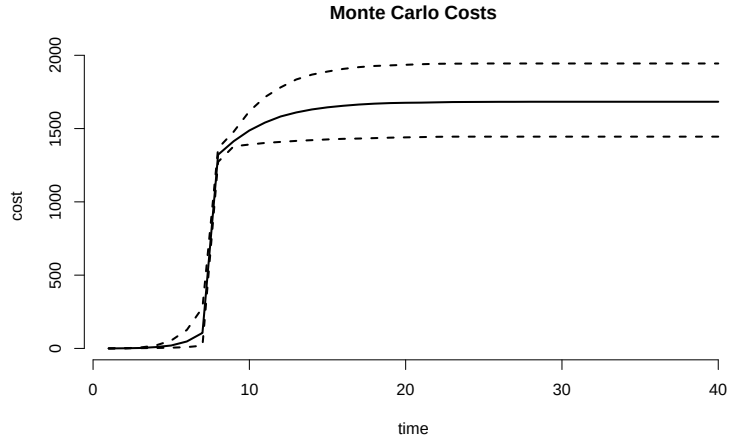


Figure 3: Monte Carlo simulated trajectories (top) for the numbers of susceptible, infected, and recovered individuals and cost(s) (bottom) trajectories under optimal (fixed) vaccination strategy.

is implemented. These costs then stabilize at a lower level than that observed under the no–vaccination strategy in Figure 1 (bottom). We can easily extract information on the distribution of final costs of the no–vaccination and optimal (fixed) vaccination policies for comparison as follows.

```
> T <- length(opt.MCepi$Median$C)
```

```
> optC <- getcost(opt.MCepi)
> initC <- getcost(init.MCepi)
> data.frame(rbind(initC, optC), row.names = c("init",
+     "opt"))

       q0.025     mean median  q0.975
init 2142.850 2297.106   2299 2437.05
opt  1444.925 1684.102   1683 1944.10
```

The optimal (fixed/static) strategy gives a (mean) savings of approximately 613.004 units, or 27%. The same information is available through the generic `print` and `summary` commands, e.g.,

```
> opt.MCepi

Call:
MCepi(init = init, params = true, vac = vac.opt, costs = costs)

Distribution of vaccinations administered:
    q0.025   mean median q0.975
vac    368 538.38    548    664

Distribution of final costs:
       q0.025     mean median q0.975
cost 1444.925 1684.102   1683 1944.1
```

Now we move on to adaptive management. In this case we assume that we do not have perfect information, so we will want to simultaneously estimate the epidemic parameters as well as find an optimal management strategy. The first, and perhaps most important, step here is to set up the function which is going to dictate the (true) evolution of the epidemic. This is done via `epistep`, which has dynamics as given in Eq. (4–6), earlier, as the default. However, the user can specify this function however they wish. We describe another `epistep` function in Section 4.

We also need to start with an initial guess, *i.e.*, priors, for the epidemic parameters that we want to estimate. The default option is to do this by choosing appropriate hyperparameters for the priors explained earlier. The defaults used by `manage` are those given in Section 2.2.

Here, we run the `manage` function with default values for hyperparameters and epistep function to adaptively design a vaccination strategy to manage the epidemic. At each time step in the evolution of the epidemic the `manage` function uses MCMC to sample from the posterior distribution of the parameters (b, k, ν, μ) given the available history of the epidemic and any already–implemented intervention. Then, a thinned subset of these samples are used propagate uncertainty in the parameter estimates through to the costs of the vaccination strategies. These costs are obtained by performing Monte Carlo forward simulations of the epidemic from the current time point into the future

with those parameters. As explained in Section 2.3 we choose the to implement the strategy that most frequently minimizes the cost. After the intervention is implemented, the state in the next time step is determined by `epistep`, and the process is repeated.

```
> out.man <- manage(init, epistep, vacgrid, costs)
```

To explore the results of the simulation we again plot the evolution of the epidemic (Figure 4, top) as well as the cost trajectory under the optimal vaccination (Figure 4, bottom), both for a single run. We can now compare the case with adaptive management to the case without vaccination (Figure 1, bottom), as well as the case of optimal vaccination with perfect information (Figure 3, bottom). As before, we can extract information on the final cost of the epidemic as follows.

```
> getcost(out.man)
```

```
[1] 1732
```

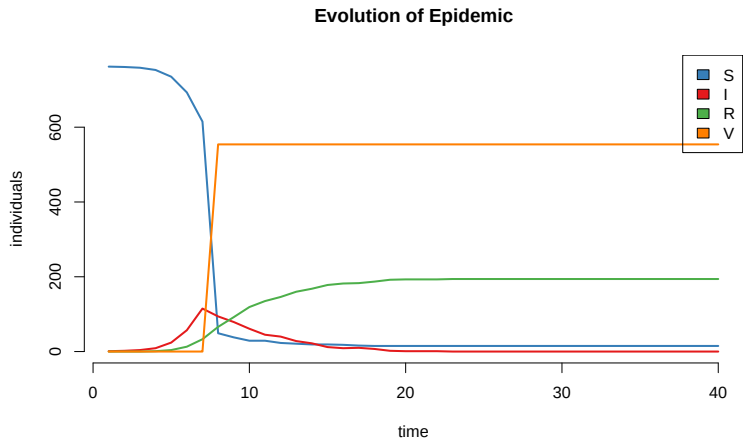
Since the `manage` function is not performing a Monte Carlo experiment, the number returned is a scalar. We can compare this to the cost of the best fixed strategy obtained with knowledge of the true underlying parameters governing the epidemic, given above, via a Monte Carlo experiment. Notice that the adaptive strategy is comparable to the best fixed strategy obtained when the true parameterization is known. Later, we shall perform a Monte Carlo version of the adaptive management strategy to make a more meaningful comparison.

We can also see the final distribution of the SIR model parameters. This is shown in Figure 5. A summary is provided by the generic `print` and `summary` commands:

```
> out.man
```

```
Call:
manage(init = init, epistep = epistep, vacgrid = vacgrid, costs = costs)

Distribution of SIR model parameters:
       b                  k                 nu
 Min.   :0.001853   Min.   : 0.2679   Min.   :0.2876
 1st Qu.:0.002233   1st Qu.: 1.5915   1st Qu.:0.3546
 Median :0.002326   Median : 3.5477   Median :0.3719
 Mean   :0.002334   Mean   : 4.1464   Mean   :0.3729
 3rd Qu.:0.002438   3rd Qu.: 5.3039   3rd Qu.:0.3902
 Max.   :0.002958   Max.   :20.9020   Max.   :0.4608
       mu
 Min.   :2.316e-06
 1st Qu.:6.773e-04
 Median :1.553e-03
 Mean   :2.232e-03
```

```
> plot(out.man)
```




```
> plot(out.man, type = "cost")
```

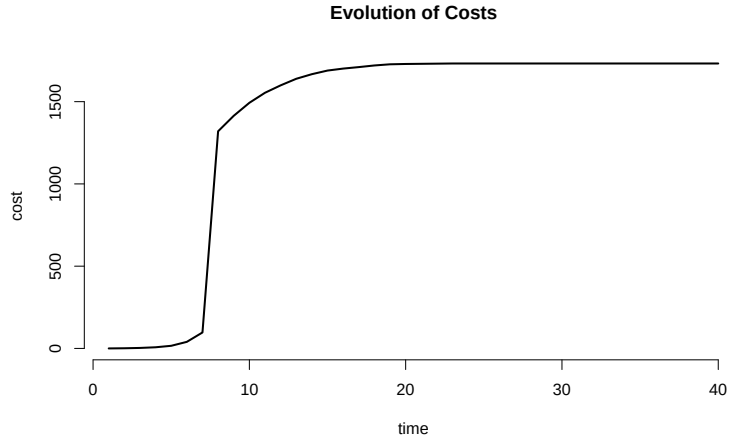


Figure 4: Trajectory (top) in terms of the numbers of susceptible, infected, recovered, and vaccinated individuals, and the corresponding cost (bottom) of the epidemic under adaptive management.

```
 3rd Qu.:3.095e-03
 Max.   :1.665e-02

Vaccinations administered and cost:
  vac cost
```

```
> true <- as.list(formals(epistep)$true)
> plot(out.man, type = "params", true = true)
```

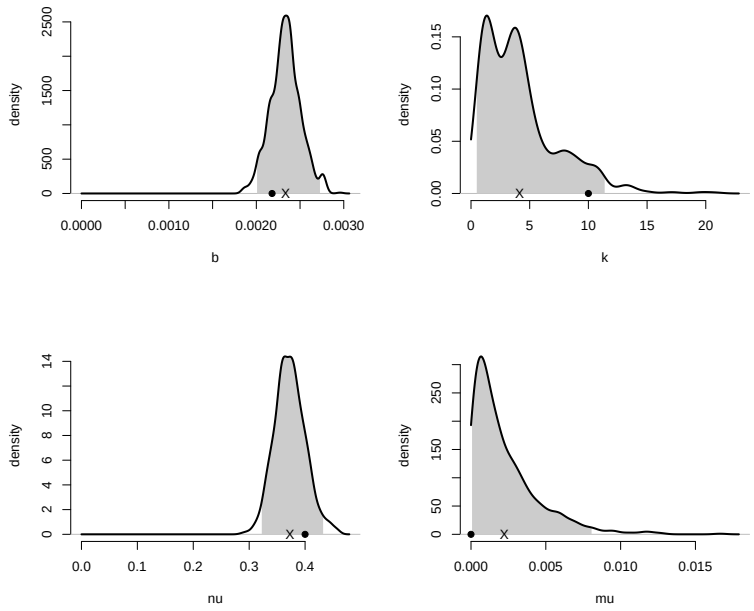


Figure 5: Final posterior distributions of estimated parameters (clockwise from the left: transmission rate, overdispersion parameter, mortality rate, recovery rate). "True" parameter values are indicated by a dot; mean posterior values are indicated by an 'x'; and the central 95% region of the distribution is shaded.

```
1 554 1732
```

Next, we illustrate a Monte Carlo experiment where epidemics are initialized and proceed randomly through the adaptive management strategy illustrated above so that we can see the average behavior, costs, and associated variability. The function `MCmanage` facilitates this experiment, and it essentially calls the `manage` function many times (which can be controlled by the `MCreps` argument). In order to provide a (relatively) quick demonstration we have set a low default of `MCreps = 30` and have used low defaults for the other Monte Carlo parameters to `management`.

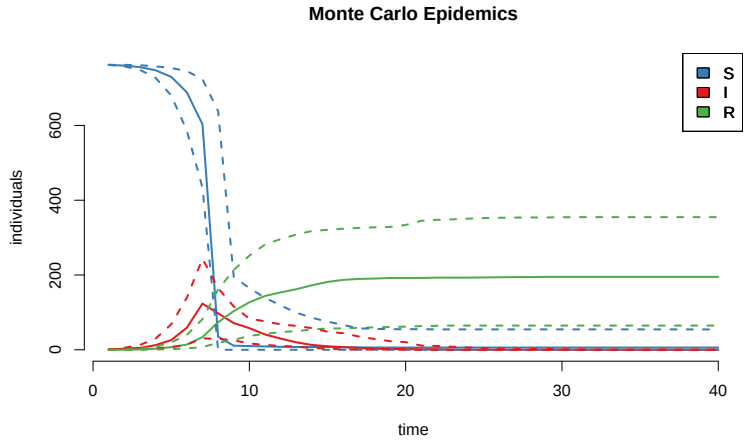
```
> out.MCmanage <- MCmanage(init, epistep, vacgrid,
+     costs)
```

To reproduce the results in Merl, et al. [20], use

```
MCvits = 100, MCMCpits = 10000, vacsamps = 100, MCreps = 100
```

and otherwise use the defaults. The object that is returned is of class `"MCepi"` with fields similar to those that are output from the `MCepi` function which implements a static (fixed) vaccination strategy. Thus, the same generic plot commands can be used. Figure 6 shows plots summarizing the distribution

```
> plot(out.MCmanage)
```




```
> plot(out.MCmanage, type = "costs")
```

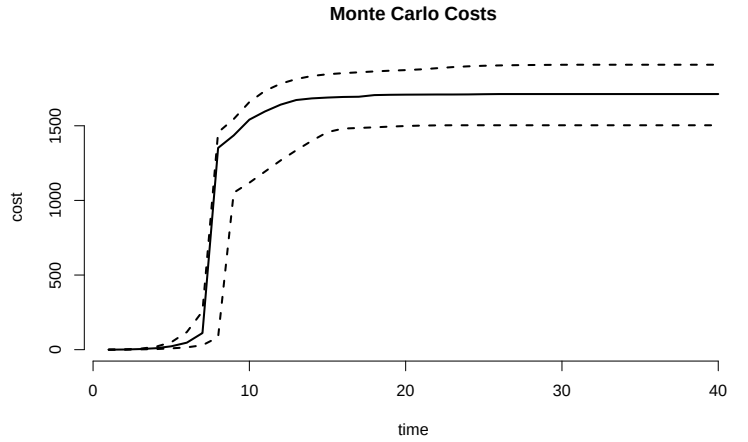


Figure 6: Monte Carlo simulated epidemic trajectories (top) for the numbers of susceptible, infected, and vaccinated individuals, and the associated cost(s) (bottom) under adaptive management

of epidemic trajectories (top) and costs (bottom) under the adaptive manage-

ment. Distributional information on number of vaccine units dispensed can be extracted as follows.

```
> getvac(out.MCmanage)
```

```
  q0.025  mean median q0.975
1 401.95 541.6    557  652.2
```

Figure 7 shows the distribution of fractions of individuals vaccinated and the stopping level for the epidemic trajectories under adaptive management in the Monte Carlo experiment. We can compare the costs (and quantile bounds) to that of the best fixed vaccination strategy calculated assuming that the true parameterization is known.

```
> cinit <- getcost(init.MCepi)
> copt <- getcost(opt.MCepi)
> cman <- getcost(out.MCmanage)
> data.frame(rbind(cinit, copt, cman), row.names = c("init",
+     "opt", "man"))
```

```
       q0.025     mean median   q0.975
init 2142.850 2297.106 2299.0 2437.050
opt  1444.925 1684.102 1683.0 1944.100
man  1503.450 1721.867 1712.5 1908.975
```

The cost of the non-intervention strategy is shown again for calibration purposes. As before, the generic `print` and `summary` commands can be used to obtain the same information. Notice that, in this short Monte Carlo experiment, the adaptive management scheme approximates the best static vaccination scheme obtained then the true parameterization is known in advance.

It is interesting to compare to what the cost of managing the epidemic would have been *without* estimating the parameters as the epidemic progressed, but rather by guessing what the appropriate parameters might be. Suppose our best guess at the parameters underestimated the true transmission probability b and overestimated the true recovery probability ν.

```
> bad <- list(b = 0.001, k = 10, nu = 0.9, mu = 0)
```

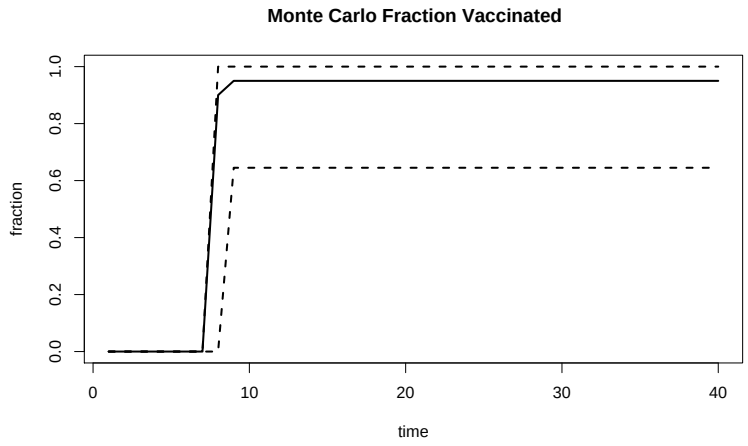
The optimal (static) policy under this parameterization can be constructed, as demonstrated above.

```
> costs.bad <- optvac(init, bad, vacgrid, costs)
> pol.bad <- getpolicy(costs.bad)
> pol.bad
```

```
     row col frac stop cost
best   2  10  0.1  677  360
```

Then, we can calculate the distribution of costs of managing the true epidemic with a policy developed under our best guess of the parameterization.

```
> plot(out.MCmanage, type = "fracs")
```




```
> plot(out.MCmanage, type = "stops")
```

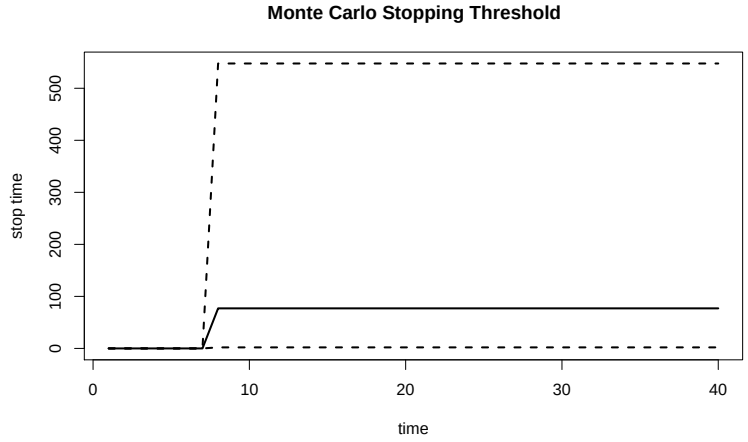


Figure 7: Fractions vaccinated (top) and stopping levels (bottom) for the simulated epidemic trajectories shown in Figure 6 that were under adaptive management. In both cases, the $(2.5, 50, 97.5)$-% quantiles are shown.

```
> bad.MCepi <- MCepi(init, true, pol.bad[3:4], costs)
> cbad <- getcost(bad.MCepi)
> cbad

  q0.025     mean median  q0.975
1 2129.9 2282.289   2284 2422.05
```

Comparing these costs with the ones obtained above, under adaptive management, we can see that a poor guess can lead to a significantly worse strategy—nearly 57% larger on average. Clearly, good quality (online) estimates of the SIR model are crucial to ensuring a cost–effective approach to the management of an epidemic.

4 Alternative Transmission Model

For reproducibility of the results in this section we have, again, set the random seed as follows.

```
> seed <- 12345
> set.seed(seed)
```

We would like to see how well the fixed and adaptive strategies can do when faced not only with parameter mis-specification, but when there may be some component of the underlying transmission model that is not accounted for in the implemented SIR model upon which its vaccination strategies are based. Towards this end, we chose a fairly simple extension of the SIR model where infection does not pass directly from individual to individual. Instead, they become infected by encountering a reservoir of the infectious agent (for instance bacteria or fungi in water or soil), and then infected individuals contribute infectious agents to the reservoir. The continuous time dynamics are:

$$\begin{aligned}
\frac{\mathrm{d}S}{\mathrm{d}t} &= -aS\left(\frac{C}{C+C_0}\right) \\
\frac{\mathrm{d}I}{\mathrm{d}t} &= aS\left(\frac{C}{C+C_0}\right) - (\nu+\mu)I \\
\frac{\mathrm{d}R}{\mathrm{d}t} &= \nu I \\
\frac{\mathrm{d}D}{\mathrm{d}t} &= \mu I \\
\frac{\mathrm{d}C}{\mathrm{d}t} &= \rho I - mC
\end{aligned} \tag{13}$$

where the concentration of the infective agent in the reservoir is given by C; the transmission rate is modeled by a saturating function of C, so that as $C \to \infty$ the transmission rate approaches the constant a at a rate determined by C_0; infective agents die or are removed from the reservoir at rate mC; and the *per capita* rate at which new infected agents are added to the reservoir is ρ.

This system can be discretized in a similar manner to the system discussed earlier, so that the single step transmission dynamics are given by:

$$\tilde{I}|s, C \sim \mathrm{Bin}(s, p_i(a, C, C_0)), \tag{14}$$

where

$$p_i(a, C, C_0) = 1 - exp\left(\frac{-aC}{C + C_0}\right) \tag{15}$$

and $\text{Bin}(n, \pi)$ is the standard binomial distribution. The dynamics for $\tilde{R}$ and $\tilde{D}$ are exactly as before (Eqs. (5–6)). The single step dynamics for the reservoir are expected to be more smooth than the epidemic in the population at large. For instance, for bacteria the concentration could be in the thousands or millions of individuals. In this case the evolution of the reservoir is approximately

$$C(t + 1) - C(t) = [\rho I(t)] - d_c. \tag{16}$$

Here d_c is the number (per unit reservoir) of infectious agents (stochastically) removed from the reservoir in a unit time, which has distribution

$$d_c \sim \text{Bin}(C, p_{d_c}), \tag{17}$$

where $p_{d_z} = 1 - e^{-m}$.

In R, we may implement the above transition model by encoding it in an alternative `epistep` function for use with the `amei` package methods as follows.

```
> alt.epistep <-
+ function(SIR, last=list(rem=0, rec=0, infect=0, dead=0, Z=0),
+          true=list(a = 0.05, mu = 0.05, nu = 0.1, m = 0.4,
+          rho = 200, C = 500))
+ {
+   ## calculate the infection probability based on the
+   ## resevoir, and randomly infect susceptibles
+   Z <- last$Z
+   fz <- Z/(Z+true$C)
+   pi <- 1 - exp(-true$a * fz)
+   infect <- rbinom(1, SIR$S, pi)
+
+   ## update recovereds and deaths
+   pr <- 1 - exp(-true$nu)
+   rec <- rbinom(1,SIR$I,pr)
+   pd <- 1 - exp(-true$mu)
+   dead <- rbinom(1, SIR$I-rec, pd)
+
+   ## resevoir dynamics
+   pz <- 1 - exp(-true$m)
+   dz <- rbinom(1, Z, pz)
+   bz <- round(SIR$I*true$rho)
+   Z <- Z - dz + bz
+
+   ## the returned list is passed in as "last" in a
```

```
+   ## subsequent call to this "epistep" function
+   return(list(rem=(rec+dead), rec=rec, infect=infect,
+               dead=dead, Z=Z))
+ }
```

Here we first use the `manage` function with a `NULL` vaccination strategy (and also, optionally, a `NULL` cost structure), in order to see how the behavior of this system compares to the default `epistep` model implemented within `amei`. We can then also look at the estimated "effective" SIR parameters.

This system can resemble the standard SIR model, especially when m is large, so infective agents do not remain in the reservoir for long.

```
> init1 <- list(S0 = 150, I0 = 1, R0 = 0, D0 = 0)
> true <- list(a = 0.065, mu = 0, nu = 0.3, m = 0.99,
+     rho = 500, C = 500)
> alt.epistep1 <- alt.epistep
> formals(alt.epistep1)$true <- true
> out.alt <- manage(init1, alt.epistep1, NULL, NULL,
+     T = 80)
```

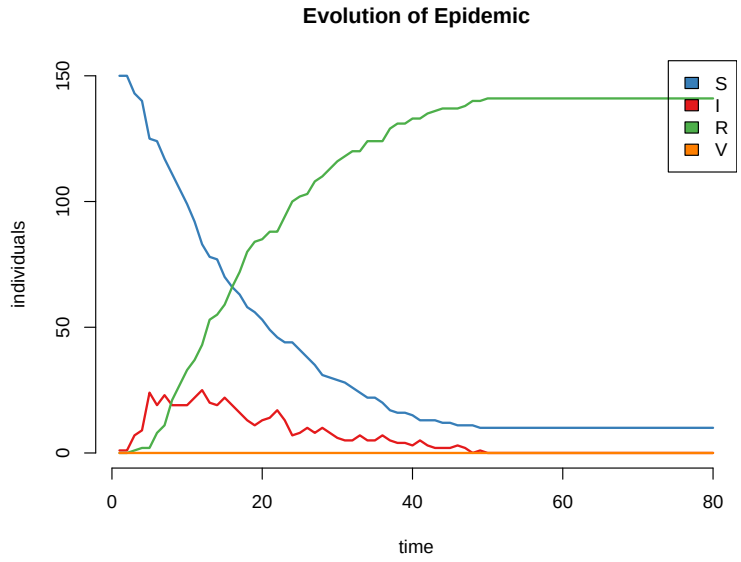
The *top* of Figure 8 shows the resulting dynamics under the default parameterization offered by the formals of the `alt.epistep` function. To illustrate how this new system can (significantly) differ from the SIR dynamics consider the case where we take m very to be small. In this case is possible for new infections to occur even if there had been no infected individuals in a previous time step since they infective agent may persist in the reservoir for a long time without infected individuals being present.

```
> true <- list(a = 0.065, mu = 0, nu = 0.3, m = 0.005,
+     rho = 500, C = 500)
> alt.epistep2 <- alt.epistep
> formals(alt.epistep2)$true <- true
> out.alt2 <- manage(init1, alt.epistep2, NULL,
+     NULL, T = 80)
```

The effect of these new dynamics may be readily seen in Figure 8, *bottom*, as the epidemic progresses. In this second case (with small m), it is also very unlikely that any susceptible individuals will be left at the end of an epidemic. This contrasts with both the stochastic SIR model and these alternative dynamics with m large, as it is possible that, due to stochastic effects, the infection will die out before all susceptibles have been exposed.

It is interesting to examine the posterior parameter distributions for these two cases (Figure 9) if we assume that these underlying alternative dynamics are well approximated by the simpler SIR model. As one might expect, the estimates of the recovery and death rates in both cases are quite similar, especially since these dynamics are the same as those implemented within `amei`. However, the estimates of b and k are quite different in the two cases. The 95% credible

```
> plot(out.alt)
```




```
> plot(out.alt2)
```

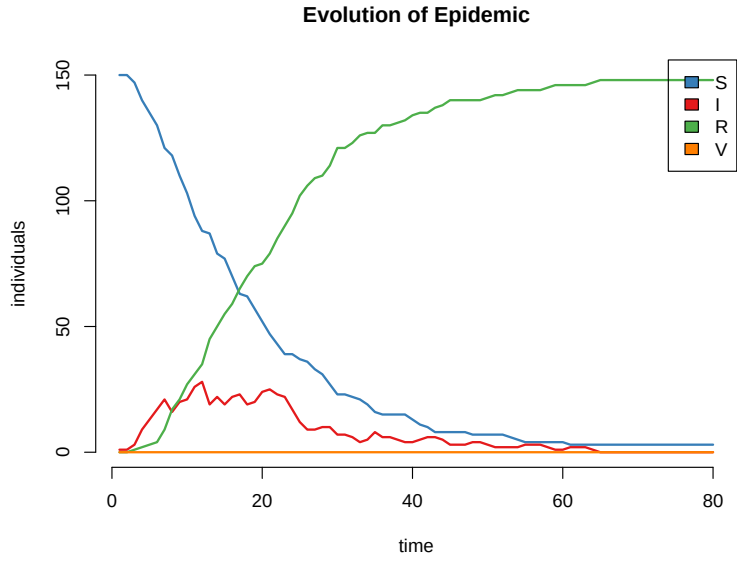
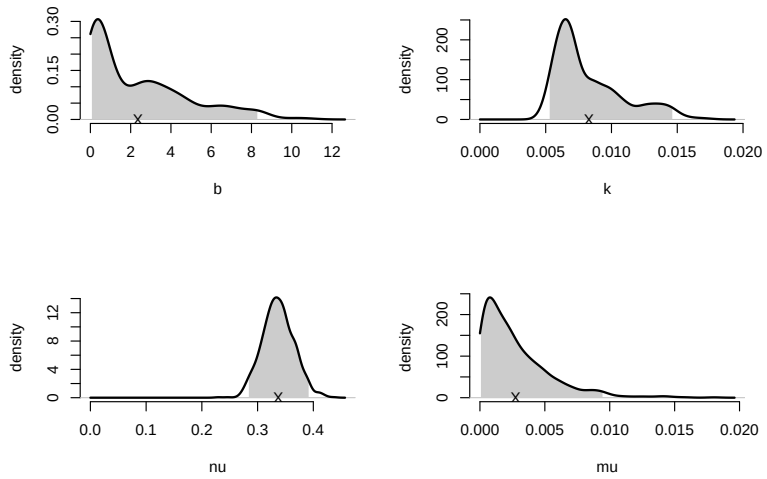


Figure 8: Trajectories in terms of the numbers of susceptibles, infecteds, recovereds, and dead individuals under the alternative model in Eq. (13) with parameters $\{(a = 0.065, mu = 0.0, nu = 0.3, rho = 500, C = 500\}$ and with (top) $m = 0.99$ (bottom) $m = 0.005$.

intervals hardly overlap, indicating that the two cases result in dynamics that

```
> plot(out.alt, type = "params", showd = TRUE)
```




```
> plot(out.alt2, type = "params", showd = TRUE)
```

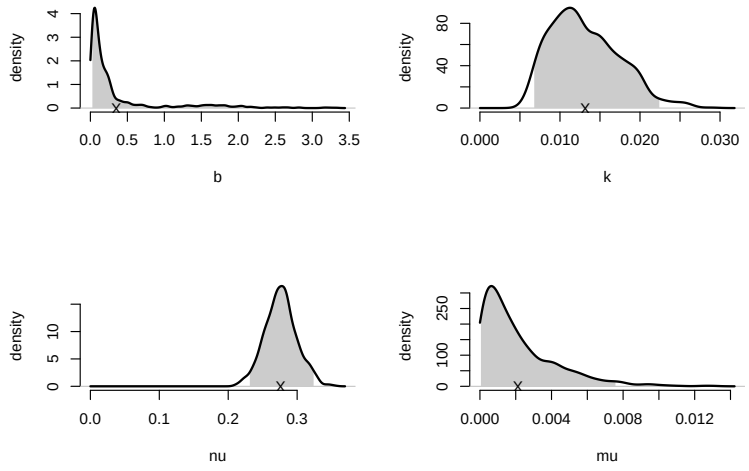


Figure 9: Posterior parameter distributions for b, k, μ, and ν when the simulated source epidemic follows the alternative transmission model with parameters $\{(a = 0.065, mu = 0.0, nu = 0.3, rho = 500, C = 500\}$ and with (top) $m = 0.99$ (bottom) $m = 0.005$.

are quantitatively different.

We now move on to our task of examining how well the fixed and adaptive strategies can do when faced with an epidemic which is evolving according to a transmission function outside the class of SIR models (1–3) used to calculate the (optimal) vaccination strategy. The first step here is to let the fixed strategy

"cheat". That is, we allow the fixed strategy to see a full epidemic spread according to the true model and parameterization and without intervention, so as to estimate the best SIR model approximation. To do this we use the `manage` function with `NULL` vaccination strategy and a `NULL` cost structure, and default "true" parameters.

```
> init <- list(S0 = 600, I0 = 1, R0 = 0, D0 = 0)
> time = 80
> posterior <- manage(init, alt.epistep, NULL, NULL,
+     T = time, bkrate = 100)
```

Figure 10 shows trace plots of the samples obtained. From the output we can extract the mean parameterization obtained at the final time point of the epidemic.

```
> mean.params <- as.list(apply(posterior$samp, 2,
+     mean))
```

Based on these parameters, and thus assuming an SIR model, we can calculate the optimal static vaccination policy.

```
> costs <- list(vac = 2, death = 4, infect = 1)
> vacgrid <- list(fracs = seq(0, 1, 0.1), stops = seq(2,
+     init$S0 - 50, 50))
> alt.optvac <- optvac(init, mean.params, vacgrid,
+     costs, T = time)
> alt.best <- getpolicy(alt.optvac)
```

The cost surface is shown in Figure 11.

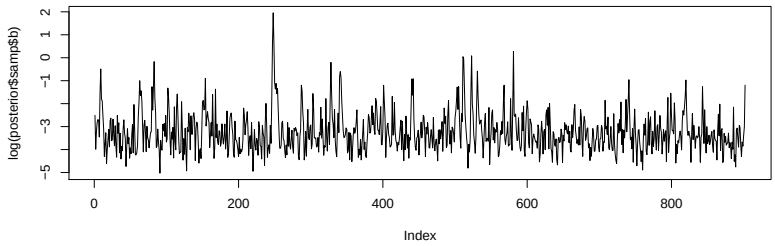
Then, based on the extracted policy, we can use `MCepi` to simulate the true evolution of the epidemic (via `alt.epistep`) many times in order to build up an understanding of the distribution of costs of the optimal static policy calculated under the simplified SIR parametrization.

```
> alt.vac.opt <- alt.best[3:4]
> alt.MCepi <- MCepi(init, alt.epistep, alt.vac.opt,
+     costs, T = time)

> getcost(alt.MCepi)

    q0.025     mean median  q0.975
1 1594.975 1773.912   1773 1972.05
```

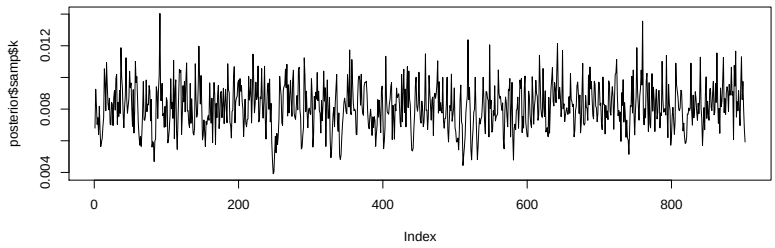
Now for the comparison. We provide the `alt.epistep` function to `MCmanage` in order to build up an understanding of how the optimal adaptive strategy fares under a restricted SIR model. Here we use the default, uninformative priors.

```
> alt.MCmanage <- MCmanage(init, alt.epistep, vacgrid,
+     costs, T = time)
```

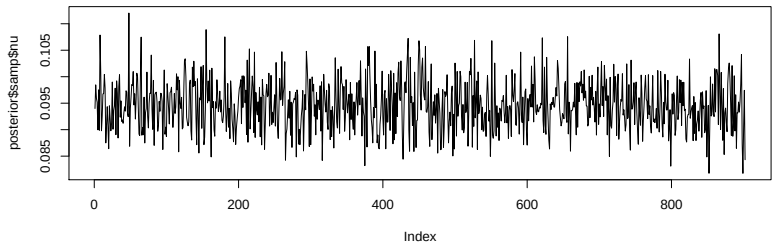
```
> plot(log(posterior$samp$b), type = "l", main = "")
```




```
> plot(posterior$samp$k, type = "l", main = "")
```




```
> plot(posterior$samp$nu, type = "l", main = "")
```




```
> plot(posterior$samp$mu, type = "l", main = "")
```

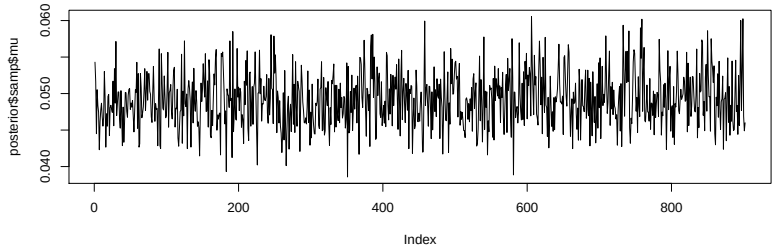


Figure 10: Traces of samples from the posterior distribution of the paramters for the (misspecified) SIR model.

```
> plot(alt.optvac)
```

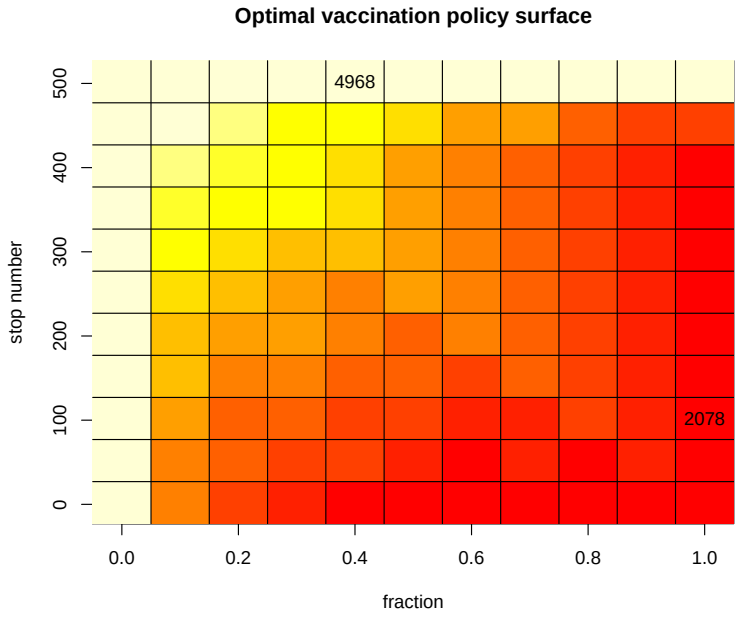


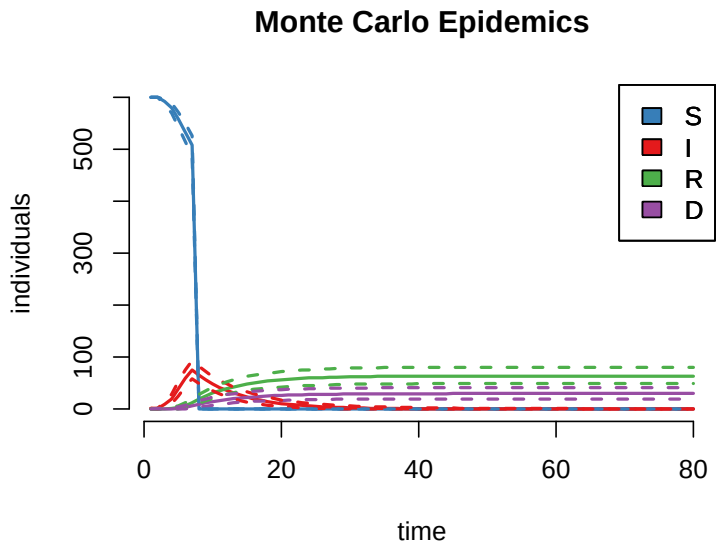
Figure 11: The cost surface calculated for fixed strategies using parameters estimated from a single run of the alternative epidemic model.

```
> getcost(alt.MCmanage)

   q0.025     mean median   q0.975
1 1645.45 1758.933 1731.5 1962.425
```

An alternative (and possibly fairer) approach would be to allow the adaptive strategy to cheat as well by choosing the prior to be tightly concentrated around around the `mean.params` estimated above. We can see by looking at the full trajectories in Figures 12 and 13 that the strategies implemented by both the fixed and adaptive methods are very similar. On average, the fixed strategy, which has the benefit of having seen a complete epidemic in order to estimate the "real" parameters, does a bit better, as we would expect. The adaptive strategy can in some cases, do better than the fixed strategy, as can be seen by looking at the lowest quantile of the expected costs. Indeed, the comparison shows on average the adaptive strategy has lower costs. However, the adaptive strategy here has a wider interval for the final costs, due to the uncertainty in the parameters. Much of this uncertainty could be reduced by, for instance, providing the adaptive algorithm with informative priors, as mentioned above.

We can also look back to Figure 11 to see the kind of improvement that the optimal strategy can have over the case when no one (or very few individuals) are vaccinated. We can extract the expected cost under this worst fixed strategy

```
> plot(alt.MCepi, showd = TRUE)
```




```
> plot(alt.MCepi, type = "costs")
```

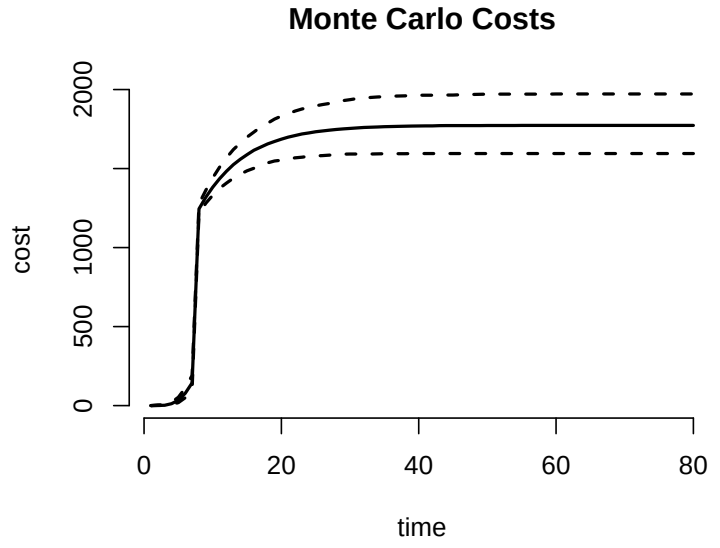
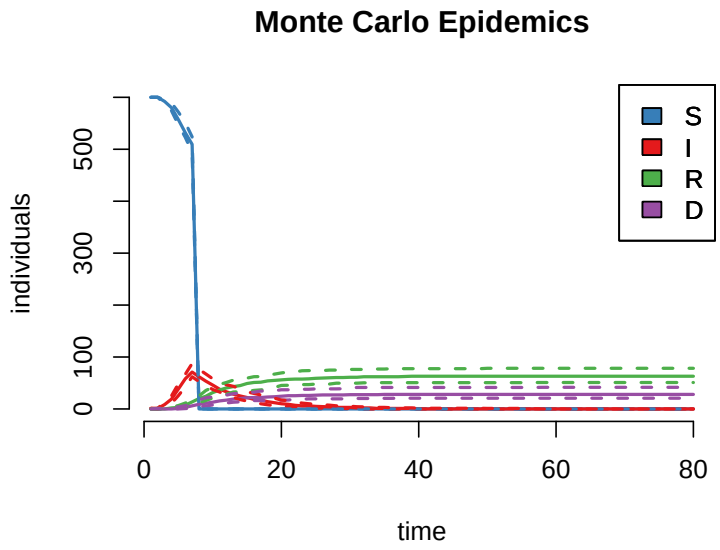


Figure 12: Trajectories in terms of *top* the numbers of susceptibles, infecteds, recovereds, and dead individuals under the alternative model in Eq. (13) and *bottom* costs under the default parameterization under the optimal static vaccination policy.

and compare it to the best fixed strategy directly.

```
> plot(alt.MCmanage, showd = TRUE)
```




```
> plot(alt.MCmanage, type = "costs")
```

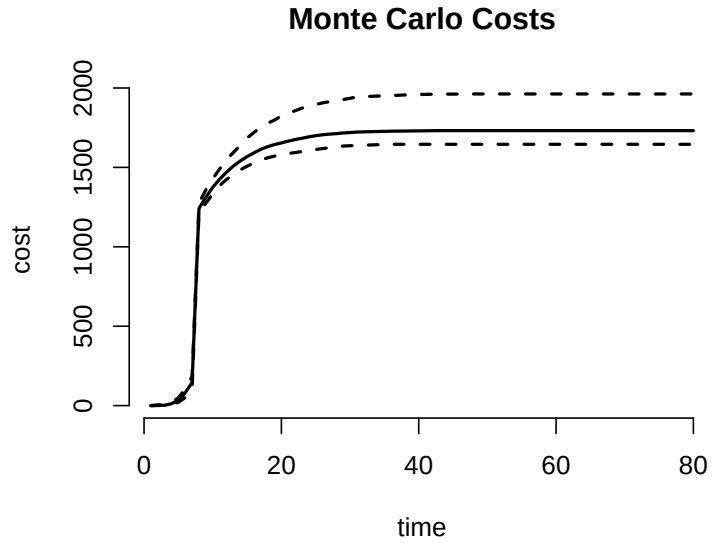


Figure 13: Trajectories in terms of *top* the numbers of susceptibles, infecteds, recovereds, and dead individuals under the alternative model in Eq. (13) and *bottom* costs under the default parameterization under the adaptive vaccination policy.

```
> alt.worst <- getpolicy(alt.optvac, which = "worst")
> rbind(alt.best, alt.worst)
```

```
      row col frac stop cost
best   11   3  1.0  102 2078
worst   5  11  0.4  502 4968
```

Here it is easy to see how much better using an optimal fixed strategy with good parameter settings is compared to not vaccinating. Since the adaptive strategy has comparable costs to the fixed strategy over many trials of the epidemic, we can conclude that the adaptive strategy gives similar reductions in costs over the null vaccination policy as the fixed policy, even though the adaptive algorithm must begin planning and implementing the intervention with considerably less information than the fixed strategy it is compared with here.

We have shown here that it is straightforward to build functions with fairly simple extensions to the SIR model to use with the `amei` package. Using one example, we can see that the model implemented within `amei` to plan the interventions is flexible enough to allow effective vaccination strategies, even when this model does not match the one which is known to govern the true underlying epidemic dynamics. We have also shown how one can estimate "effective" SIR parameters from an alternative dynamic epidemic model. These estimates can be used to build a fixed policy, or the posterior samples could even be used as parameters to enhance the adaptive strategies.

5 Discussion and related work

Our `amei` package for R implements a statistical framework that enables concurrent estimation of epidemic parameters and optimal intervention strategies. In particular, it allows parameter uncertainty to be taken into consideration when planning an intervention.

In the current implementation, we look for adaptive strategies of only one type—vaccination of proportions of the population until susceptibles fall below some threshold. However, if we were to instead allow the fraction of the population targeted for vaccination to be a function of future disease states we could regard Eq. 12 as a stochastic iteration equation and use stochastic dynamic programming [7] to calculate the optimal intervention associated with a set of parameter values. Such an approach may be useful for situations in which knowledge of the disease state is available, but for whatever reason sequential inference is not possible. In the situation considered here, in which the static strategy is sequentially updated based on the current disease state and parameter estimates, the adaptive strategy that emerges is flexible in that it consists of a state–dependent sequence of target fractions, but does not involve the additional computational burden associated with stochastic dynamic programming.

If there is not much data observed during an epidemic (for instance if the epidemic is occurring in a small population) then estimating disease parameters within any framework could, of course, be difficult. One could use data about previous epidemics to estimate parameters, for instance by maximum likelihood

methods, to plan an intervention. However, as we showed in Section 3, should these estimates be poor then the cost of intervening in the epidemic can be greatly increased. By instead using previous epidemic data to inform the priors within a Bayesian framework, such as the one we describe here, we can take into account other research and observations without risking the high costs inherent in using incorrect parameter estimates for planning an intervention.

Recently other authors (e.g., [8]) have also proposed using Bayesian methods to estimate disease parameters and propagate uncertainty in parameter estimates through to an optimal of vaccination strategy. For instance, Elderd, et al. [8], considered a Susceptible–Exposed–Infectious–Recovered (SEIR) model for a small pox epidemic with mass action infection dynamics, and either mass or trace vaccination. However, in their approach the likelihood requires numerical solutions of the system of DEs (which can be computationally intensive). Ball, et al. [3, 4] have also written technical papers on the optimal vaccination strategies in epidemics, but do not consider adapting the policy over time.

Compared to these approaches, our proposed method has several advantages. Our negative binomial discreetization of the SIR model is much less computationally intensive, and could easily be expanded to include an exposed class. Additionally, this model is comparitively simple since it connects intuitively with a straightforward SIR model, and is thus likely to be much more approachable to policy makers and practicioners outside of of statistics. Our approach is also adaptive, allows a much more flexible cost framework, and is generalizable to other types of interventions. Since the intervention strategy implemented by `amei` is iteratively updated upon the arrival of new data, the vaccination schedule is inherently adaptive to the state of the epidemic. This allows significant changes in the vaccination strategy mid-intervention. This may be helpful, for example, in a scenario where an intervention may be discontinued during a lull, a subsequent refinement of parameter estimates or a surge of new infections may dictate that the vaccination campaign should be re-initiated.

Not only can more comlicated infection dynamics be used for simulating the dynamics of a disease, but in addition the method implemented in `amei` can be modified to include more complicated disease dynamics such as latent states or vector-communicated diseases, as well as more complicated intervention strategies that allow combinations of vaccination, quarantine, and culling. Also note the possibility of calculating policies based on minimization of some quantile of the realized cost rather than the mean cost. This would lead to minimization of worst-case-scenarios, which may be useful in practice.

There are other R packages that involve the simulation of epidemics, and inference for models like SIR. As far as we can tell, ours is the only one which considers (on–line) adaptive management of interventions. However, two packages that take approaches similar to ours for inferring the parameters governing an epidemic are worth mentioning here. `stochasticGEM` [31] provides Bayesian inference for partially observed stochastic epidemics. The implementation in this package also allows for estimating parameters governing the infectious and incubation period length. Several variants of the general epidemic model are considered (e.g., [12, 16, 24, 23, 22, 28]), including the stochastic SIR with

Markovian and non–Markovian infectious periods, and SEIR models. As with `amei`, estimation is via MCMC. Höhle, et al., whose methods are implemented in the `stochasticGEM` package, have themselves released an R package called `RLadyBug` [15]. In this package, maximum likelihood and Bayesian inference can be performed to estimate the parameters and provide confidence/credible intervals and the ability to test hypotheses. Unfortunately, the package requires JAVA. However, they do provide a nice paper outlining the usage [17] with examples. Both packages contain interesting data sets on epidemics and visualization tools.

There are several other R packages that allow for the manipulation and analysis of epidemic and disease data. They include a few packages which are predominantly used for teaching purposes, e.g.: `Epi` [5] with methods for multiscale and censored data; `epiR` [30] focusing on veterinary epidemiology, and `epibasix` [27]. Two other packages that include a range of statistical functionality for data manipulation and inference are `epicalc`[6] and `epitools` [2].

Acknowledgments

This work was partially supported by NSF grant DMS 0310542 to MM, and EPSRC grant #EP/D065704/1 to RBG.

A Implementation details

At a high level, most of the functions and routines in the `amei` package are written in R. However, most the sub–routines implementing the Monte Carlo evaluation of costs for the various vaccination strategies, which are obtained by repeatedly simulating the epidemic forward in time given the current samples of the parameter estimates, are written in C for speed considerations. There is significant scope for parallelizations of these Monte Carlo routines, since each forward simulation is independent of the next. For this reason future versions of this package may leverage `Pthreads` or `MPI` to obtain significant speedups.

References

[1] R. M. Anderson and R. M. May. *Infectious Diseases of Humans: Dynamics and Control*. Oxford University Press, 1991.

[2] Tomas Aragon. `epitools`: *Epidemiology Tools*, 2008. R package version 0.5-2.

[3] F.G. Ball and O. Lyne. Optimal vaccination policies for stochastic epidemics among a population of households. *Mathematical Biosciences*, 177 & 178:333–354, 2002.

[4] F.G. Ball, P.D. O'Neill, and J. Pike. Stochastic epidemic models in structured populations featuring dynamic vaccination and isolation. *Journal of Applied Probability*, 44:571–585, 2007.

[5] Bendix Carstensen, Martyn Plummer, Esa Laara, and Michael Hills et al. `Epi`*: A package for statistical analysis in epidemiology.*, 2008. R package version 1.0.8.

[6] Virasakdi Chongsuvivatwong. `epicalc`*: Epidemiological calculator*, 2008. R package version 2.7.2.3.

[7] C. W. Clark and M. Mangel. *Dynamic State Variable Models in Ecology: Methods and Applications*. Oxford University Press, 2000.

[8] B. Elderd, Dukic V., and Dwyer G. Uncertainty in prediction of disease spread and public health responses to bioterrorism and emerging disease. *Proceedings of the National Academy of Sciences (PNAS) USA*, 103:15693–15697, 2006.

[9] M. Enserink. Rapid response could have curbed foot-and-mouth epidemic. *Science*, 294:26–27, 2001.

[10] N.M. Ferguson, M.J. Keeling, W.J. Edmunds, R. Gani, B.T. Grenfell, R.M. Anderson, and S. Leach. Planning for smallpox outbreaks. *Nature*, 425:681–685, 2003.

[11] D. Gamerman and H. Lopes. *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. Chapman & Hall/CRC, 2006.

[12] G.J. Gibson and E. Renshaw. Estimating the parameters in stochastic epidemic models using Markov chain models. *IMA Journal of Mathematics Applied in Medicine and Biology*, 16:19–40, 1998.

[13] M.E. Halloran, I.M. Longini, and Y. Nizam, A.and Yang. Containing bioterrorist smallpox. *Science*, 298:1428–1432, 2002.

[14] H. W. Hetchote. The mathematics of infectious diseases. *SIAM Review*, 42:599–653, 2000.

[15] Michael Hoehle and Ulrike Feldmann. `RLadyBug`*: Analysis of infectious diseases using stochastic epidemic models*, 2008. R package version 0.4-3.

[16] M. Höhle, E. Jorgensen, and P.D. O'Neill. Inference in disease transmission experiments by using stochastic epidemic models. *Journal of the Royal Statistical Society Series C*, 54:349–366, 2005.

[17] Michael Höhle and Ulrike Feldmann. `RLadyBug`–An r package for stochastic epidemic models. *Computational Statistics and Data Analysis*, 52(2):680–686, 2007.

[18] J.O. Lloyd-Smith, A.P. Galvani, and W. M. Getz. Curtailling transmission of severe acute respiratory syndrome within a community and its hospital. *Proceedings of the Royal Society of London*, 270:1979–1989, 2003.

[19] H. McCallum, N. Barlow, , and J. Hone. How should pathogen transmission be modelled. *Trends in Ecology and Evolution*, 16:295–300, 2001.

[20] Daniel Merl, Leah R. Johnson, Robert B. Gramacy, and Marc S. Mangel. A statistical framework for the adaptive management of epidemiological interventions. Technical Report 08–29, Duke University Department of Statistical Science, 2008.

[21] J. D. Murray. *Mathematical Biology I: An Introduction*. Springer Verlag, 2002.

[22] P.D. O'Neill. A tutorial introduction to bayesian inference for stochastic epidemic models using Markov chain Monte Carlo method. *Mathematical Biosciences*, 180:103–114, 2002.

[23] P.D. O'Neill and N.G. Becker. Inference for an epidemic when susceptibility varies. *Biostatistics*, 2(99–108), 2001.

[24] P.D. O'Neill and G.O. Roberts. Bayesian inference for partially observed stochastic epidemics. *Journal of the Royal Statistical Society, Series A*, 162:121–129, 1999.

[25] R. Patel, I.M. Longini, and M.E. Halloran. Finding optimal vaccination strategies for pandemic influenza using genetic algorithms. *Journal of Theoretical Biolog*, 234:201–212, 2005.

[26] R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2008. ISBN 3-900051-07-0.

[27] Michael A Rotondi. `epibasix`*: Elementary Epidemiological Functions for a Graduate Epidemiology/Biostatistics Course*, 2007. R package version 0.1.

[28] G. Streftaris and G.J. Gibson. Bayesian inference for stochastic epidemics in closed populations. *Statistical Modelling*, 4:63–65, 2004.

[29] M.J. Tildesley, N.J. Savill, D.J. Shaw, R. Deardon, S.P. Brooks, M.E.J. Woolhouse, B.T Grenfell, and M.J. Keeling. Optimal reactive vaccination strategies for a foot-and-mouth outbreak in the uk. *Nature*, 440:83–86, 2006.

[30] Mark Stevenson with contributions from Telmo Nunes, Javier Sanchez, and Ron Thornton. `epiR`*: Functions for analysing epidemiological data*, 2008. R package version 0.9-11.

[31] Eugene Zwane. `stochasticGEM`*:* R *Package for Fitting Stochastic General Epidemic Models*, 2007. R package version 0.0-1.