STATISTICS IN MEDICINE
Statist. Med. 2009; **volume**:1–39 *Prepared using* simauth.cls *[Version: 2002/09/18 v1.11]*

A Bayesian approach for estimating bioterror attacks from patient data

J. Ray[1]*, Y. M. Marzouk[1], P. T. Boggs[1], D. M. Gay[2], and H. N. Najm[1]

[1] *Sandia National Laboratories, PO Box 969, Livermore, CA, 94550-0969*
[2] *Sandia National Laboratories, PO Box 5800, Albuquerque, NM, 87185-1318*

SUMMARY

Terrorist attacks using an aerosolized pathogen have gained credibility as a national security concern after the anthrax attacks of 2001. Characterizing the attack quickly—by estimating the number of people infected, the time of infection, and a representative dose received—is crucial to planning a medical response. We use a Bayesian approach, based on a short time series of diagnosed patients, to estimate a joint probability density for these parameters. We first test the formulation with idealized cases and then apply it to realistic scenarios, including the Sverdlovsk anthrax outbreak of 1979. We also use simulated outbreaks to explore the impact of model error, as when the model used for generating simulated epidemic curves does not match the model subsequently used to characterize the attack. We find that in all cases except for the smallest attacks (fewer than 100 infected people), 3–5 days of data are sufficient to characterize the outbreak to a specificity that is useful for directing an emergency response. We conclude by demonstrating how a probabilistic characterization can be used to design an allocation schedule for scarce resources in the aftermath of an attack. Copyright © 2009 John Wiley & Sons, Ltd.

Keywords: Bayesian inference, anthrax, Sverdlovsk outbreak, bioterrorism, attack characterization, resource allocation, optimization under uncertainty

1. Introduction

The anthrax attacks of 2001 [1] raised the credibility of aerosolized pathogens being used in a bioterror attack. Early warning, in the form of an anomalous increase in syndromes detected by public health monitoring networks [2] or alternatively via detection by environmental sensors, holds the highest potential for reducing casualties. However, syndromic surveillance can only provide heightened awareness; it results neither in definitive evidence of an attack nor in identification of the pathogen. Further, environmental sensors may not always capture the introduction of an aerosolized pathogen into a population; examples include small releases that do not travel far, coarse particulate formulations that precipitate easily, and releases in areas which are not well instrumented. In these cases, the first intimation of an attack (i.e., identification of the pathogen, confirmation that the outbreak was not being caused by natural causes, etc.) will follow the diagnosis of the first few patients, but by then the disease may have established itself in the population. Characterizing the release (henceforth referred to

*Correspondence to: MS 9159, PO Box 969, Sandia National Laboratories, Livermore, CA, 94550-0969. Email: jairay@somnet.sandia.gov

Contract/grant sponsor: US Dept of Energy; contract/grant number: DE-AC04-94-AL85000

as the bioterror, or BT, attack) by computing the number N of people infected, the time τ of infection, and a representative dose D received by the infected individuals has important ramifications in planning a response, as underscored by the "Dark Winter" exercise [3]. The inferred characteristics can serve as initial conditions for ensembles of predictive epidemic simulations that guide the optimal allocation of medical resources under uncertainty.

Inferring the characteristics of an outbreak can be challenging. Data for inference consist of the time each diagnosed patient became symptomatic, known to within a finite time interval, and perhaps the location of each patient's residence and place of work. The time at which each patient's symptoms first appear is related to the genesis of the outbreak via the incubation period, typically modeled as a random variable whose probability distribution may be dependent on the dose received. For relevance to response planning and consequence management, inferences should be drawn early in the outbreak—from a 3–5 day time series of patient data, for instance. We also note that the incubation period distribution used for inference may be a poor model for the particular instance of the disease. This mismatch, and the paucity of data in a short observational period, suggest that the inferred characteristics will be rather approximate and that quantifying uncertainty in the characterization will be a key requirement of the inference process.

Few studies have used statistical methods to characterize the genesis of a partially observed epidemic. Walden & Kaplan [4] introduced a Bayesian formulation for estimating the size and time of a bioterror attack and tested it on a low-dose (less than ID_{25}, the dose at which a person has a 25% probability of incurring the disease) anthrax release corresponding, approximately, to the Sverdlovsk anthrax outbreak [5] of 1979. Their formulation incorporated an incubation period model developed by Brookmeyer *et al.* [6] and demonstrated the use of prior distributions on N to reduce uncertainty in the inferred characteristics. Brookmeyer & Blades [7] used a maximum likelihood approach, along with the anthrax incubation model in [6], to infer the size of the 2001 anthrax attacks [1] before estimating the reduction in casualties due to the timely administration of antibiotics. Both [4] and [7] developed similar expressions for the likelihood function, i.e., the probability of observing a patient time series given an attack at time τ with N infected people. The incubation period model in [6] was not dose-dependent, and hence no doses were inferred in these two studies.

Significantly more effort has been spent in characterizing the incubation period of inhalational anthrax. Most work has been experimental, with non-human primates subjected to anthrax challenges [8–13]. Brookmeyer *et al.* [6], on the other hand, used data from the Sverdlovsk outbreak to fit a log-normal distribution of incubation periods valid at low doses; their more recent work, based on a competing risks formulation, includes dose-dependence [14]. Wilkening [15] compares four dose-dependent models for the incubation period distribution, one of which (termed Model D) is structurally identical to Brookmeyer's [14], with updated parameters. Compared to Model D, Wilkening's Model A2 provides slightly better agreement with the spatial and temporal distribution of anthrax cases observed in Sverdlovsk. Yet experimental results by Ivins *et al.* [12] and Brachman *et al.* [13] show significant departures from the results of both models, especially in the 10^3–10^4 spore dose range (see Fig. 1). Thus both A2 and D must be considered approximate, though useful, predictive tools. In this study, we will explore the impact of model error by using Model D to simulate epidemic curves arising from BT attacks while using Model A2 for inference. A more detailed discussion of the anthrax incubation period models is provided in Section 2.2.

The issue of dose-response functions—whether a person exposed to a number of spores will actually contract the disease—will not be addressed in this study. We concentrate on inferring the number of people who are actually infected, not merely exposed to the pathogen. The problem of estimating the probability of infection from D spores was addressed by Brookmeyer *et al.* [14] as well as by Glassman [16] and Druett *et al.* [17]. Haas [18] has established that exposure to low doses can still pose a statistically significant risk to large populations.

The BARD [19] effort also seeks to identify (provide early warning of) a BT attack from the presentation of symptoms. The observables consist of respiratory visits to emergency departments, as might be obtainable from syndromic surveillance systems such as RODS [20]. The model that relates these observables to the testing of competing hypotheses (normal morbidity versus a BT attack-generated spatiotemporal morbidity pattern) of the outbreak includes a Gaussian dispersion plume [21] and Wilkening's A2 model [15]. However, BARD's use in an urban context is only approximate since Gaussian plumes are suited mainly for open spaces [21]. In the tests documented in [19], BARD was used to detect outbreaks with over 900 pre-diagnosed cases but performed poorly for small outbreaks. The paper by Legrand *et al.* [22] investigates simulated anthrax outbreaks with a view of inferring the location and time of release and the quantity released. It was demonstrated that about 15 diagnosed patients were sufficient to infer the characteristics of the release with enough accuracy to make an appreciable impact on a targeted prophylaxis campaign. They also investigated the effect of model uncertainty, i.e., when the model used for inference is at variance with the model used for generating the synthetic data. The effect of a Gaussian plume model (versus a more sophisticated Gaussian puff model) on the accuracy of inferences was also investigated.

In this study, we develop a Bayesian formulation for inferring BT attack characteristics in the form of probability distributions for N, τ, and D, using data from the first 3–5 days of an outbreak. Note that we infer a single representative dose (of anthrax) for the entire infected population, rather than a distribution (e.g., a mean and an interquartile range). This is driven by the tests performed in [22] where an inference approach based on a spatially variable dose performed no better than an older, single-representative-dose version of the inference technique presented in this paper. The observables consist of the number of patients who show symptoms and are diagnosed by a certain point in time. The approach is intended to be used within the context of medical resource planning in the aftermath of a BT attack. We therefore rely on simple temporal input data, reducing the complexity of data collection and the potential for significant observational errors [22]. All tests are performed with anthrax as the pathogen. Compared to [4] and [19], we introduce a new degree of detail to outbreak data and its analysis. Unlike [4], we consider dose-dependent incubation periods and populations infected by a broad range of doses, commensurate with atmospheric dispersion, and infer a representative dose for the population. Since aerosol releases in confined spaces can lead to high doses (comparable to or greater than ID_{50}), the inferred dose serves as a useful indicator of the indoor versus outdoor nature of the release. Model uncertainty is examined here in order to assess how large an error one might encounter under realistic conditions. We also explore how the accuracy and uncertainty of estimates are affected by the size of the outbreak, the dose received, and the frequency with which patient data is collected. Further, we identify correlations between the inferred parameters of the attack, demonstrating realistic cases in which scarce data might support multiple characterizations. These characterizations were not explored in [4, 19]. We then use our method to analyze the Sverdlovsk outbreak of 1979 [5].

In the final section of this study, we solve a problem of optimization under uncertainty, using the probabilistic characterization of the BT attack to guide the allocation of scarce medical resources. In particular, we use the joint posterior distribution of N, τ, and D, and the resulting predictive distributions of subsequent symptomatics, to guard against an over- or under-allocation of resources. Efficient allocation is especially important if further BT attacks are expected. We present some proof-of-concept results, then conclude with a discussion of how more realistic assumptions can be incorporated into the optimization approach, thus creating a useful tool for an emergency manager.

2. The Inverse Problem

2.1. Formulation

We now formulate a Bayesian parameter estimation problem for the characteristics of a BT attack. A detailed derivation can be found in [23]; we reproduce a summary here.

Consider a time series of infected patients $\{t_i, n_i\}$, $i = 0 \ldots M$, where n_i is the number of people developing symptoms in the time interval $(t_{i-1}, t_i]$. For simplicity, we let the intervals be of uniform length $\Delta t = t_i - t_{i-1}$. We will consider two values of Δt, 6 hours and 24 hours, to explore the impact of time resolution in the observations. At time t_0 the first patient(s) become symptomatic; given the finite resolution of our time series, we allow that this patient may have developed symptoms anytime between t_0 and $t_{-1} = t_0 - \Delta t$. M is the total length of the time series and is expected to be small, e.g., 3–5 days. We seek a probabilistic model for these observables, conditioned on an attack that infects N people at time τ with a uniform dose of D spores. By convention, we set t_0 to zero, and thus τ, the time of infection, is always negative.

The dose-dependent incubation period is described by its cumulative distribution function (CDF) $C(T, D)$, where T, the incubation period, is the time elapsed since infection. The probability of an infected individual developing symptoms in the interval $(t_{i-1}, t_i]$ is thus $\{C(t_i - \tau, D) - C(t_{i-1} - \tau, D)\}$. Let $L = \sum_{i=0}^{M} n_i$ be the total number of people who have developed symptoms by the end of the observation period t_M. Then $N - L$ infected people are still asymptomatic; the probability of someone remaining asymptomatic at t_M is the survival probability, $P_{\rm surv}(t_M - \tau, D) = 1 - C(t_M - \tau, D)$. Since the incubation times of each individual are conditionally independent given N, τ, D and the disease model, the probability of the entire time series $\{t_i, n_i\}$ obeys a multinomial distribution with $M+2$ outcomes. One outcome corresponds to remaining asymptomatic at t_M; the $M + 1$ others correspond to developing symptoms in a preceding time interval. The resulting conditional probability distribution is given by the following expression:

$$\begin{aligned} & P\left(\{t_i, n_i\}_{i=0}^{M} | N, \tau, D\right) \\ &= \frac{N!}{(N-L)! \prod_{i=0}^{M} n_i!} \times \{P_{\rm surv}(t_M - \tau, D)\}^{N-L} \\ &\quad \times \prod_{i=0}^{M} \left(C(t_i - \tau, D) - C(t_{i-1} - \tau, D)\right)^{n_i} \\ &\equiv \mathcal{L}\left(N, \tau, D\right). \end{aligned} \tag{1}$$

In the last line of this equation, we rewrite the probability of the observables as a likelihood function $\mathcal{L}(N, \tau, D)$. We then use Bayes rule to obtain the posterior probability of the attack parameters:

$$p\left(N, \tau, \log_{10}(D) | \{t_i, n_i\}_{i=0}^{M}\right) \propto \mathcal{L}(N, \tau, D) \pi_N(N) \pi_\tau(\tau) \pi_D(\log_{10}(D)). \tag{2}$$

Note that we have written the posterior density in terms of $\log_{10}(D)$ rather than D; this is in keeping with [15, 16, 24], where response to infection is generally modeled as a function of the log-dose. Here π_N, π_τ, and π_D are prior densities on N, τ, and $\log_{10}(D)$. Presuming a lack of additional information (and without any reason to believe the contrary), we use broad uniform priors on all three parameters. The joint posterior density can then be marginalized to obtain individual probability density functions (PDFs) for N, τ and $\log_{10}(D)$. Integrals yielding these marginal densities are evaluated using the VEGAS algorithm [25], an iterative adaptive Monte Carlo method implemented in the GNU Scientific Library [26].

2.2. Anthrax incubation models

This section briefly reviews two mathematical models of the incubation period of inhalational anthrax. One of these models is used in the inference procedure. The second is used to simulate anthrax attacks in Sec. 3.2 where we investigate the effect of model uncertainty, i.e., the uncertainty in the inference when the model used for inference is an inexact representation of the processes that generate the data. We also present a comparison of the models versus experimental results (mostly from non-human primates) to provide an estimate of the accuracy (and applicability to humans) of both the models. These models are from Wilkening [15]; details of their derivation can be found in [15, 27]. In both the models, the time of onset of symptoms in a person exposed instantaneously to D anthrax spores is considered to be a random variable, described by its cumulative distribution function (CDF).

The CDF for Wilkening's Model D is given by [15, 27]

$$C_{\text{ModelD}}(T, D) = \int_0^T F(T - s; D, \lambda, \theta) g(s) ds, \tag{3}$$

which is a convolution of $F(T; D)$—the probability that at least one spore out of a dose of D spores will germinate into a vegetative anthrax cell by time t—and $g(s)$, which is the PDF of the time s taken, post-germination, to reach a bacterial load at which symptoms appear. F and g are defined as

$$\begin{aligned} F(T; D, \lambda, \theta) &= \frac{1}{p}\left(1 - \exp\left(-\frac{D\lambda}{\lambda+\theta} Q(T)\right)\right), \quad \text{where} \\ Q(T) &= 1 - \exp\left(-(\lambda+\theta)T\right), \end{aligned} \tag{4}$$

$$p = 1 - \exp\left(-\frac{D\lambda}{\theta+\lambda}\right) \tag{5}$$

and

$$g(s) = \frac{1}{\sqrt{2\pi}\sigma_s s} \exp\left(-\frac{1}{2}\frac{\log^2(s/M_s)}{\sigma_s^2}\right). \tag{6}$$

The probability of showing symptoms in infinite time, denoted p, is also called the attack rate. These distributions depend on a number of parameters:

- N_{thresh}, a threshold bacterial load in a person that causes symptoms;
- t_2, the bacterial load doubling time in a given medium (e.g., mediastinal lymph nodes where the spores germinate), which can be obtained from *in vitro* laboratory experiments;
- t_M, which is the time required to reach a bacterial load of N_{thresh}

$$t_M = t_{\text{lag}} + \frac{t_2}{\log(2)} \frac{N_{\text{thresh}}}{D}$$

  ;
- t_{lag}, a lag time in bacterial growth experiments (typically 1 hour);
- σ_s^2, the variance of the log of the time required to reach the symptomatic bacterial load;
- θ, the probability rate of clearance of a spore (by the immune system), specified in terms of probability of clearance per spore per day;
- λ, the probability rate of germination of a spore, specified in terms of probability of germination per spore per day.

In the present models, M_s, the median time to symptoms, is set to t_M. The values of the parameters for Model D are $\theta = 0.109$ day^{-1}, $\lambda = 8.79 \times 10^{-6}$ day^{-1}, $t_{\text{lag}} = 1$ hour, $t_2 = 2.07$ hour, $N_{\text{thresh}} = 10^9$ and $\sigma_s = 0.544$ day^{-1}.

Sartwell [28] found that the incubation period for a number of diseases was log-normally distributed, which is at odds with Eq. 3. Wilkening's Model A2 captures this alternative by assuming a log-normal distribution,

$$C_{\text{ModelA2}}(T,D) = \frac{1}{2}\left[1 + \text{erf}\left(\frac{\ln(T/T_0)}{\sqrt{2}S}\right)\right], \quad S = 0.804 - 0.079\log_{10}(D), \tag{7}$$

where T_0, the median incubation time, is obtained by solving an integral equation derived from Eq. 3

$$0.5 = \int_0^{T_0} F(T_0 - s; D, \lambda, \theta) g(s) ds.$$

However, in solving for T_0, Wilkening used a slightly different set of parameters: $\theta = 0.11$ day^{-1}, $\lambda = 8.84 \times 10^{-6}$ day^{-1}, $t_{\text{lag}} = 1$ hour, $t_2 = 2.06$ hour, and $\sigma_s = 0.542$ day^{-1}. The reason for the slight change in parameters as well as the difference between Models A2 and D is discussed below.

Parameters in Eqs. 3 and 7 were obtained by fitting the models to the median incubation periods observed in experiments with non-human primates (performed by Henderson *et al.* [8] and Friedlander *et al.* [11]) and to the data from the Sverdlovsk anthrax outbreak. The average dose in the Sverdlovsk outbreak, however, had to be inferred from atmospheric dispersion models and the probability of exhibiting symptoms (in infinite time) given a dose of D spores. This is the procedure adopted by Wilkening [15]. Using Glassman's model [16] for the probability of infection, one obtains an average dose of 2.4 spores. Alternatively, if one employs Eq. 5 (which is similar in form to Druett's [17] and was used by Brookmeyer in [14]) one obtains a dose of 300 spores. Wilkening retained both possibilities and incorporated them into separate models. Model D is based on a dose of 300 spores at Sverdlovsk while A2 assumes 2.4 spores.

In Fig. 1, we plot the median incubation period predicted by Models A2 and D as a function of dosage D. The dosage at Sverdlovsk, estimated as 2.4 spores (represented by $\bullet$) is used to calculate parameters for Model A2 (solid line); the alternative estimate of 300 spores (represented by a filled ∇) is used for Model D (dashed line). Studies by Henderson [8] with 2.1×10^5, 3.9×10^5 and 7.6×10^5 spores (represented as filled $\diamond$) and Friedlander with 3.5×10^5 spores (represented by filled $\triangle$) were also used to calculate the parameters of both models. Studies by Ivins *et al.* [12] (unfilled $\triangle$) and Gleiser *et al.* [10] (unfilled $\square$) were conducted with very few primates and consequently are plotted only for reference. Primate experiments by Brachman [13] simulated the effect of prolonged regular exposure to low doses, as might be the case in a contaminated wool-sorting mill. The primates experienced extended periods during which they received no spores at all. The dose was defined as the total number of spores inhaled and was generally low, between 1000 and 10,000 spores. We plot the resulting ranges of incubation periods observed at various dosages, also for reference.

We see that the tests by Gleiser *et al.* and Ivins *et al.* agree with both models, which in turn agree with each other. However, significant differences arise when $D \lesssim 10^3$ spores. (Note that the vertical axis is logarithmic.) Brachman's tests show median incubation periods which are at odds with the models' predictions; however the mode of infection (a continuous low-level infection process spread over days or months) was very different from the rapid (timescale of an hour) challenge one would expect in a BT attack. Both models show a "kink" at $D \approx 10^3$; this is because they are evaluated with a lower value of λ (1.3×10^{-6} day^{-1}), corresponding to a primate ID_{50} of 55,000 spores, for comparison with primate results at the high dose limit, while the low dose predictions were developed with a human ID_{50} of 8600 spores for comparison with Sverdlovsk data. To the best of the authors' knowledge, this is the sum total of experimental data obtained from anthrax challenges of non-human primates where incubation times were measured. We have omitted a study by Klein *et al.* [29] in which an incubation period increase

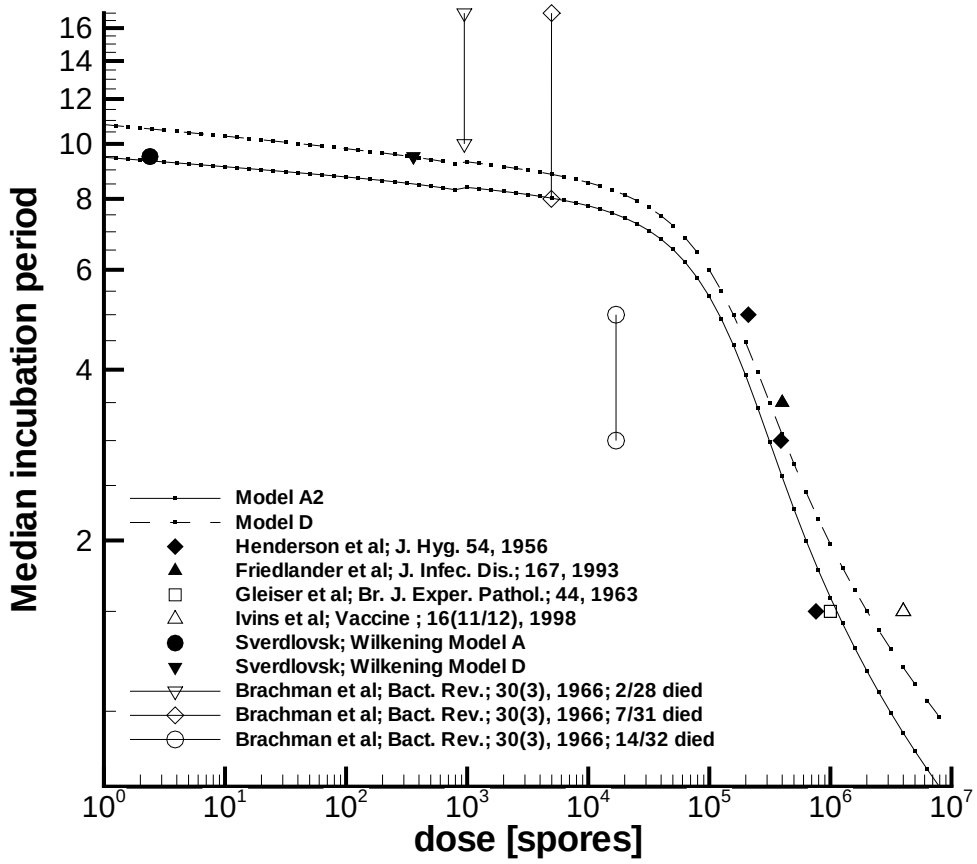

Figure 1. The median incubation period (in days) for anthrax as a function of dose D. The solid line is Model A2, which assumes a dose of 2.4 spores at Sverdlovsk; the dashed line is Model D, which assumes 300 spores. The solid symbols are median incubation periods obtained from experimental investigations or from Sverdlovsk data. The filled circle (Sverdlovsk; Wilkening Model A) refers to both Models A1 and A2, though only Model A2 is used in the current study. Symbols which are not filled denote experiments where the population of primates was too small to draw statistically meaningful results. The experiments by Brachman *et al.* [13] are shown by vertical lines between symbols. In these experiments, only the lower and upper bounds of the incubation period were provided. These ranges were not used for determining model parameters and are only provided for reference.

was observed with increasing doses, because only one primate was subjected to each dose, making the behavior statistically unreliable.

2.3. Inference of attack parameters with ideal cases

In this section we test the Bayesian estimation procedure described above. We use Wilkening's Model A2, described in Sec. 2.2, to simulate symptomatic times for inhalational anthrax outbreaks of different sizes. The same model is used for inference; that is, there are no systematic errors between the inference and data-generation models. Thus, posterior uncertainties may be ascribed to (1) incomplete observation of the outbreak, specifically finite

Table I. Time series obtained from six different outbreaks, simulated with the parameters $\{N, \tau, D\}$ as noted at the bottom of the table. The table has been divided into 24-hour sections, where the n_i in each section are summed to produce the low-resolution time series (24-hour resolution) used to investigate the effect of temporal resolution. Time is measured in days and dose in spores.

time	Simulation A	Simulation B	Simulation C	Simulation D	Simulation E	Simulation F
0.00	1	1	1	1	2	1
0.25	0	2	2	7	13	7
0.50	0	1	1	12	18	24
0.75	1	1	1	39	39	29
1.00	2	2	2	50	38	60
1.25	0	3	3	77	64	96
1.50	1	2	3	77	84	153
1.75	2	1	1	98	116	164
2.00	1	1	2	126	130	193
2.25	1	1	2	162	137	223
2.50	2	3	3	146	141	258
2.75	3	1	4	148	160	302
3.00	2	1	3	149	190	299
3.25	1	3	3	163	175	312
3.50	1	1	2	181	182	304
3.75	1	1	3	162	201	335
4.00	2	1	2	165	200	373
4.25	1	5	5	177	238	340
4.50	1	4	4	169	202	327
4.75	3	2	2	217	216	332
5.00	1	1	1	167	217	350
5.25	1	3	4	182	237	321
5.50	1	1	5	163	207	316
N	100	100	100	10,000	10,000	10,000
τ	-0.75	-2.25	-2.25	-0.5	-1.0	-1.25
D	1	100	10,000	1	100	10,000

time resolution Δt and a short time series, and (2) the probabilistic character of disease incubation. We investigate how the quality of the inference varies with the size of the outbreak and the dose received. We also investigate whether a higher-resolution time series spanning a given observation period performs significantly better than a lower-resolution one.

In Table I, we list time series at 6-hour resolution: the number of patients showing symptoms collected over 6-hour intervals obtained from 6 simulated outbreaks, henceforth called Simulations A–F. Each infected patient received an identical dose D. N indicates the number of people infected and τ is the time of attack, measured in days prior to the exhibition of symptoms in the first diagnosed patient.

We use the procedure outlined in Sec. 2.1 to develop posterior PDFs for N, τ, and $\log_{10}(D)$ in Simulations A–F. Figs. 2, 3, and 4 plot the resulting marginal densities for $\{N, \tau, \log_{10}(D)\}$. These are conditioned on the 6-hour resolution time series listed in Table I. In Table II, we summarize the maximum a posteriori (MAP) estimates and 90% posterior credibility intervals (CIs) for N, τ, and $\log_{10}(D)$ obtained with 5 days of data. We see that the marginal MAP estimate of N (the value of N corresponding to the peak of $p\left(N|\{t_i, n_i\}_{i=0}^{M}\right)$) is generally close to the correct value after 5 days of data (see Table II). Even with 3 days of data, the PDFs in Figs. 2, 3, and 4 are quite informative. Increasing the length of the observation period to 5 days usually sharpens the PDF, reflecting a reduction in uncertainty. This trend holds true for small

attacks ($N = 10^2$) as well as for large ones ($N = 10^4$). An exception is Simulation F, which will be discussed below. The marginal MAP estimate of the time of attack τ is also close to the correct value, except for the small-N low-dose Simulation A. Larger attacks (Simulations D, E, and F) have narrower PDFs for τ compared to Simulations A, B, and C. Higher values of n_i in Eq. 1 (which generally result from large N attacks) provide structure in $\mathcal{L}$ and allow a more accurate estimation of the attack.

The dose D is the most difficult parameter to infer. PDFs for Simulations A, B and C in Figs. 2 and 3 show that it is virtually impossible to estimate the dose for small ($N = 10^2$) attacks; appreciable posterior probability is spread over 5 orders of magnitude. Table II confirms that MAP estimates of the dose in these small attacks are incorrect. Larger attacks ($N = 10^4$) yield more informative PDFs for D. Note that the sensitivity of $C(T, D)$ to D is rather small for Model A2 (see the expression for S in Eq. 7), suggesting that dependence of the likelihood function on D will be weak unless n_i or M is large.

Simulations D, E, and F (Figs. 3 and 4) demonstrate how early observations of an outbreak may support multiple hypotheses, and at times favor a "wrong" hypothesis over the correct one. For instance, Simulation D exhibits peaks in $p(N)$ at $N \approx 4 \times 10^3$ and $N \approx 10^4$. Peaks in the PDF of $\log_{10}(D)$ occur at 1 spore and between 10^4 and 10^5 spores. For this simulation, both marginal PDFs overwhelmingly favor a large N, low-dose attack, which is the correct characterization. A similar ambiguity is observed in Simulation E. Marginal PDFs in Simulation F (Fig. 4) are much more strongly bimodal, however. In Fig. 5 we plot the joint posterior density $p(N, \log_{10}(D))$ to examine correlations among these parameters; it clearly shows two distinct islands—one corresponding to a large-N low-dose attack, and the other corresponding to a small-N high-dose attack. Up to Day 5, the data favor the wrong hypothesis (a larger, low-dose attack) over the correct one. Note also that the large low-dose attack corresponds to larger (i.e., later) values of τ, as evidenced by the posterior density $p(\tau)$ for Days 3–5 (Fig. 4, right column). With more data (Day 6 and 7), the correct values for $\{N, \tau, \log_{10}(D)\}$are recovered, with peaks at $N \approx 10^4$, $\tau \approx -1.2$, and $\log_{10}(D) \approx 4$. However, such a long observation period would not be relevant for consequence planning purposes. We stress that a Bayesian analysis is free to identify competing hypotheses, and that the degree of belief assigned to each is determined by the data and the prior information. In a partially observed attack, the MAP estimate may be erroneous, especially if data are scarce. One possible remedy is the use of informative priors for N, τ, and/or $\log_{10}(D)$ instead of the broad uniform priors used here. Otherwise, natural ambiguities may remain and should be accounted for in consequence management plans based on these inferences.

Coarser time resolution ($\Delta t = 24$ hours instead of 6 hours) was investigated in [27] and generally yielded only a mild degradation in the smoothness of the PDFs. In simulations where a multimodal PDF evolves into a unimodal PDF over time (e.g., Simulations D, E, and F), evolution is more rapid when the observations are collected in 6-hour intervals.

To summarize, solution of the inference problem successfully provides N and τ for small and large attacks. After 5 days of observations, the error in the MAP estimate for N (versus the true value) is less than 50% for small attacks (100 infected people) and less than 5% for large attacks (10,000 infected people). The same holds true for τ, except the errors are smaller than for N. D can be estimated only for large attacks (errors less than about 25%). This may be due to the low sensitivity of the early symptomatic patient stream to dose, an observation that was also made by Legrand *et al.* [22] where they characterized an anthrax BT attack with a low-dose incubation period model, without too much error. Posterior PDFs are sharper for large attacks and for high-dose attacks. Higher temporal resolution may smooth the PDFs slightly. When conditioning on a short time series, the Bayesian method may suggest multiple hypotheses, supported to differing degrees by the data. In some simulations, e.g., Simulation F, the data might initially support the wrong hypothesis, but the correct characterization is recovered as more data become available.

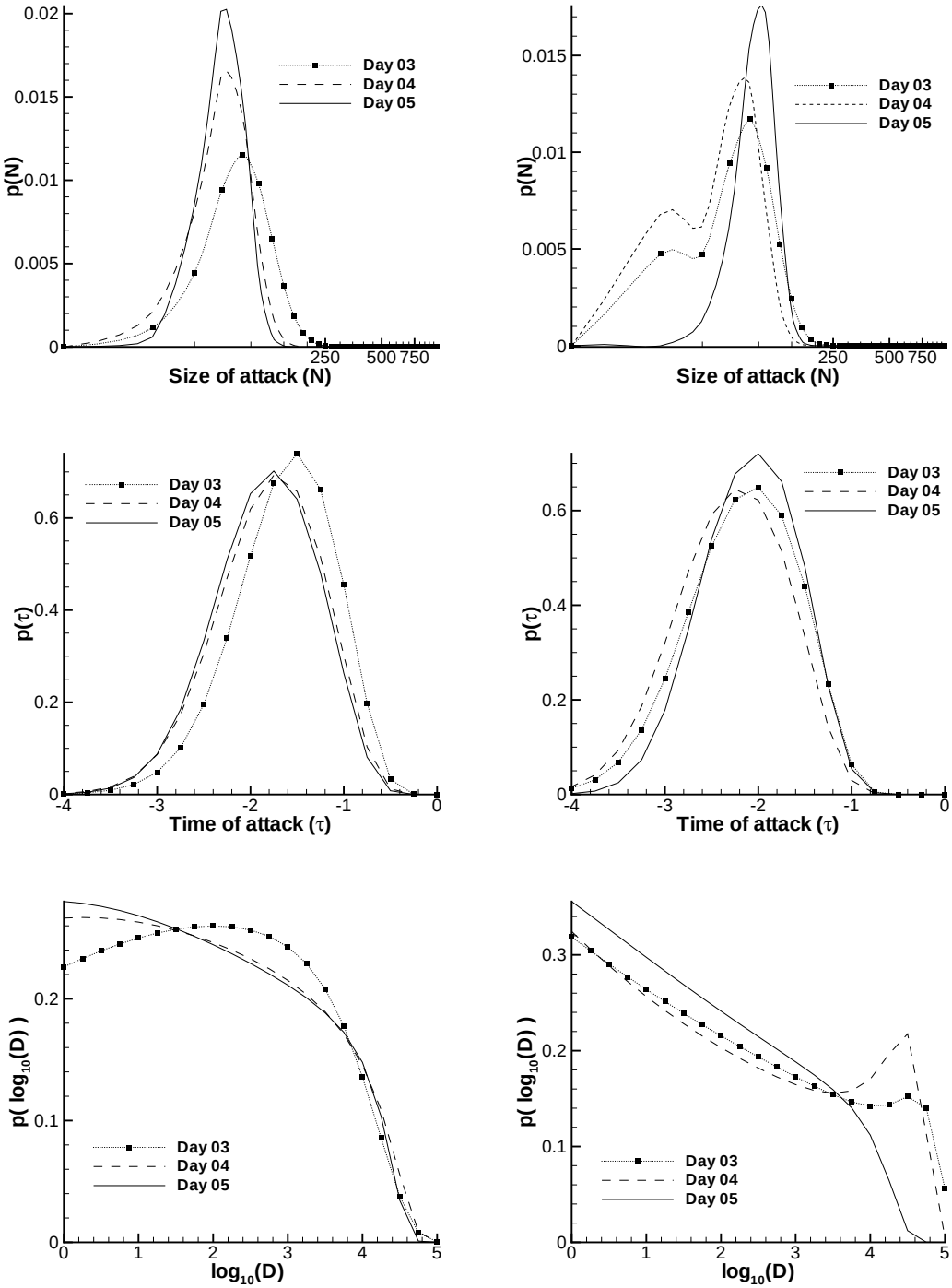

Figure 2. Posterior PDFs for N (top), τ (middle), and $\log D$ (bottom) based on the time series for Simulation A (left) and Simulation B (right), as tabulated in Table I. Data are collected at 6-hour intervals in both simulations. The correct values for $\{N, \tau, \log_{10}(D)\}$in Simulation A are $\{10^2, -0.75, 10^0\}$; in Simulation B they are $\{10^2, -2.25, 10^2\}$. In both simulations, PDFs are reported after 3-, 4- and 5-day observational periods (dotted, dashed, and solid lines respectively).

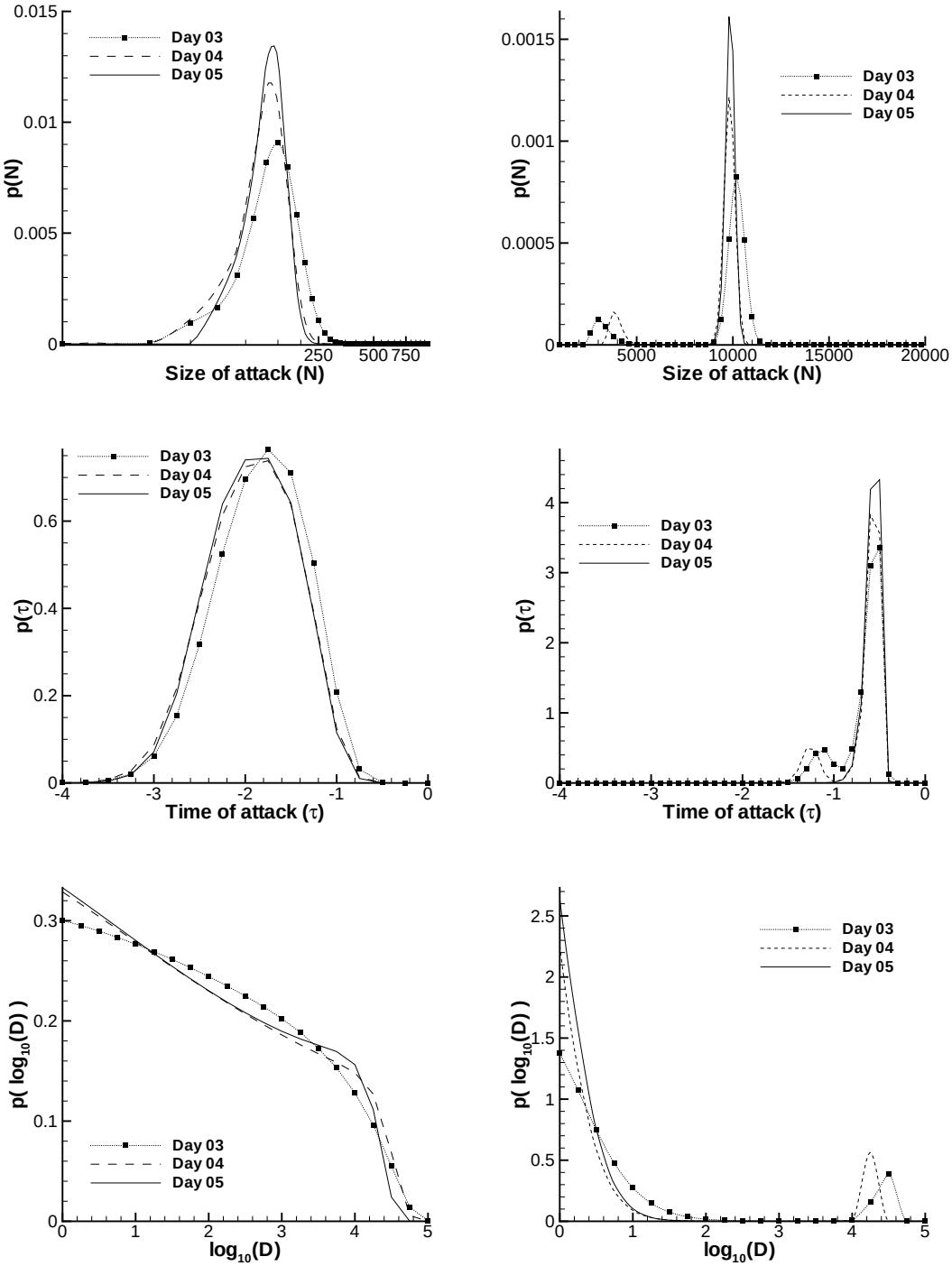

Figure 3. Posterior PDFs for N (top), τ (middle), and $\log D$ (bottom) based on the time series for Simulation C (left) and Simulation D (right), as tabulated in Table I. Data are collected at 6-hour intervals in both simulations. The correct values for $\{N, \tau, \log_{10}(D)\}$in Simulation C are $\{10^2, -2.25, 10^4\}$; in Simulation D they are $\{10^4, -0.05, 10^0\}$. In both simulations, PDFs are reported after 3-, 4- and 5-day observational periods (dotted, dashed, and solid lines respectively).

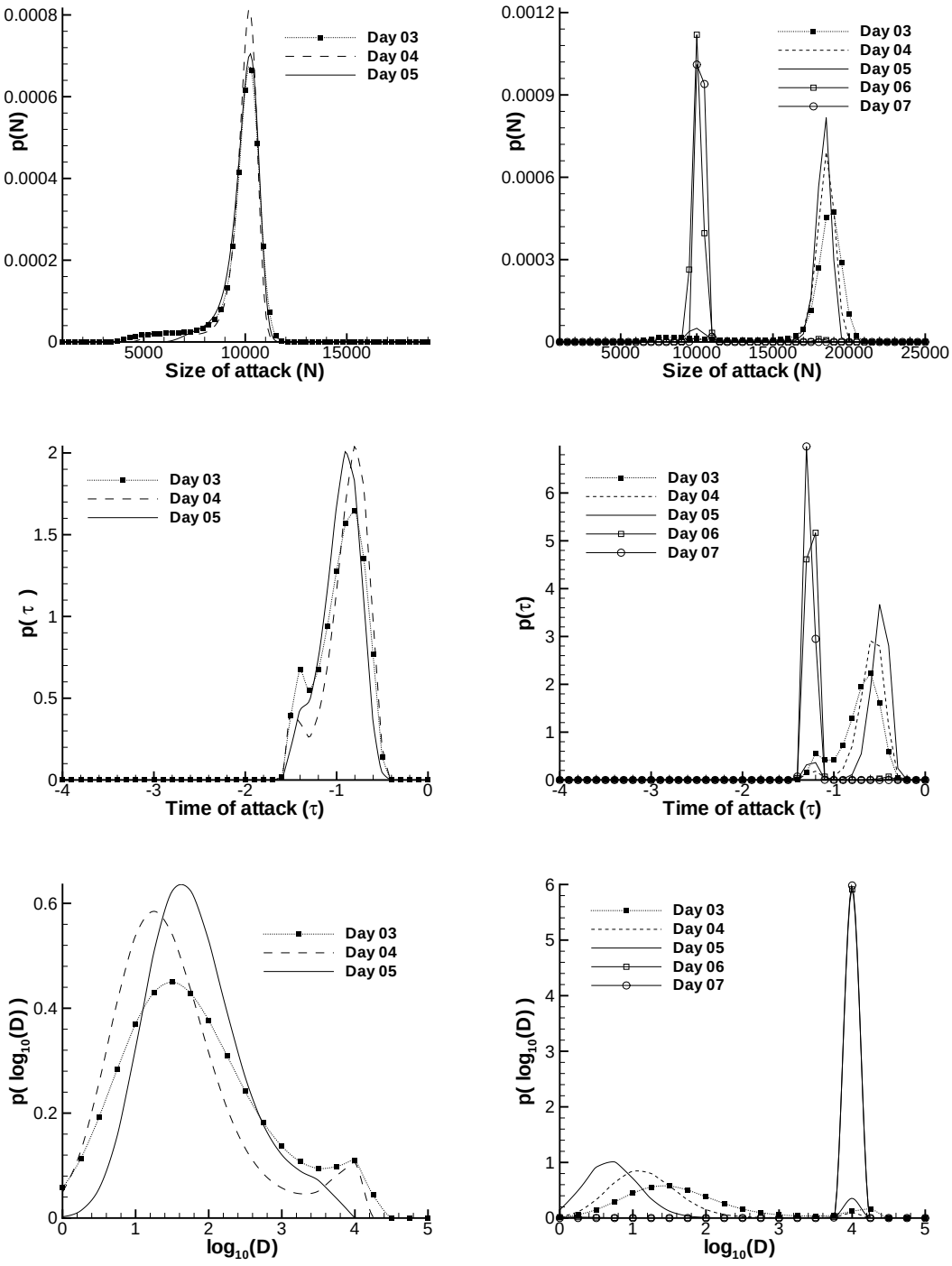

Figure 4. Posterior PDFs for N (top), τ (middle), and $\log D$ (bottom) based on the time series for Simulation E (left) and Simulation F (right), as tabulated in Table I. Data are collected at 6-hour intervals in both simulations. The correct values for $\{N, \tau, \log_{10}(D)\}$in Simulation E are $\{10^4, -1.0, 10^2\}$; in Simulation F they are $\{10^4, -1.25, 10^4\}$. In both simulations, PDFs are reported after 3-, 4- and 5-day observational periods (dotted, dashed, and solid lines respectively), but Simulation F also includes PDFs at Day 6 (solid lines with unfilled squares) and at Day 7 (solid lines with unfilled circles).

Prepared using simauth.cls

Table II. Simulations A–F; MAP estimates and 90% credibility intervals (in parentheses) for N, τ, and $\log_{10}(D)$, conditioned on the high-resolution time series at Day 5. The number in the curly brackets {} is the correct value.

Simulation	N	τ	$\log_{10}(D)$
A	70, (39.45 – 123.3) {100}	-1.75, (-2.90 – -1.04) {−0.75}	0.0, (0.18 – 4.12) {0}
B	110, (65.7 – 148.4) {100}	-2.0, (-3.1 – -1.33) {−2.25}	0.00, (0.14 – 3.97) {2}
C	150, (88.78 – 194.7) {100}	-1.75, (-2.85 – -1.22) {−2.25}	0.0, (0.153 – 4.13) {4}
D	9800, (9439 – 10,350) {10,000}	-0.50, (-0.85 – -0.44) {−0.50}	0.00, (0.024 – 1.03) {0}
E	10,200, (8396 – 10,890) {10,000}	-0.9, (-1.41 – -0.67) {−1.00}	1.75, (0.87– 3.23) {2}
F	18,500, (10,500 – 19,290) {10,000}	-0.5, (-0.99 – -0.34) {−1.25}	0.75, (0.16 – 3.84) {4}

3. Inference of Attack Parameters Under Variable Doses

In this section we conduct eight tests corresponding to more realistic conditions. In the first four (Simulations I, Ia, II, and IIa) we relax the assumption of a constant dose D; instead, the infected people receive a range of doses commensurate with atmospheric dispersion. However, the disease is still assumed to evolve per Wilkening's Model A2, with the same model providing $C(T, D)$ to the inference procedure. In the second set of tests (Simulations III, IIIa, IV, and IVa), we retain distributed doses and additionally relax the second assumption: data are generated with Model D, while the inference procedure still uses Model A2 to evaluate the incubation period distribution. This mismatch introduces a degree of realism into the inference process since the host-pathogen interaction for humans and anthrax will seldom be characterized accurately.

In order to obtain a realistic distribution of doses in a geographically distributed population, we first simulate an explosive point release of spores at a height of 100 meters with a Gaussian plume model, thus exposing different numbers of people to varying doses as described in Appendix I. We see from Fig. 15 that given a quantity of spores, the number of people infected depends on the total population in the domain, the orientation of the plume, and the population distribution. A release does not lead to many infected people if the high concentration isopleths of the plume miss the localized regions of high population density.

The series of symptomatic times arising from such a simulated attack will reflect the evolution of inhalational anthrax in an infected population that receives a range of doses. These data will be "fit" using the model described in Sec. 2.1, which assumes a uniform dose for all infected individuals. The uniform dose thus inferred is, in a sense, a representative dose for the entire infected population; however, it is *not* rigorously linked to the median or mean of the actual distribution of doses. To avoid confusion, we therefore refer to the inferred (uniform)

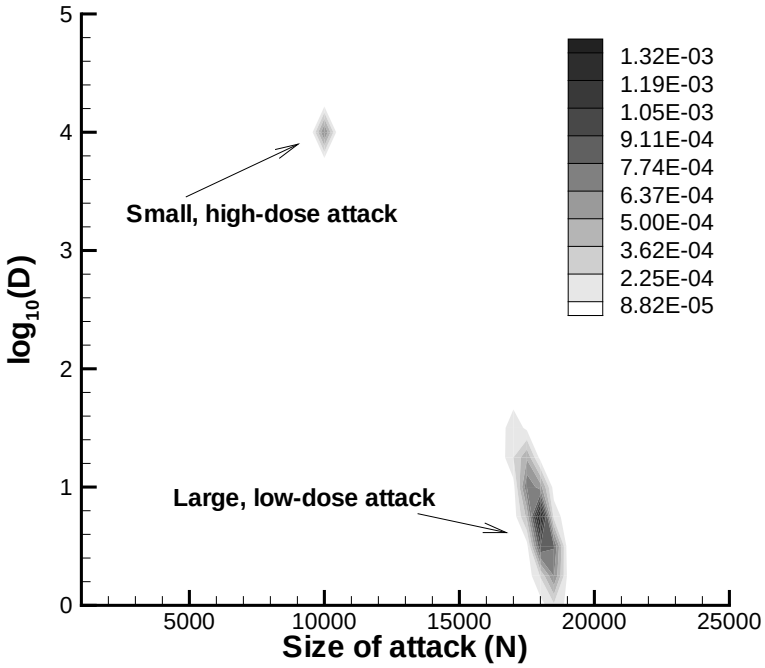

Figure 5. The joint probability density $p(N, \log_{10}(D))$ obtained after 5 days of data for Simulation F. We clearly see a dual characterization—a larger low-dose attack and a smaller high-dose attack.

dose as the "representative" dose.

Tables III and IV list the time series obtained from all eight simulations. The time series have a resolution of 6 hours, with successive 24-hour intervals indicated in the tables. As noted in Appendix I, these simulations correspond to two choices of population size ($p_{\text{exposed}} = 10^3$ for Simulations Ia, II, IIIa and IV; $p_{\text{exposed}} = 10^4$ for Simulations I, IIa, III and IVa) combined with two choices of plume orientation ($\theta = 170°$ for Simulations I, Ia, III and IIIa; $\theta = 125°$ for Simulations II, IIa, IV and IVa). The latter orientation directs the plume over a more population-dense region. Tables III and IV also report quantiles of the dose distribution $D_1, D_{25}, D_{50}, D_{75}$, and D_{99}. That is, 1% of the population receives a dose of D_1 spores or less, 75% of the population receives less than D_{75} spores, and D_{99} is near the maximum dose. In Fig. 16 (Appendix I), we plot dose distributions corresponding to the simulations listed in Tables III and IV. Note that while the doses may easily span two orders of magnitude, about 80% of the infected people lie within a one-decade range of doses. The inferred representative dose D may reflect this range, and thus for limited verification purposes, we will compare D to the actual median dose D_{50}. Estimating a single D is, of course, a source of model error, adding to the uncertainty caused by incomplete observations and the inherent stochasticity of the data. This model error is not expected to diminish with additional data, and one of the aims of this investigation is to quantify it.

3.1. Inference of attack parameters without incubation model mismatch

We begin with results from Simulations Ia, I, II, and IIa—i.e., eliminating the assumption that each infected person receives the same dose of anthrax spores, but simulating and inferring

Table III. Time series obtained from eight simulated outbreaks with variable doses. Simulations I, Ia, II, and IIa are simulated using Wilkening's Model A2, with the attack parameters—N, τ, and the dose distribution—indicated at the bottom of the table. Simulations III, IIIa, IV and IV are simulated using Wilkening's Model D. $\bar{D}$ is the average dose for the N infected individuals. The table has been divided into 24-hour sections, where the values n_i in each section can be summed to produce the low-resolution time series used to investigate the effect of temporal resolution. The dose distribution is represented by its quantiles D_1, D_{25}, D_{50}, D_{75}, and D_{99}; x% of the population receives a dose of D_x or less. Table IV continues the time series from Day 5 to Day 8.

Time (days)	Sim. Ia	Sim. I	Sim. II	Sim. IIa	Sim. IIIa	Sim. III	Sim. IV	Sim. IVa
0.0	1	3	2	5	1	1	1	3
0.25	2	3	2	8	1	8	5	14
0.50	0	6	1	8	0	20	6	36
0.75	4	12	5	27	1	16	13	81
1.00	1	14	7	46	3	9	12	77
1.25	2	26	12	57	2	18	14	94
1.50	2	28	9	85	2	28	13	123
1.75	6	49	16	94	1	30	13	132
2.0	6	57	9	133	2	37	17	129
2.25	5	65	20	134	2	27	15	159
2.50	7	68	12	139	4	41	17	126
2.75	6	53	18	163	2	39	14	149
3.0	11	80	15	138	3	34	9	131
3.25	8	62	15	180	2	32	14	129
3.50	9	89	21	140	3	25	16	136
3.75	8	106	16	164	6	33	12	100
4.00	17	70	20	180	4	27	14	125
4.25	12	65	21	136	5	33	11	104
4.50	9	87	8	147	3	33	6	110
4.75	3	87	8	151	5	23	11	106
5.0	6	76	7	127	6	23	15	90
N	318	2989	454	4537	161	1453	453	4453
τ	-1.5	-1.5	-1.5	-1.25	-0.75	-0.75	-0.75	-0.5
$\bar{D}$	2912.8	2776.8	13,870.5	13,150.4	3603.5	3660.77	16,941	16,532
$D_1 \times 10^{-2}$	0.53	0.65	1.39	1.32	3.41	2.65	3.1	3.0
$D_{25} \times 10^{-3}$	1.23	1.15	3.96	3.47	1.99	2.13	9.8	9.45
$D_{50} \times 10^{-4}$	0.29	0.26	1.34	1.24	0.33	0.35	1.65	1.57
$D_{75} \times 10^{-4}$	0.41	0.39	1.91	1.87	0.48	0.48	2.09	2.07
$D_{99} \times 10^{-4}$	0.83	0.87	5.79	5.91	0.92	0.95	6.74	6.52

disease progression with Wilkening's Model A2.

Figs. 6, 7, 8, and 9 show posterior PDFs for $\{N, \tau, \log_{10}(D)\}$ conditioned on the time series in Tables III and IV. Table V reports the MAP estimates and the 90% CIs for $\{N, \tau, \log_{10}(D)\}$ after 5 days of data. Since the true doses are distributed, we use the log of the median dose, $\log_{10}(D_{50})$, as a reasonable value for comparison to the posterior $\log_{10}(D)$.

First consider Figs. 6 and 7, corresponding to Simulations Ia and I. These attacks have similar dose distributions but differ by an order of magnitude in N. In both simulations, the MAP estimate of τ nearly coincides with the true value after only 3 days of data. In Simulation Ia, the MAP estimate of N deviates from the true value by approximately 20%, but the 90% CIs bracket the correct N quite easily. In Simulation I, the PDF for N initially favors an inaccurate characterization (a peak at $N \approx 4000$) but by Day 5, assumes a bimodal

Table IV. Continuation of Table III beyond Day 5. Time series obtained from 4 simulated outbreaks with variable doses. Simulations I, Ia, II, and IIa are simulated using Wilkening's Model A2, with the attack parameters—N, τ, and the dose distribution—indicated at the bottom of the table. $\bar{D}$ is the average dose for the N infected individuals. The table has been divided into 24-hour sections, where the values n_i in each section can be summed to produce the low-resolution time series used to investigate the effect of temporal resolution. The dose distribution is represented by its quantiles D_1, D_{25}, D_{50}, D_{75}, and D_{99}; x% of the population receives a dose of D_x or less.

Time (days)	Simulation Ia	Simulation I	Simulation II	Simulation IIa
5.25	9	70	16	129
5.50	8	91	8	109
5.75	10	79	9	147
6.00	9	86	12	126
6.25	8	82	13	108
6.50	7	55	9	114
6.75	7	69	7	90
7.0	6	75	8	96
7.25	8	61	6	88
7.50	4	67	6	77
7.75	6	65	8	75
8.00	2	62	6	69
N	318	2989	454	4537
τ	-1.5	-1.5	-1.5	-1.25
$\bar{D}$	2912.8	2776.8	13,870.5	13,150.4
$D_1 \times 10^{-2}$	0.53	0.65	1.39	1.32
$D_{25} \times 10^{-3}$	1.23	1.15	3.96	3.47
$D_{50} \times 10^{-4}$	0.29	0.26	1.34	1.24
$D_{75} \times 10^{-4}$	0.41	0.39	1.91	1.87
$D_{99} \times 10^{-4}$	0.83	0.87	5.79	5.91

shape with a peak close to the correct characterization. Dose is the most difficult parameter to estimate in Simulation Ia—the marginal PDF of $\log_{10}(D)$ remains rather broad at all times. In the larger-N Simulation I, however, the posterior on $\log_{10}(D)$ at least indicates that the attack is not a low dose (i.e., $D_{50} \leq \mathrm{ID}_{25}$) event. Also in Simulation I, conditioning on the high resolution time-series provides more structure to the PDF; the posterior densities on $\log_{10}(D)$ and even on τ are more prominently bimodal, indicating that inference is inconclusive, and more observations will be required to obtain a unique characterization. For reference, both Figs. 6 and 7 include a further set of PDFs conditioned on data through Day 7; MAP estimates from these posteriors generally show even closer agreement with the true values of $\log_{10}(D_{50})$ and N.

Inference is considerably less challenging in Simulations II and IIa, corresponding to Figs. 8 and 9. Because the doses are higher ($D_{50} > \mathrm{ID}_{50}$), the variance of the incubation period distribution is smaller. The time of attack τ is captured with only 3 days of data, as is a representative $\log_{10}(D)$ for the large N attack (Simulation IIa). With 5 days of data, MAP estimates for N are close to the correct values in both simulations, as is the MAP estimate of $\log_{10} D$ in Simulation II. Here, conditioning on the higher-resolution time series yielded little gain over the lower-resolution time series. In Simulation II, MAP estimates of τ based on 6-hour data are in fact inaccurate on Days 3 and 4, recovering the correct characterization after 5 days of data.

In general, therefore, many of the behaviors discussed in Sec. 2.3 are repeated in the present

Simulation	N	τ	$\log_{10}(D)$
Ia (6-hr resolution)	400, (233.6 – 581.9)	-1.5, (-2.00 – -0.795)	3.0, (0.37 – 3.99)
Ia (24-hr resolution)	400, (230.4 – 582.2) {318}	-1.5, (-2.04 – -0.78) {−1.5}	2.75, (0.32 – 4.00) {3.46}
I (6-hr resolution)	4100, (2334 – 4439)	-1.4, (-1.57 – -0.64)	4.00, (0.715 – 4.147)
I (24-hr resolution)	4000, (2281 – 4358) {2989}	-1.4, (-1.59 – -0.70) {−1.5}	4.00, (0.91 – 4.173) {3.41}
II (6-hr resolution)	400, (305.5 – 981.6)	-1.5, (-1.98 – -1.08)	4.25, (0.68 – 4.72)
II (24-hr resolution)	400, (327.0 – 984.7) {454}	-1.6, (-2.10 – -1.03) {−1.5}	4.25, (0.36 – 4.69) {4.13}
IIa (6-hr resolution)	3900, (3686 – 4340)	-1.3, (-1.48 – -1.14)	4.25, (4.05 – 4.72)
IIa (24-hr resolution)	4000, (3709 – 4433) {4537}	-1.5, (-1.55 – -1.18) {−1.25}	4.25, (4.04 – 4.72) {4.09}

Table V. Simulations I, Ia, II, IIa; MAP estimates and 90% credibility intervals (in parentheses) for N, τ, and $\log_{10}(D)$ conditioned on data through Day 5. Correct values for N and τ are in { }. The inferred representative dose is compared with $\log_{10}(D_{50})$, also in { }.

simulations. The representative dose D is difficult to estimate for small N attacks, while the time τ is always easy to infer. We can bound the size N of the attack quite accurately for all simulations. MAP estimates of N obtained from 5 days of data are always within 20% of the correct value. Further, the 90% CIs at Day 5 for N, τ, and $\log_{10}(D)$ almost always bracket the true attack parameters. Finer temporal resolution Δt may better capture the evolution of the outbreak, but has a relatively minor impact on summaries of the posterior; MAP estimates obtained from the low and high-resolution time series are similar, as are the 90% CIs. Thus, while the errors incurred in fitting variable-dose data with a constant-dose inference model are not negligible, the current formulation provides a reasonable and useful characterization of the BT attack.

3.2. Inference of variable dose attack with incubation model mismatch

We now proceed to Simulations III, IIIa, IV, and IVa. As noted above, these simulations introduce a systematic difference between the simulated evolution of the disease in infected persons and the model used to interpret the observed data. We simulate BT attacks using Wilkening's Model D (i.e., sampling the incubation period distribution in Eq. 3), but infer the attack parameters using Model A2. As in Sec. 3.1, the infected population receives a distribution of doses (see Appendix I) but the model used in the inference process assumes a constant dose.

Figs. 10 and 11 show posterior PDFs for $\{N, \tau, \log_{10}(D)\}$ conditioned on the time series in Table III. As described in the preceding section, finer resolution in the time series does not have a great impact on the posterior, and hence we only plot PDFs resulting from daily observations in each case. Several features are worth highlighting. First, the dose is identified much more closely in Simulations IV and IVa, where both N and $\log_{10}(D)$ are higher, than in Simulations

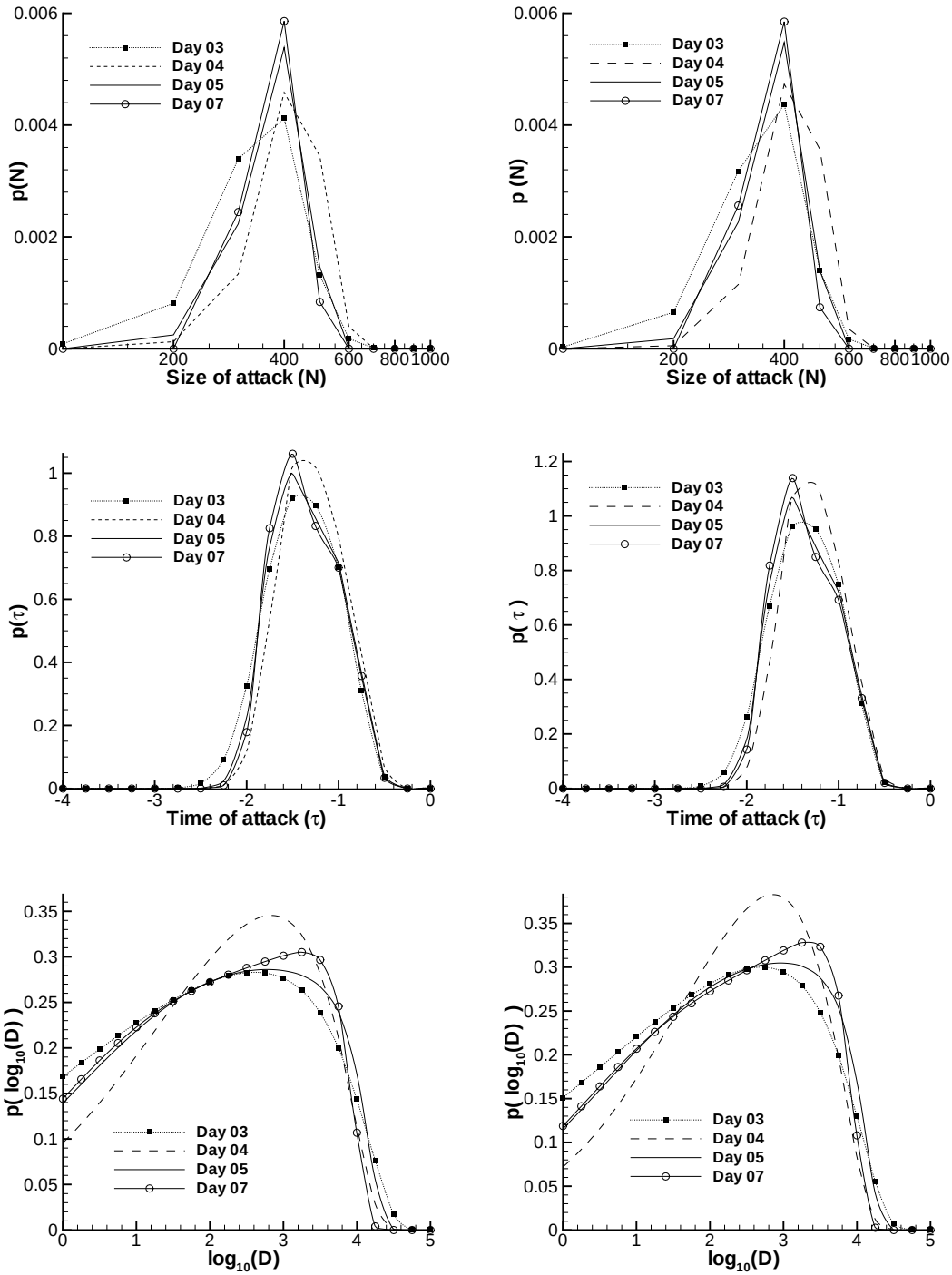

Figure 6. Posterior PDFs for N (top), τ (middle), and $\log_{10} D$ (bottom) based on the time series for Simulation Ia, as tabulated in Tables III and IV. Lower-resolution data (collected in 24-hour intervals) yield the PDFs on the left, while higher-resolution data yield the PDFs on the right. Correct values for $\{N, \tau, \log_{10}(D)\}$ are $\{318, -1.5, 3.46\}$, where the "correct" representative dose is taken to be $\log_{10}(D_{50})$. In both simulations, PDFs are reported after 3-, 4-, 5-, and 7-day observational periods.

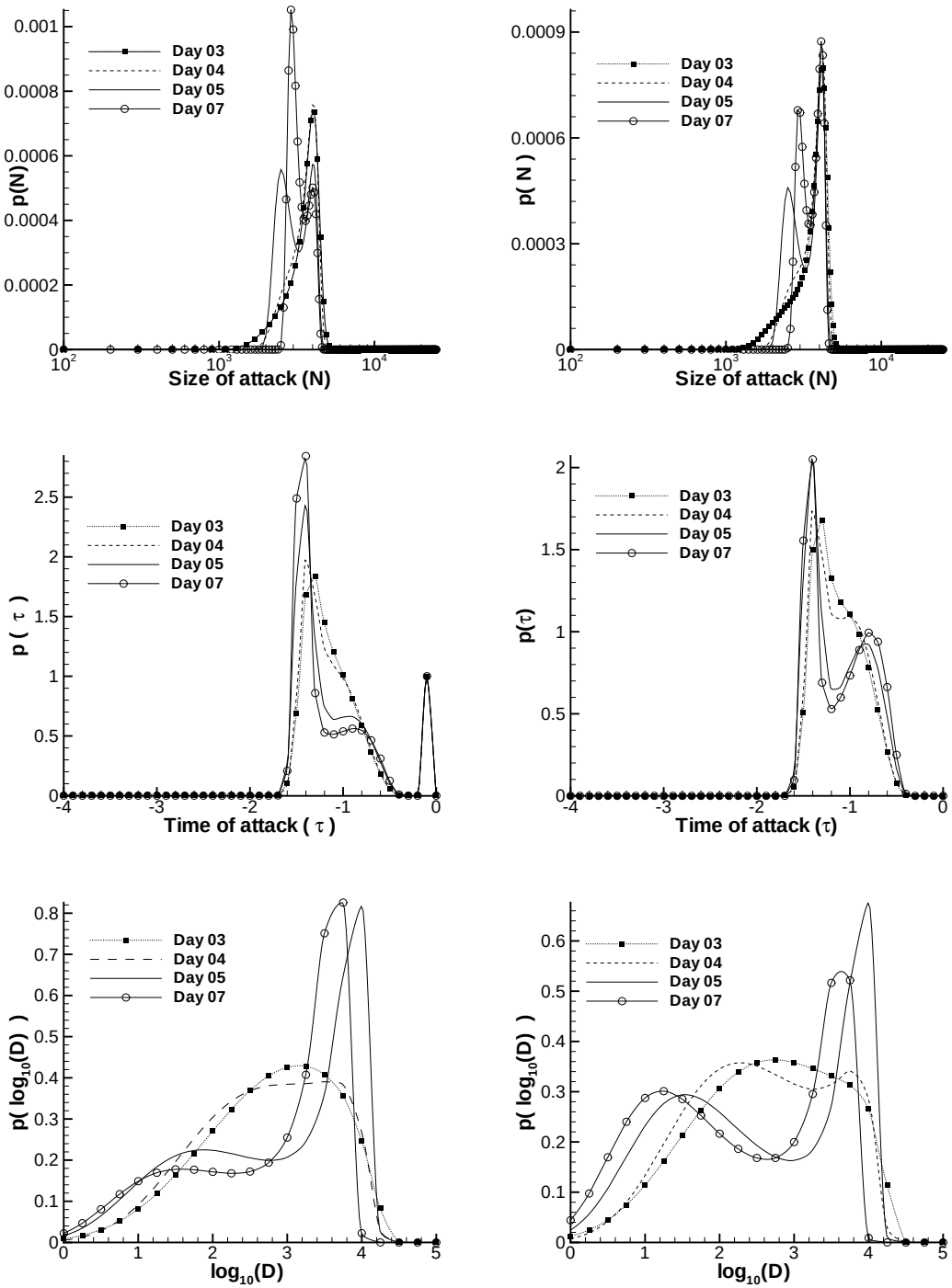

Figure 7. Posterior PDFs for N (top), τ (middle), and $\log_{10} D$ (bottom) based on the time series for Simulation I, as tabulated in Tables III and IV. Lower-resolution data (collected in 24-hour intervals) yield the PDFs on the left, while higher-resolution data yield the PDFs on the right. Correct values for $\{N, \tau, \log_{10}(D)\}$ are $\{2989, -1.5, 3.41\}$, where the "correct" representative dose is taken to be $\log_{10}(D_{50})$. In both simulations, PDFs are reported after 3-, 4-, 5-, and 7-day observational periods.

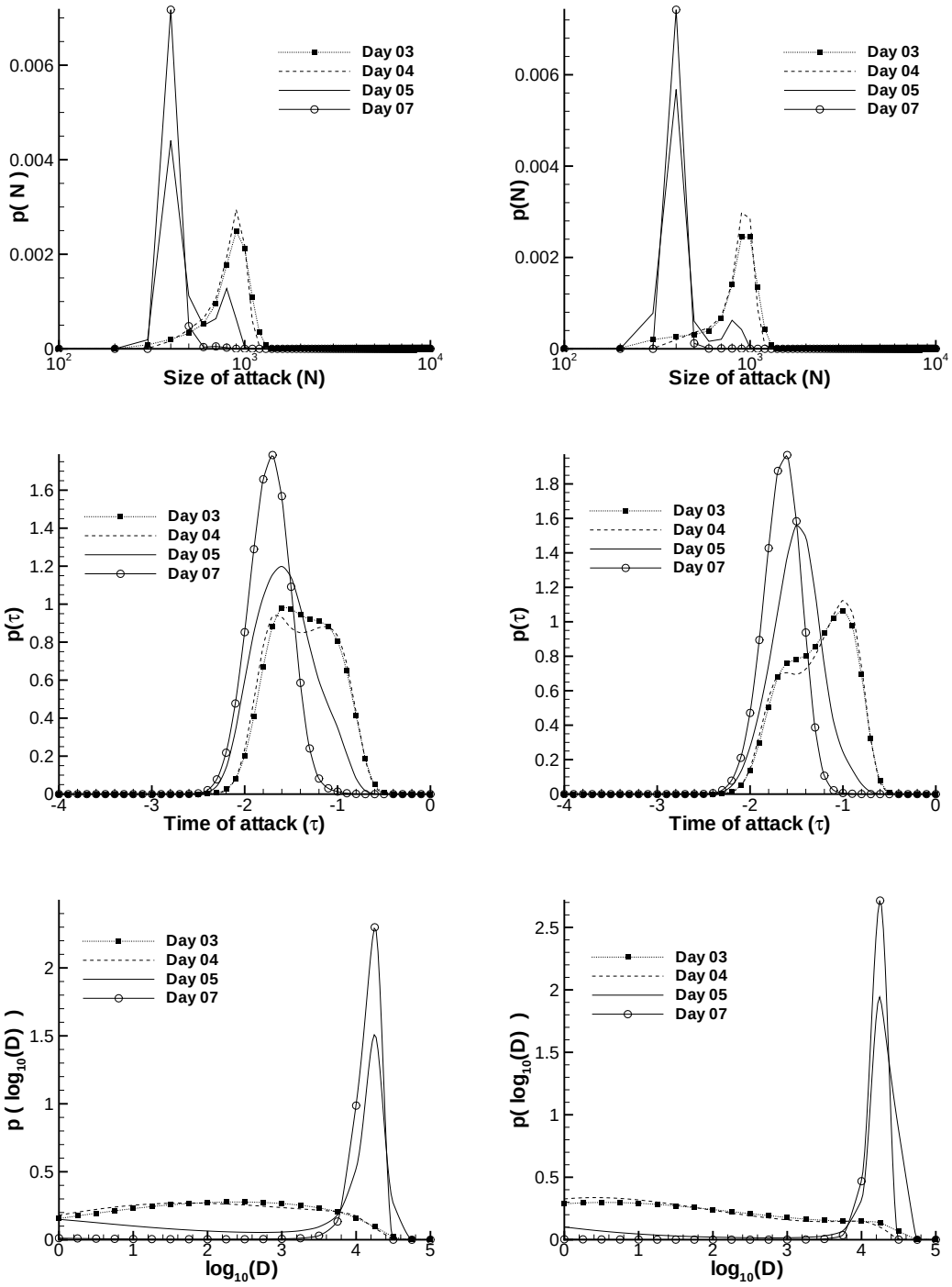

Figure 8. Posterior PDFs for N (top), τ (middle), and $\log_{10} D$ (bottom) based on the time series for Simulation II, as tabulated in Tables III and IV. Lower-resolution data (collected in 24-hour intervals) yield the PDFs on the left, while higher-resolution data yield the PDFs on the right. Correct values for $\{N, \tau, \log_{10}(D)\}$ are $\{454, -1.5, 4.13\}$, where the "correct" representative dose is taken to be $\log_{10}(D_{50})$. In both simulations, PDFs are reported after 3-, 4-, 5-, and 7-day observational periods.

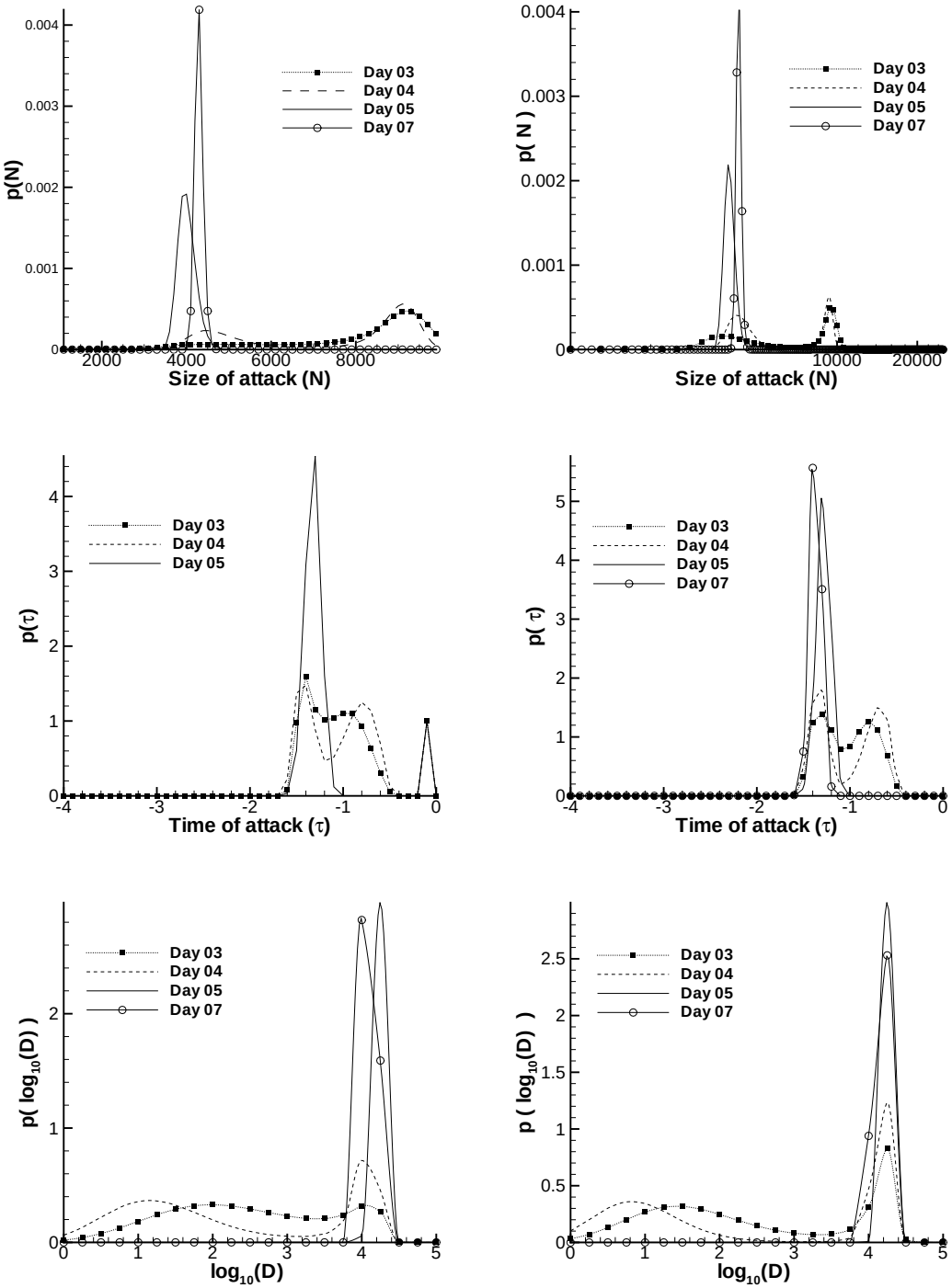

Figure 9. Posterior PDFs for N (top), τ (middle), and $\log_{10} D$ (bottom) based on the time series for Simulation IIa, as tabulated in Tables III and IV. Lower-resolution data (collected in 24-hour intervals) yield the PDFs on the left, while higher-resolution data yield the PDFs on the right. Correct values for $\{N, \tau, \log_{10}(D)\}$ are $\{4537, -1.25, 4.09\}$, where the "correct" representative dose is taken to be $\log_{10}(D_{50})$. In both simulations, PDFs are reported after 3-, 4-, 5-, and 7-day observational periods.

III and IIIa. Indeed, $p(\log_{10}(D))$ in the low-dose small-N Simulation IIIa remains broad at all times. In Simulation III, after only 3 days of data, we observe a dual characterization of the outbreak: $N \approx 700$ and, to a larger extent, $N \approx 2000$. However, $p(N)$ becomes unimodal as additional data become available. In fact, PDFs for all three parameters in all four simulations are unimodal by Day 5. The resulting MAP estimates and 90% CIs for $\{N, \tau, \log_{10}(D)\}$ are reported in Table VI. In contrast to Sec. 3.1, MAP estimates for N and τ are not within 20% of the true values. With the exception of Simulation IIIa, N is smaller than it should be, and in all simulations τ is more negative than it should be.

A qualitative explanation for these discrepancies is advanced as follows. Since Model A2 predicts shorter incubation periods than Model D (recall Fig.1), the epidemic curve as simulated with Model D will rise more slowly that predicted by Model A2. When these data are interpreted using Model A2, it is reasonable to expect the posterior to compensate for the slower rise by underestimating N, i.e., by suggesting a smaller outbreak. Simultaneous estimation of D and τ raises a few additional complications, however. Recall that the posterior of D is centered quite close to its true value in Simulations IV and IVa, and to a lesser extent in Simulation III. But in the likelihood function, this dose enters the wrong model. Using a "correct" dose in Model A2 is akin to using a much larger dose in Model D; both situations yield shorter incubation periods. Now draw a parallel with Simulation F in Sec. 2.3. There, we found that a large-dose small-N attack and a small-dose large-N attack gave rise to very similar patient data during the first five days of an outbreak. Moreover, we found that N and τ were positively correlated (and that both were negatively correlated with D): the small-N mode of the posterior also favored more negative τ, i.e., attacks that occurred approximately one day earlier. The very same correlations affect inference in the present simulations. Incubation model mismatch is roughly equivalent to an overestimation of D, which is compensated for by underestimating N and τ.

In summary, Table VI shows that MAP estimates for N are typically within a factor of two below the true result and that τ is estimated roughly a day too early.

4. The Sverdlovsk Anthrax Outbreak of 1979

We now address the characterization of the Sverdlovsk anthrax outbreak. It is suspected that on 2 April 1979, a high-grade anthrax formulation was accidentally released from a military facility in Sverdlovsk (today, Yekaterinburg), Russia. The resulting outbreak lasted 42 days, and patient data were collected on a daily basis [5]. Characterizing the Sverdlovsk case presents significant challenges. It corresponds to a low-dose "attack" infecting fewer than 100 people. Wilkening [15] estimates that the average dose was either around 2–3 spores, based on his Model A, or around 300 spores based on his Model D; Meselson [5] estimates 100–2000 spores as the likely dose. The first patient presented symptoms on 4 April 1979. Around 12 April, tetracycline was administered around Sverdlovsk; around 15 April, people were vaccinated. Prophylactic measures may have prevented the onset of symptoms in some people and increased the incubation period in others. Further, the available data almost certainly contain some recording errors. Errors in the data, the effect of prophylaxis (which is not modeled in our likelihood function), and the small size of the infected population are expected to stress our inference procedure.

In Fig. 12 we plot the posterior densities of N and τ based on the data in [5]. Model A2 is used for inference. After 9 days of data, the time of release was easy to infer: the MAP estimate of τ is -2 (i.e., 2 April 1979) and the 90% CI for τ is $[-3.22, -1.38]$. PDFs for the dose (omitted here) were indeterminate; the 90% CI for $\log_{10}(D)$ spans $[0.18, 3.5]$, and moreover the average dose at Sverdlovsk is unknown. The MAP estimate for N centers around 50, though the earlier

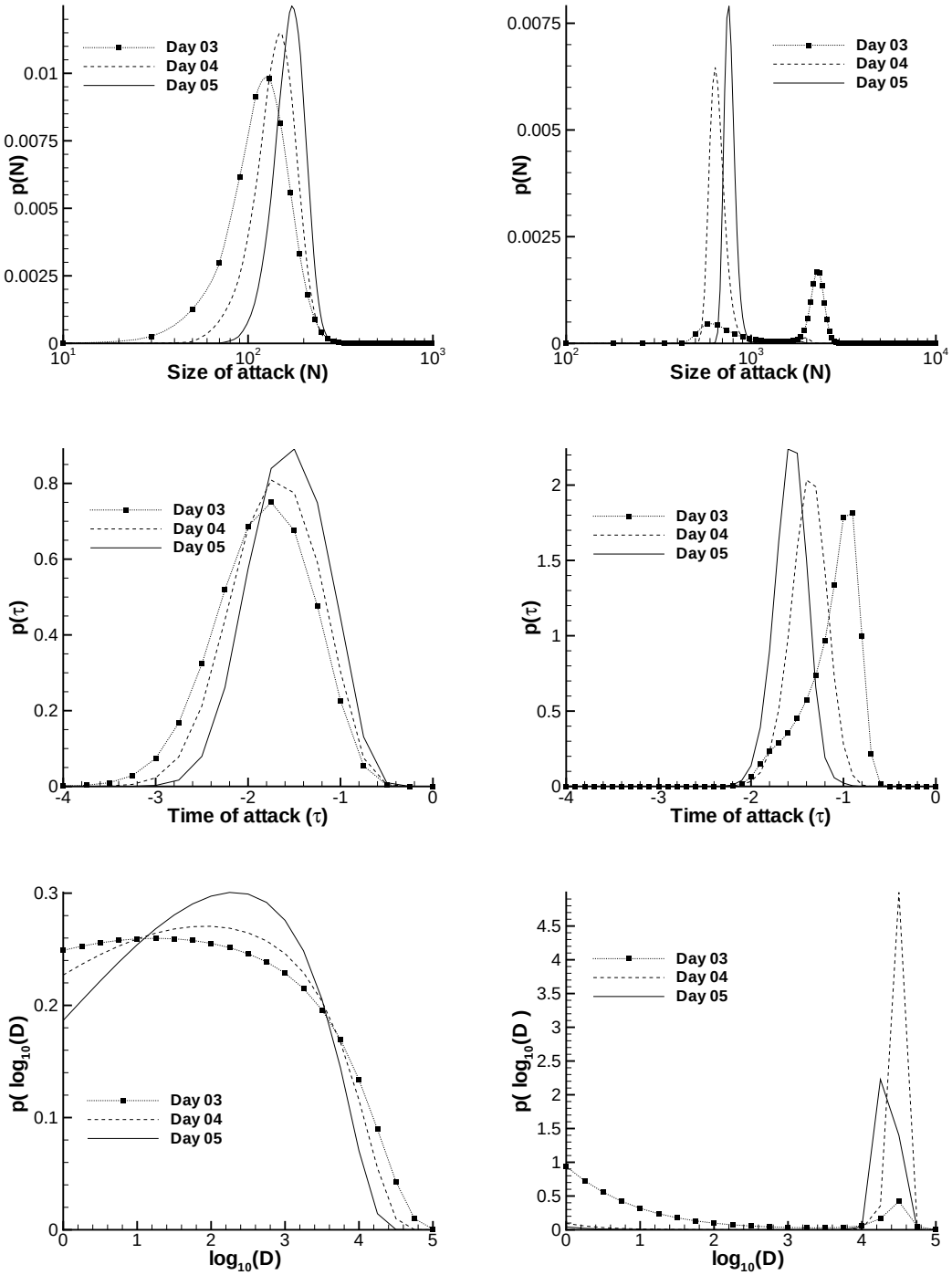

Figure 10. Posterior PDFs for N (top), τ (middle), and $\log_{10} D$ (bottom) based on daily time series for Simulation IIIa (left) and Simulation III (right). Correct values for $\{N, \tau, \log_{10}(D)\}$ are $\{161, -0.75, 3.52\}$ (Simulation IIIa) and $\{1453, -0.75, 3.54\}$ (Simulation III), where the "correct" representative dose is taken to be $\log_{10}(D_{50})$. In both simulations, PDFs are reported after 3-, 4-, and 5-day observational periods (dotted, dashed and solid lines respectively).

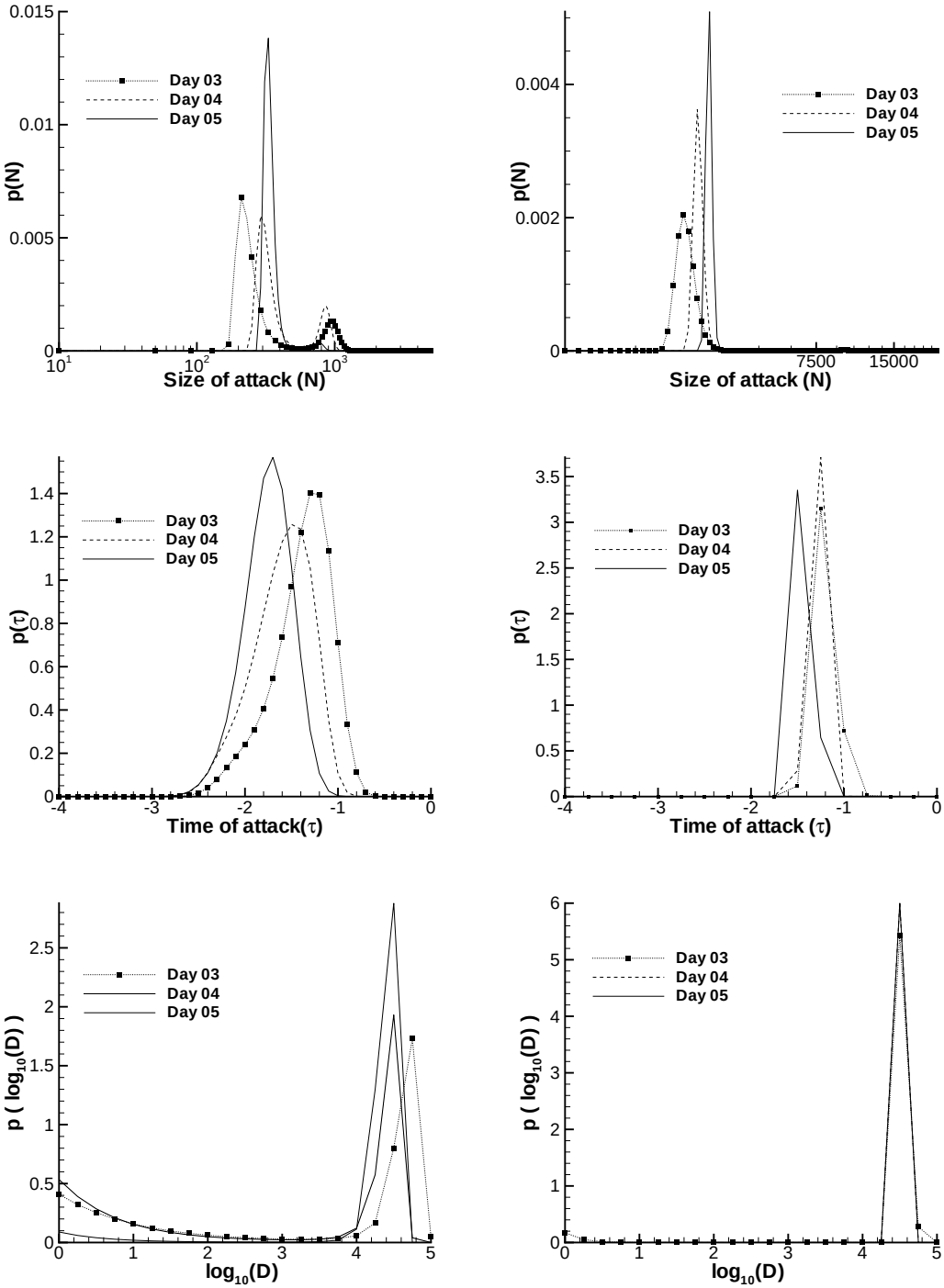

Figure 11. Posterior PDFs for N (top), τ (middle), and $\log_{10} D$ (bottom) based on daily time series for Simulation IV (left) and Simulation IVa (right). Correct values for $\{N, \tau, \log_{10}(D)\}$ are $\{453, -0.75, 4.22\}$ (Simulation IV) and $\{4453, -0.5, 4.20\}$ (Simulation IVa), where the "correct" representative dose is taken to be $\log_{10}(D_{50})$. In both simulations, PDFs are reported after 3-, 4-, and 5-day observational periods (dotted, dashed and solid lines respectively).

Simulation	N	τ	$\log_{10}(D)$
IIIa (6-hr resolution)	170, (130.1 – 243.6)	-1.5, (-2.3 – -0.86)	2.0, (0.23 – 3.74)
IIIa (24-hr resolution	170, (125.1 – 238.8)	-1.5, (-2.4 – -0.94)	2.5, (0.255 – 3.78)
	{161}	{−0.75}	{3.52}
III (6-hr resolution)	780, (722 – 945.5)	-1.7, (-2.03 – -1.42)	4.25, (4.02 – 4.723)
III (24-hr resolution	760, (701 – 891.7)	-1.6, (-1.91 – -1.31)	4.25, (4.04 – 4.724)
	{1453}	{−0.75}	{3.54}
IV (6-hr resolution)	330, (297.2 – 668.6)	-1.7, (-2.23 – -1.40)	4.5, (1.4 – 4.72)
IV (24-hr resolution)	330, (296.3 – 705.3)	-1.7, (-2.26 – -1.38)	4.5, (1.45 – 4.72)
	{453}	{−0.75}	{4.22}
IVa (6-hr resolution)	2900, (2728 – 3056)	-1.5, (-1.90 – -1.1)	4.5, (4.275 – 4.725)
IVa (24-hr resolution)	2900, (2741 – 3064)	-1.5, (-1.97 – -1.26)	4.5, (4.275 – 4.725)
	{4453}	{−0.5}	{4.20}

Table VI. Simulations III, IIIa, IV, and IVa: MAP estimates and the 90% credibility intervals (in parentheses) for N, τ, and $log_{10}(D)$ conditioned on data through Day 5. Correct values for N and τ are in { }. The "correct" representative dose is taken to be $\log_{10}(D_{50})$, also in { }.

PDFs underestimate N. The 90% CI for N after 9 days of data is $[41.15, 66.49]$. Thus by 13 April (i.e., Day 9, the start of the prophylaxis campaign and 2 days before the vaccination campaign), the PDF of N strongly suggests that the outbreak will affect fewer than 200 people. In comparison, 70 people are believed to have died [5, 6] and 80 are believed to have been infected [6], though the true number is unknown. However, approximately 59,000 people in the Chkalovskiy *raion* were impacted by the medical interventions; 80% were vaccinated at least once [5].

Guillemin [30] documents the public health response undertaken by the Soviet authorities once the Sverdlovsk epidemic was detected, illustrating the difficulties and pitfalls faced by medical responders when the origin and the extent of an epidemic are unknown. Indeed, Soviet authorities held that the epidemic was caused by infected meat and spent considerable effort searching for it. The response also engaged many medical personnel and officials from outside Sverdlovsk. Yet by Day 4 (8 April 1979) it was clear that the epidemic was small (Fig. 12, left) and could be handled by local authorities. (Sverdlovsk was a military-industrial city with a population of 1.2 million [5].) Guillemin [30] also describes efforts to decontaminate buildings and trees by hosing them down with disinfectants; yet with knowledge of τ (Fig. 12, right) and meteorological conditions, the bounds of the affected region could have been established (as Meselson did in 1994 [5]) and the public health response suitably targeted. A quantitative model and an inferential capability could therefore have been of assistance in 1979. These lessons are equally applicable to contemporary bioterror scenarios.

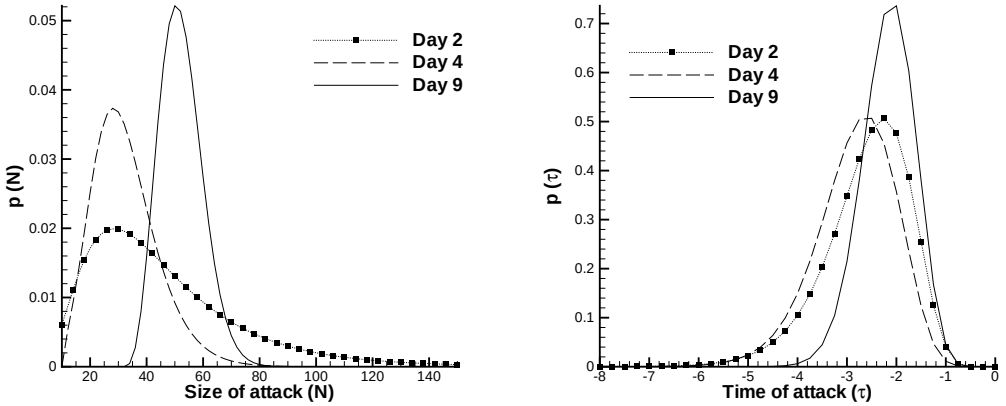

Figure 12. PDFs of N (left) and τ (right) for the Sverdlovsk outbreak.

5. Optimal Allocation of Resources During an Anthrax Attack

In this section we demonstrate how the posterior distribution of the release parameters (as calculated in the previous sections) may be used to allocate scarce resources in the aftermath of an attack. Note that we do not seek to promote a particular resource allocation policy; rather we wish to show, in general, how a probabilistic characterization of a BT attack may be used to allocate resources and hedge against misallocation due to uncertainty in future demands. For the purposes of our study, "resources" will refer to medical personnel and equipment that are difficult to gather and transport, rather than easily transportable products such as antibiotics. Thus the resources of interest are those that are necessary for treating symptomatic patients, e.g., those required to perform thoracentesis and similar surgical procedures that were widely required in the 2001 anthrax attacks [1, 31]. These are usually not available in great excess in most urban centers, and we target situations involving catastrophic attacks where mobilization of national or regional resources may be required. We do not consider resources required for the prophylactic treatment of people who may have been exposed, where antibiotic treatment alone can be highly effective. Further, we will pay particular attention to the *schedule* of allocation of the resources rather the size of the total allocation. Since the total demand for resources is uncertain, it is likely that resources will be supplied in steps over the duration of an epidemic rather than granted outright.

Overall, our approach uses the posterior distribution to pose a problem of optimization under uncertainty. With any number of days of data, the joint posterior of N, τ, and $\log_{10}(D)$ is given by Eq. 2. We simulate M samples from this distribution to yield an ensemble of attack parameters, and for each member of the ensemble—each triple $(N, \tau, \log_{10}(D))$—we then simulate Q realizations of the *future* patient time series $\{t_i, n_i\}$, i.e., the number of people who become symptomatic on a daily basis. The entire set thus comprises $K \equiv MQ$ samples from the posterior predictive distribution of the epidemic curve. An example is given in Figure 13. The left side shows 900 samples from the posterior predictive distribution, following three days of data; the right side shows sampled epidemic curves following seven days of data. These ensembles reflect uncertainty in the scope of the outbreak that an emergency manager would have to address. Note that longer observation of an outbreak (e.g., seven days of data) results in a narrower posterior and a corresponding reduction in the predicted spread of scenarios.

A key task facing the emergency manager is to allocate scarce resources in order to minimize

the number of deaths in a population that has been attacked. To demonstrate an allocation methodology in this context, we will make a number of simplifying assumptions. More complicated mortality-rate models and additional constraints, beyond those considered below, are certainly possible; indeed, a large number of additional constraints would likely specify any specific response scenario. However, these constraints can be added without changing the overall structure or principle of our optimization problem. We discuss a few of these in Sec. 5.4.; see [32] for a more complete consideration of these issues along with more computational results. For the current demonstration, we adopt the following assumptions and notation:

1. There has been a single attack, i.e., one release of the anthrax spores on one target.
2. We have δ days of data. To be clear, the datum on each day is the number of people who became symptomatic on that day. Patients may not come to the hospital on the first day they become symptomatic; the day of initial symptoms is elicited from the patients when they arrive. Given these data, we sample the posterior predictive distribution to generate K scenarios, corresponding to the number of people who will arrive at the hospital for the next T days. Denote these predictions by N_k^j, which is the number of people who become symptomatic on day j in scenario k.
3. Patients symptomatic on day j who have access to medical personnel and equipment are assumed to have a mortality rate h_j. Patients who do not have access to 'resources' (for instance, patients without access to procedures like pleural drainage) are assumed to have a higher mortality rate $u_j > h_j$.
4. Assume $h_j < h_{j-1}$. In other words, those who show symptoms later in the attack have a better chance of successful treatment, because having heard news of the attack, they will tend to seek treatment sooner once they start to feel ill. This assumption is useful to eliminate multiple optimal solutions to the problem posed in Sec. 5.1. In practice, successive h_j may differ only by a small ϵ.
5. Similarly, assume $u_j < u_{j-1}$.
6. Let r_j be the allocation of resources on day j and let R be the total quantity of resources available.
7. Assume that one unit of resource can be used to treat one person.
8. For simplicity, assume no delays in shipping resources to the target.

We determine the best allocation of resources $\{r_j\}$ using a "least-regret" [33] optimization strategy. The strategy consists of two phases, outlined as follows. First, with δ days of data, we simulate K scenarios from the posterior predictive distribution. Each scenario k^*, e.g., a set $\{N_k^j\}$ for $k = k^*$, is a possible time series of future symptomatic patients. Supposing that scenario k^* represents actual knowledge of the future, we can compute the optimal allocation of resources for $t > \delta$ and the corresponding time series of patient deaths. We repeat this computation for all K scenarios, thus solving K optimization problems. These optimization problems are precisely formulated in Sec. 5.1. In the second phase, we must determine a single allocation schedule $\{r_j\}$ given posterior uncertainty in the patient time series. This overall allocation schedule may not perform as well, for each individual scenario, as the corresponding optimal allocation schedule calculated in the first phase. But we seek to minimize the *expected* deviation between the minimum number of deaths in each scenario (calculated in the first phase) and the actual number of deaths given $\{r_j\}$. This "least-regret" optimization problem is precisely formulated in Sec. 5.2.

The least-regret procedure yields a complete allocation schedule $\{r_j\}$ for the entire time interval of interest. The entire schedule can be used for estimating future transportation and resource requirements; one "ships" the amount specified for the current day. On the next day, the posterior distribution is updated with data from an extra day's observation of the epidemic. The optimization procedure described above is repeated, with R reduced by the size of the allocation just made. Since we expect the posterior predictive distribution to narrow with time,

the range of optimal allocations will also narrow and the least-regret solution should improve.

5.1. Optimal Resource Allocation

Given scenario k and predictions N_k^j, we wish to minimize the number of deaths D_k by computing an allocation schedule r_j^k. Using the model described above, the number of patients showing symptoms on day j who will ultimately die in spite of treatment is $h_j r_j^k$; the number who will die untreated is $u_j(N_k^j - r_j^k)$. Resource allocations must be non-negative and their sum cannot exceed R. In addition, we do not want to supply more resources than can be used. Thus $r_j^k \leq N_k^j$. We can now write our resource allocation problem as follows:

$$\begin{array}{lll} \underset{r_1^k \ldots r_T^k, D_k}{\text{minimize}} & D_k & \\ \text{subject to:} & \sum_{j=1}^T r_j^k & \leq R \\ & r_j^k & \geq 0 \\ & r_j^k & \leq N_k^j \\ & D_k & = \sum_{j=1}^T h_j r_j^k + \sum_{j=1}^T u_j(N_k^j - r_j^k). \end{array} \tag{8}$$

This is a standard linear programming problem that can easily be solved by many existing algorithms. The number of variables is $T + 1$ and will typically be less than 50. There are $2T + 1$ inequality constraints and 1 equality constraint. Thus, even though there may be a large number (K) of these problems to solve, each problem is tiny and the entire set can be solved very quickly. Let D_k^* denote the optimal number of deaths obtained for the k^{th} scenario. Also, let

$$r_j^{\max} = \max_k \{r_j^k\}.$$

This is the maximum allocation under any scenario on day j. We will use it as a constraint in the least-regret problem below.

5.2. Least-Regret Allocation

As noted above, the least-regret problem is to find a resource allocation schedule that minimizes the deviation between the optimal number of deaths and the actual number of deaths, averaged over the posterior predictions. Using a squared error to measure the deviation, we obtain the following:

$$\begin{array}{lll} \underset{r_1 \ldots r_T, D_1 \ldots D_K}{\text{minimize}} & \frac{1}{2}\sum_{k=1}^K (D_k - D_k^*)^2 + \sigma \sum_{j=1}^T r_j^2 & \\ \text{subject to:} & \sum_{j=1}^T r_j & \leq R \\ & r_j & \geq 0 \\ & r_j & \leq r_j^{\max} \\ & D_k^j & \geq h_j N_k^j \\ & D_k^j & \geq h_j r_j + u_j(N_k^j - r_j) \\ & D_k & = \sum_{j=1}^T D_k^j \end{array} \tag{9}$$

The derivation of the constraint equations above is slightly non-intuitive. In the course of solving the optimization problems, it becomes necessary to evaluate a broad range of resource allocation schedules. For some of these schedules, there may be a certain day j when resources are over-allocated—that is, $r_j > N_k^j$ for some scenario k. Indeed, the optimal solution may well over-allocate for some scenarios, yet still be the best overall schedule. In such a case, the casualties on day j for scenario k are given by

$$D_k^j = \begin{cases} h_j N_k^j & \text{if } r_j > N_k^j \\ h_j r_j + u_j(N_k^j - r_j) & \text{if } r_j \leq N_k^j \end{cases} .$$

Here D_k^j is the number of people exhibiting symptoms on day j (in scenario k) who will ultimately die. Such a piecewise-linear expression can lead to a difficult optimization problem. Instead, we incorporate the daily deaths as a constraint in the optimization problem

$$\begin{aligned} D_k^j &\geq h_j N_k^j \\ D_k^j &\geq h_j r_j + u_j(N_k^j - r_j) \end{aligned}$$

and evaluate the deaths in a scenario k as the sum of daily deaths, i.e., $D_k = \sum_{j=1}^{T} D_k^j$. A quick check will show that the first expression will prevail when $r_j > N_k^j$ and the second otherwise. We also include a cost of resource allocation $\sigma \sum_{j=1}^{T} r_j^2$ in the objective function (else, in the absence of any penalty for overallocation, the optimization process has no reason not to allocate all the available resources to alleviate the current attack). Depending upon the value of σ, one may conserve resources to a lesser or greater degree. This feature may be useful in several contexts; these are discussed further in Sec. 5.4. The resulting optimization problem is far larger than the one in Sec. 5.1, but it is a convex quadratic program and thus relatively easy to solve.

5.3. Results

Now we apply the least-regret formulation to determine a resource allocation schedule given uncertainty in the characterization of an outbreak. We will consider two circumstances, one in which the resources are grossly insufficient and another in which they are available in excess. In both circumstances, we will evaluate the posterior distribution and calculate the resulting allocation schedule at two points in time during the course of the epidemic.

We simulate an anthrax attack using the approach described in Appendix I. 22384 individuals are infected with a range of doses (average dose: 1470 spores). The first symptoms appear a day after infection. The (daily) time-series of symptomatics for the first 10 days is $\{3, 123, 719, 2046, 2202, 2194, 2058, 1918, 1656\}$; note that this reflects the epidemic curve obtained from a population that received a range of doses, as opposed to a uniform "representative" dose. We then simulate from the joint posterior distribution of $(N, \tau, \log_{10}(D))$, given in Eq. 2, using Markov chain Monte Carlo (MCMC). In particular, we use a single-component random-walk Metropolis sampler [34]. Convergence of the MCMC chain was verified by comparing marginal densities for N, τ and $\log_{10}(D)$ against those developed by the VEGAS algorithm. The MCMC chain was thinned, after some burn-in, to yield $M = 100$ $(N, \tau, \log_{10}(D))$ samples. Each of these samples was used to generate $Q = 10$ scenarios, resulting in a total of $K = 1000$ posterior predictive samples. Eq. 7 (Model A2) was used for the incubation period distribution. We use $h_0 = 0.45$ (5 out of the 11 patients infected in the 2001 anthrax attacks died [31]) while u_0 is set to 0.85 (there were 70 casualties in Sverdlovsk, while 80 were detected with infection [5,6]). $\epsilon = 10^{-6}$. The posterior distribution was simulated after 3 days and after 7 days of data. The linear and quadratic programming problems in Sec. 5.1 and 5.2 were encoded in `AMPL` [35] and solved using the `SNOPT` optimization routine [36]. As described above, we consider conditions of both scarce and excess resources. The former corresponds to $R = 10^4$, less than half of the required amount, while the latter corresponds to $R = 5 \times 10^4$, i.e., excess by roughly a factor of two.

In Fig. 13 we plot the resource allocation schedules obtained after 3 days of data (left) and 7 days of data (right). In the upper row, we show the allocation obtained with excess resources, i.e., $R = 5 \times 10^4$; in the bottom, we show the allocation with scarce resources, $R = 10^4$. In each of the plots we include an ensemble of posterior predictive scenarios as described earlier. Overlaid on the scenarios is the average of the optimal allocations, $\frac{1}{K}\sum_{k=1}^{K} r_j^k$, plotted with a thick solid line. We refer to this as the "average allocation." The actual evolution of the outbreak is plotted with triangles. Two least-regret allocation schedules are plotted—one obtained with $\sigma = 0$ (unfilled circles) and one with $\sigma = 20$ (filled circles). One can easily show

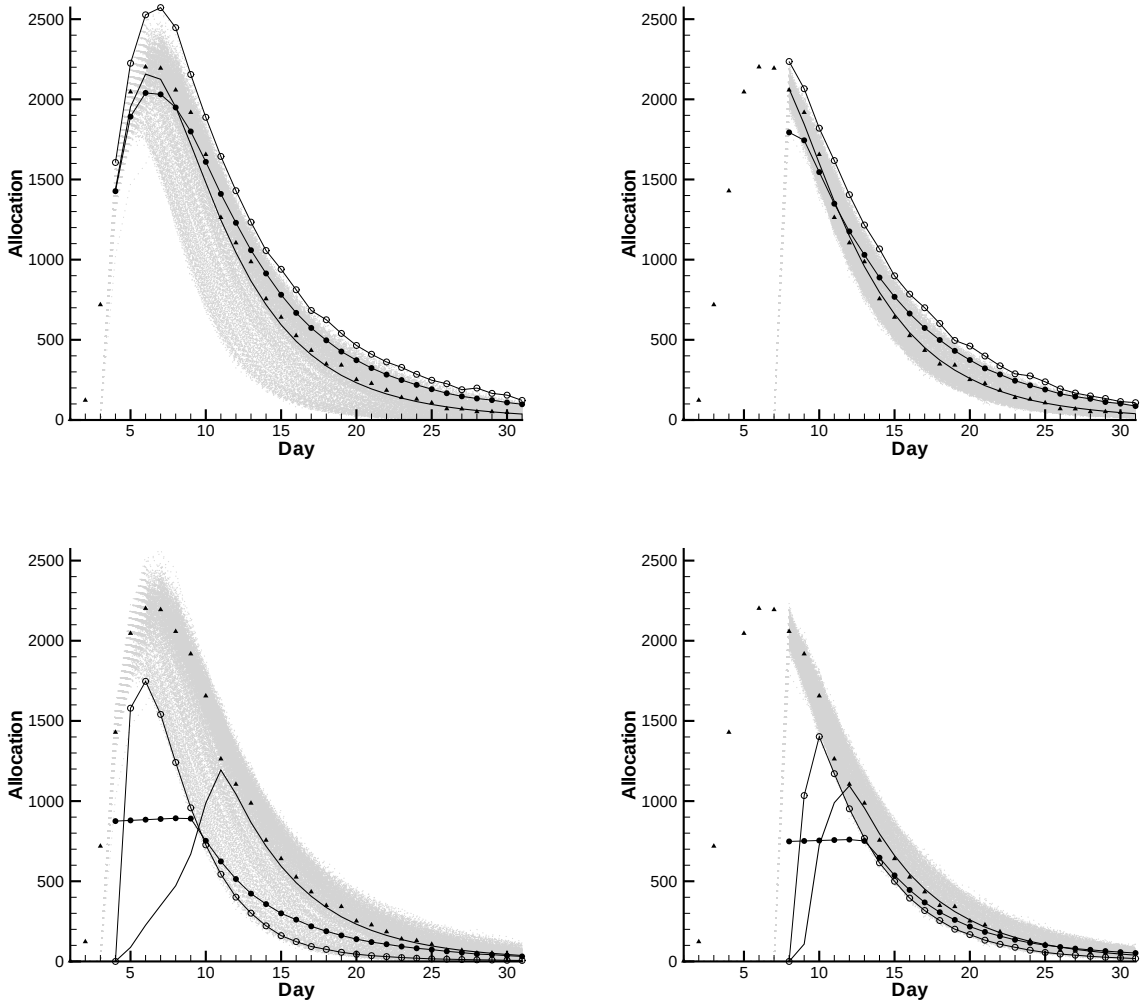

Figure 13. Ensemble of epidemic scenarios (gray region) obtained by sampling the posterior distribution of $(N, \tau, \log_{10}(D))$ constructed from 3 days (left) and 7 days (right) of observations. The average of the optimal allocations (called the "average allocation") is plotted as an unmarked solid line. The least-regret allocation calculated with $\sigma = 0$ (i.e., no penalty for over-allocation) is plotted with unfilled circles, while filled circles denote the least-regret allocation with $\sigma = 20$. The actual evolution of the outbreak is plotted using triangles. The number of infected people is $N = 22,384$. For clarity, scenarios in the top and bottom deciles of the ensemble have not been plotted. Top: excess-resources case, $R = 5 \times 10^4$. Bottom: scarce-resources case, $R = 10^4$. The tendency of the method to delay the allocation of resources is clearly visible. The top and bottom 5% of the scenarios were not plotted to reduce clutter in the graphs.

that as σ increases, the solution is completely dominated by the resource penalty and thus the solution tends to $(\sum_j u_j N_k^j - D_k^*)u_j N_k^j/(2\sigma)$ on day j. For intermediate values of σ, however, a lowering/flattening of the peak resource allocation schedule can be expected.

When resources are in excess, following 3 days of data (upper-left plot in Fig. 13), the least-regret allocation schedule with $\sigma = 0$ (unfilled circles) essentially envelops the scenarios, up to a maximum integrated area dictated by R. Since resources are in excess and there is no penalty for over-allocation, the optimization algorithm tries to accommodate all the scenarios and succeeds in doing so. When over-allocation is penalized, the least-regret allocation (filled circles) assigns fewer resources (with respect to the average allocation) in the immediate future, but more resources later. We also see a perhaps fortuitous similarity between the average allocation and the actual evolution of the epidemic. In the upper-right plot, we allocate resources after 7 days of data are available. Due to the more informative posterior, the spread of scenarios is narrower. Otherwise, the behaviors of the two least-regret allocation schedules are qualitatively similar to those observed with 3 days of data.

Perhaps more interesting is the bottom row of plots in Fig. 13, which considers the case where resources are scarce. We see that the average allocation is quite different from the evolution of the scenarios; on a per-scenario basis, the formulation in Sec. 5.1 seeks to allocate resources later in time. The least regret allocation (with $\sigma = 0$; unfilled circles), however, follows the trend of the epidemic approximately, systematically under-allocating every day due to scarcity of resources. Furthermore, with 3 days of data, the schedule assigns no resources to Day 4, again due to scarcity. When a penalty for allocation is introduced ($\sigma = 20$; filled circles) we see a flattening of allocation schedule in the early epoch, followed by a more generous allocation later in the epidemic. Also, resources are immediately allocated on Day 4. While the allocations developed with 3 days of data (bottom left) show a significant discrepancy between allocation and demand, the performance of the allocation schedule improves as more data become available. The plot on the bottom right (7 days of data; scarce resources) shows that the least-regret allocation with $\sigma = 0$ is very close to the actual demand, and that the schedule obtained with $\sigma = 20$ is not very far behind.

In Fig. 14 we compare the predicted performance of the least-regret allocation schedules. In the absence of perfect knowledge of how the epidemic would evolve beyond the current day, the allocation seeks to guard against the probability of outlier scenarios yielding a large set of casualties. The overall impact of the two choices of allocation (least-regret versus the simple average) can be evaluated by considering the number of deaths that occur in each case. These are plotted as PDFs of casualties *in excess* of those caused by the optimal allocation for each scenario, when resources are scarce. The PDFs were calculated empirically via kernel density estimation from 1000 samples of $(D_k - D_k^*)$ obtained from our scenarios, with allocations computed with 3 and 7 days of data. The average allocation performs quite poorly, in that it yields a very broad, long-tailed PDF—indicating a significant probability of extremely bad outcomes. The least-regret allocations, on the other hand, result in PDFs that are concentrated at the low end of the scale with fairly compact support, thus reducing the probability of excess casualties. Introducing the penalty $\sigma > 0$ broadens the PDF and slightly increases the probability of larger casualties; however, it remains smaller than that due to the average allocation. In Fig. 14 (inset), we plot the PDFs following 7 days of data. They are qualitatively similar to their 3-day counterparts, except that more information allows a better allocation of resources, further reducing the probability of excess casualties. Indeed, the support of the $\sigma = 0$ PDF extends only to 40 deaths in excess of the per-scenario optima.

Overall, it appears that the least-regret approach is far more robust and flexible than the average allocation approach. It better mitigates excess casualties and provides a means of smoothing allocations while reducing the expected number of deaths.

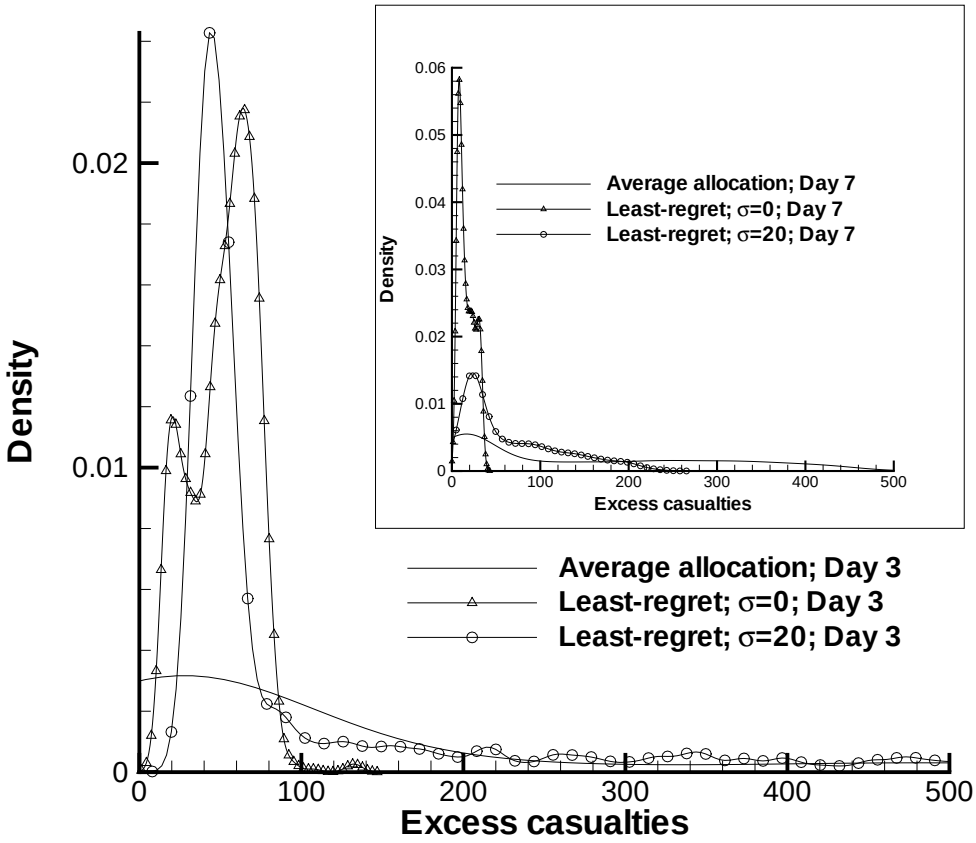

Figure 14. PDFs of "excess" casualties—over the per-scenario optimal allocation of resources—when resources are scarce. Resource allocations were performed using 3 days of data. We see that the average allocation results in a low but broad PDF, resulting in a significant probability of a large number of deaths over the optimal number. The least-regret allocation sharpens the PDFs significantly, thus reducing the probability of an extremely bad outcome. Inset: PDFs of excess casualties after 7 days of data. Compared to the allocations following 3 days of data, these densities are far sharper and clustered at a smaller number of excess deaths. The effect of $\sigma > 0$ is to widen the range of the PDF, resulting in a larger expected number of deaths.

5.4. Extensions

The approach described above can be easily extended to accommodate more realistic constraints. These are discussed in [32], where substantial testing has also been performed. Extensions include:

Stochastic optimization model. Other optimization models can be considered, including more standard stochastic optimization with recourse models. In addition, other objective functions can be formulated. Stochastic optimization models can rapidly become extremely large, so we will probably be limited to only a few stages. See [32] for preliminary results on how stochastic optimization may be used with the current formulation within the context of resource allocation for two time-staggered BT attacks.

Reuse of resources. Inevitably, a certain percentage of patients who receive treatment will die, presumably at a particular rate. One can make more realistic use of the resources if the possibility of resource reuse is taken into account. This constitutes a linear constraint on the available resources and thus does not change the character of the optimization problems. Similarly, we did not specifically take into account that a patient arriving on day j may receive treatment on day $j' > j$ if the doctors think that this is be best use of the resources. We do not think that this will have a significant effect on the results.

Multiple attacks. An important situation facing an emergency manager is the possibility of multiple attacks. Although, from a certain point of view, it is immaterial whether a life is saved in the first target or the second, there may be "social" constraints dictating that both attack sites must get a proportional amount of the resources. One can devise an "equilibration of pain" constraint that allows such an allocation. This constraint is also linear and requires that the relative number of infected people treated at each target should be approximately the same. Recall that the use of the parameter σ in the objective function of Eq. 9 can be used in several ways. For example, although it could be used to model a logistics constraint, it could also be used to limit the allocations in the first attack as a hedge against a subsequent attack. Our results show immediately what the effect of such a strategy would be on the total number of expected deaths.

Shipping delays. It is easy to constrain the optimal resource allocation problem (Sec. 5.1) so that shipped resources reach the target after a fixed delay. In fact, the general resource allocation problem, studied extensively in the operations research literature, can handle multiple sources of resource and multiple destinations, each with different costs and times of delivery. These problems are typically linear programming problems and can also be easily solved and used within our least-regret framework.

Effect of delayed prophylaxis One could also include the effect of delaying the treatment of patients, i.e., treating patients showing symptoms on day j on day $j + 1$ because of a lack of resources. This requires a model for the (reduced) chances of recovery, given a delay in treatment, created, perhaps, based on the discussion and data in [37].

6. Conclusions

We have developed a Bayesian approach to characterize BT attacks from a time series of diagnosed patients. Our tests with anthrax show that an observation period of 3–5 days may be sufficient to estimate the number of asymptomatic infected people, the time of infection, and a representative dose, and to provide quantified uncertainty intervals around these estimates. Note that since the data consist of the times of symptoms of diagnosed patients, the inference can be performed only after the first patient is successfully diagnosed. Sensitivity studies suggest that when the disease model is not accurate, we may arrive within a factor of two of the size of the attack. The resolution of the time series of diagnosed patients has a small impact if the disease model is accurate; otherwise, model errors dominate.

This Bayesian approach may be extended and improved in many ways. One extension is to include hospital visit delay times in the analysis. While we have used the time at which each patient became symptomatic to demonstrate our method, existing medical informatics systems typically record the time of clinical encounter instead. The difference between the two—the hospital visit delay—can be modeled in a Bayesian manner (see [38] for a model of the hospital visit delay for inhalational anthrax) and included in the analysis. This would capture the effect of the patients who have turned symptomatic, but have not yet reported to the hospitals, on the inferred quantities. However, it would introduce an additional source

of variability and require further sensitivity analysis. Improved medical informatics systems that capture the time of symptoms would make our methodology more useful, eliminating the variability introduced by hospital visit delays. A second source of delay — the time required to perform tests, etc., that lead to a confirmed diagnosis — can be another source of uncertainty, but could be modeled in a manner similar to the hospital visit delay and included in the inference procedure. Another potential shortcoming of our approach is our inability to infer a distribution of doses (commensurate with a spatially distributed population and atmospheric dispersion); instead we rely on inferring a representative dose. However, the impact of this "shortcoming" on the accuracy of the inference is unclear; the tests performed by Legrand *et al.* [22] between their spatiotemporal inference technique (which, of course, included a fully spatially distributed dosage pattern) and an older, single-representative-dose version of our current approach [23] revealed no major differences in the quality of the inference. However, unlike [22], our purely temporal approach cannot infer certain, mostly spatial, characteristics of an aerosolized BT attack, e.g., the location and height of a release.

The ability to "fuse" disparate sources of data via prior distributions contributes significantly to the robustness of Bayesian inference in data-starved environments. Informative prior distributions for N and τ, drawn from syndromic surveillance data, may increase the efficiency of the inference process. Also, the present approach can immediately be applied to other noncontagious diseases, as well as to contagious diseases with long incubation periods, such as smallpox, where secondary cases do not appear in the early time series of patient data. If secondary cases are included, one could extend the current formulation into a non-stationary Bayesian inference problem on a social/transportation network.

The importance of quantitatively characterizing a BT attack was explicitly identified in the "Dark Winter" exercise [3]. ("Dark Winter" was a war-game/table-top exercise conducted in 2001 to assess the ability of key decision makers, e.g., the President of the United States, etc., to respond to a smallpox BT attack. Participants included, among others, The Hon. Sam Nunn, the former senator, The Hon. R. James Woolsey, the former Director of the Central Intelligence Agency and General J. Titelli (U.S.A, retd). The "lessons learned" section of [3] lists the information gaps that the participants faced when formulating the response to the ensuing epidemic. (Estimates of the size of the problem at hand are often crucial in choosing between competing response strategies, and this exercise was no exception.) Participants sought the ability "...to immediately predict the likely size of the epidemic on the basis of the initial cases; to know how many people were exposed." Thus the primary utility of our inference procedure is in the context of a response plan, especially when resources have to be husbanded and used carefully (e.g., if more attacks are feared). To this end, we have demonstrated how a probabilistic characterization—capturing the inevitable uncertainties in the severity and timing of a BT attack—can guide the optimal allocation of scarce resources. Preliminary results on how a similar approach can be used within the context of *multiple* attacks can be found in [32]. Since our inference methodology is purely temporal, data requirements are simple, thus reducing the opportunities for introducing significant measurement errors. For example, a spatiotemporal approach would require the geographical location of each patient *at the instant of infection*; in a mobile population, locations can be a significant source of error since a detailed movement schedule of the infected patients is rarely available. In the absence of such data, the location may be modeled probabilistically, and results from preliminary work by the BARD group can be found in [39, 40]. The same technique was also used by Legrand *et al.* [22] in their investigation into the spatiotemporal inference of BT attacks. A discussion of the ambiguities and difficulties introduced by location data, within the context of the Sverdlovsk outbreak, can be found in [30, 41]. Certainly a spatiotemporal approach, correctly applied, has the potential to uncover more information (e.g., spatial information, which can be critical for prioritizing prophylaxis [22], decontamination, etc.) from a partially observed epidemic, but the simpler temporal approach considered here may constitute a more

robust and practical tool for early response and resource allocation.

ACKNOWLEDGEMENTS

The authors acknowledge the help of Dr. Joseph Egan, Health Protection Agency, Porton Down, Salisbury, Wiltshire, UK, who read a draft version of the paper and provided numerous comments and suggestions. We thank Dr. Petri Fast, formerly of Lawrence Livermore National Labortory, Livermore, CA and Dr. Mark Kraus of NORAD-NORTHCOM, Peterson AFB, CO for helpful discussions and preliminary testing of our method. We also thank the referees for their many useful remarks. This work was supported by Sandia National Laboratories' LDRD (Laboratory Directed Research and Development) program, with funds allocated under the "Enabling Predictive Simulation" Investment Area. Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under Contract DE-AC04-94-AL85000.

REFERENCES

1. Jernigan JA, et al. Bioterrorism-related inhalational anthrax: The first 10 cases reported in the United States. *Emerging Infectious Diseases* 2001; **7**(6):933–944.
2. Goldenberg A, Shmueli G, Caruana RA, Fienberg SE. Early statistical detection of anthrax outbreaks by tracking over-the-counter medication sales. *Proceedings of the National Academy of Sciences, USA* 2002; **99**:5237–5240.
3. O'Toole T, Mair M, Inglesby TV. Shining light on Dark Winter. *Clinical Infectious Diseases* 2002; **34**:972–983.
4. Walden J, Kaplan EH. Estimating time and size of bioterror attack. *Emerging Infectious Diseases* 2004; **10**(7):1202–1205.
5. Meselson M, Guillemin J, Hugh-Jones M, Langmuir A, Popova I, Shelokov A, Yampolskaya O. The Sverdlovsk anthrax outbreak of 1979. *Science* 1994; **266**:1202–1208.
6. Brookmeyer R, Blades N, Hugh-Jones M, Henderson DA. The statistical analysis of truncated data: application to the Sverdlovsk anthrax outbreak. *Biostatistics* 2001; **2**:233–247.
7. Brookmeyer R, Blades N. Statistical models and bioterrorism : Application to the U.S. anthrax attacks. *Journal of the American Statistical Association* 2003; **98**(464):781–788.
8. Henderson DW, Peacock S, Belton FC. Observations on the prophylaxis of experimental pulmonary anthrax in the monkey. *Journal of Hygiene* 1956; **54**:28–36.
9. Albrink WS, Goodlow RJ. Experimental inhalational anthrax in the chimpanzee. *American Journal of Pathology* 1959; **35**(5):1055–1065.
10. Gleiser CA, Berdjis CC, Hartman HA, Gochenour WS. Pathology of experimental respiratory anthrax in *Macaca Mulatta*. *British Journal of Experimental Pathology* 1963; **44**:416–426.
11. Friedlander MM, Welkos SL, Pitt ML, Ezzell JW, Worsham PL, Rose KJ, Ivins BE, Lowe JR, Howe GB, Mikesell P, *et al.*. Postexposure prophylaxis against experimental inhalation anthrax. *Journal of Infectious Diseases* 1993; **167**(5):1239–1243.
12. Ivins BE, Pitt MLM, Fellows PF, Farchaus JW, Benner GE, Waag DM, Little SF, Anderson GW, Gibbs PH, Friedlander AM. Comparative efficacy of experimental anthrax vaccine candidates against inhalational anthrax in rhesus macaques. *Vaccine* 1998; **16**(11/12):1141–1148.
13. Brachman PS, Kaufmann AF, Dalldorf FG. Industrial inhalational anthrax. *Bacteriological Reviews* 1966; **30**(3):646–657.
14. Brookmeyer R, Johnson E, Barry S. Modelling the incubation period of anthrax. *Statistics in Medicine* 2005; **24**:531–542.
15. Wilkening D. Sverdlovsk revisted : Modeling human inhalational anthrax. *Proceedings of the National Academy of Science, USA* May 2006; **103**(20):7589–7594.
16. Glassman HN. Discussion on industrial inhalational anthrax. *Bacteriological Review* 1966; **30**:657–659.
17. Druett HA, Henderson DW, Packman L, Peacock S. Studies on respiratory infection. I.The influence of particle size on respiratory infection with anthrax spores. *Journal of Hygiene* 1953; **51**:359–371.
18. Haas CN. On the risk of mortality to primates exposed to anthrax spores. *Risk Analysis* 2002; **22**(2):189–193.
19. Hogan WR, Cooper GF, Wallstrom GL, Wagner MM, Depinay JM. The Bayesian aerosol release detector: An algorithm for detecting and characterizing outbreaks caused by an atmospheric release of Bacillus anthracis. *Statistics in Medicine* 2007; **26**:5225–5252.
20. Tsui FC, Espino JU, Dato VM, Gesteland PH, Hutman J, Wagner MM. Technical Description of RODS: A Real-time Public Health Surveillance System. *Journal of the American Medical Informatics Association* 2003; **10**(5):399–408.

21. Turner DB. *Workbook of Atmospheric Dispersion Estimates: An Introduction to Dispersion Modeling*. Lewis Publishers, CRC Press LLC, 2000 N.W. Corporate Blvd. Boca Raton, FL 33431, 1994.
22. Legrand J, Egan JR, Hall IM, Cauchemez S, Leach S, Ferguson NM. Estimating the location and timing of a covert anthrax release. *Public Library of Science - Computational Biology* 2009; **5**(1).
23. Ray J, Marzouk YM, Najm HN, Kraus M, Fast P. A Bayesian method for characterizing distributed micro-releases: I. The single-source case for non-contagious diseases. *SAND Report SAND2006-1491*, Sandia National Laboratories, Livermore, CA 94551-0969 March 2006. Unclassified unlimited release; available at http://www.caip.rutgers.edu/∼jaray/Inverse.html.
24. Tigertt WD, Benenson AS, Gochenour WS. Airborne q fever. *Microbiology and Molecular Biology Reviews* 1961; **25**(3):285–293.
25. Lepage GP. A new algorithm for adaptive multidimensional integration. *Journal of Computational Physics* 1978; **27**:192–203.
26. GNU Scientific Library. Http://www.gnu.org/software/gsl/.
27. Ray J, Marzouk YM, Kraus M, Fast P. A Bayesian method for characterizing distributed micro-releases: II. Inference under model uncertainty with short time-series data. *SAND Report SAND2006-7568*, Sandia National Laboratories, Livermore, CA 94551-0969 December 2006. Unclassified unlimited release; available at http://www.caip.rutgers.edu/∼jaray/Inverse.html.
28. Sartwell PE. The Distribution of Incubation Periods of Infectious Diseases. *American Journal of Hygiene* 1950; **51**:310–318.
29. Klein F, Walker JS, Fitzpatrick DF, Lincoln RE, Mahlandt BG, Jones WL, Dobbs JP, Hendrix KJ. Pathophysiology of anthrax. *Journal of Infectious Diseases* 1966; **116**(2):123–138.
30. Guillemin J. *Anthrax: The Investigation of a Deadly Outbreak*. University of California Press, 2001.
31. Jernigan JA, et al. Investigation of bioterrorism-related anthrax, United States, 2001: Epidemiological findings. *Emerging Infectious Diseases* 2002; **8**(10):1019–1028.
32. Boggs PT, Gay DM, Ray J. Probabilistics attack reconstruction and resource estimation in reload scenarios. *IEEE Xplore, Proceedings of the 2009 IEEE Homeland Security Conference*, Waltham, MA, 2009. Http://www.caip.rutgers.edu/∼jaray/Presentations/ieee09.pdf.
33. Lempert RJ, Popper SW, Bankes SC. *Shaping the Next One Hundred Years: New Methods for Quantitative Long-Term Policy Analysis*. Rand, Santa Monica, CA 90407, 2003.
34. Gilks WR, Richardson S, Spiegelhalter DJ. *Markov Chain Monte Carlo in Practice*. Chapman and Hall, 1996.
35. Fourer R, Gay DM, Kernighan BW. *AMPL: A Modeling Language for Mathematical Programming*. Second edn., Duxbury Press/Brooks/Cole Publishing Co., 2003.
36. Gill P, Murray W, Saunders M. SNOPT: an SQP algorithm for large scale constrained optimization. *SIAM Review* 2005; **47**:91–131.
37. Holty JE, Bravata DM, Liu H, Olshen RA, McDonald KM, Owens DK. Systematic review: A century of inhalational anthrax cases from 1900 to 2005. *Annals of Internal Medicine* 2006; **144**:270–280.
38. Hogan WR, Wallstrom GL. Approximating the sum of lognormal distributions to enhance models of inhalational anthrax. *Quantitative Methods in Defense and National Security*, Fairfax, VA, 2007. Available at: http://www.galaxy.gmu.edu/QMDNS2007/QMDNS2007-booklet.pdf (Accessed on 30 March 2007).
39. Cami A, Wallstrom GL, Hogan WR. Effect of commuting on the detection and characterization performance of the Bayesian Aerosol Release Detector. *Proceedings of the International Workshop on Biomedical and Health Informatics (BHI 2008)*, Philadelphia, PA, 2008. DOI: 10.1109/BIBMW.2008.4686214.
40. Cami A, Wallstrom GL, Hogan WR. Integrating a commuting model with the bayesian aerosol release detector. *Biosecure 2008, Lecture Notes in Computer Science*, 5354, 2008; 85–96.
41. Guillemin J. The political determinants of delayed diagnosis: The 1979 Sverdlovsk anthrax outbreak and the 1972 Yugoslavian smallpox epidemics. *Confronting Terrorism – 2002*, Gupta R, Perez MR (eds.), Los Alamos National Laboratory, 2002. Http://library.lanl.gov/cgi-bin/getdoc?event=CT2002&document=20.

APPENDIX

I. Methodology for Obtaining a Dose Distribution Consistent with Atmospheric Dispersion over a Geographically Distributed Population

The spatial distribution of dosages due to atmospheric release of an aerosol can be modeled using a simple Gaussian plume model [21]. An atmospheric release typically occurs over a domain with a non-uniform population distribution; we can combine the plume model with the population distribution to calculate the number of people exposed to a given dose. In this section, we describe a simple way to obtain such a population-dosage distribution.

We consider a square domain, L km on each side; in this study, $L = 10$ km. The domain is divided into N blocks per side; here $N = 100$. 25 population clusters are chosen in the form of Gaussian

kernels $A\exp(-r^2/R^2)$, where $r^2 = |\mathbf{x} - \mathbf{x_0}|^2$. The strength of the kernel A, its center $\mathbf{x_0}$, and its length scale R are randomly sampled from independent uniform distributions. The population density in any block, with its center at $\mathbf{x}$, is a sum of the strengths of all the 25 kernels. The strengths of the kernels are scaled to obtain a total population (in the domain) of P_{domain}. The population in a given block is obtained by multiplying the population density with the block area. This creates a geographically distributed population.

The number of people exposed (i.e., who inhaled the aerosol, but may or may not develop symptoms) and infected (i.e., who will develop symptoms) is dependent on the location and size of the release and direction of the wind. We release 10^{13} spores at the origin, at a height of 100 meters. A wind speed of 4 m/s and a Pasquill stability class of "B" are assumed. Pasquill stability classes indicate atmospheric stability; class B indicates a moderately unstable atmosphere with strong daytime insolation. Details of Pasquill stability classes and atmospheric dispersion are in [21]. In our study, wind directions are measured in degrees from due north; that is, a wind direction of zero degrees is a wind from due north, 90 degrees is a wind from the west, and a direction of 180 degrees is a wind from due south. The release is assumed to be an explosive point release, and the concentration of the aerosol at any point (x, y) on the ground and any time t is given by [21]

$$\chi(x,y,t) = \frac{2Q_T}{(2\pi)^{3/2}\sigma_{x'}\sigma_{y'}\sigma_{z'}} \exp\left(-\frac{(x'-ut)^2}{2\sigma_{x'}^2}\right) \exp\left(-\frac{(y')^2}{2\sigma_{y'}^2}\right) \exp\left(-\frac{(H')^2}{2\sigma_{z'}^2}\right) \tag{10}$$

where (x', y') are Cartesian coordinates in a frame of reference where the x'-axis is aligned with the wind. $\sigma_{x'}, \sigma_{y'}$ and $\sigma_{z'}$ are coefficients dependent on x' and on the Pasquill stability class. H is the height of release and χ is the concentration of the aerosol in spores per unit volume. u is the wind velocity. Q_T is the total number of spores released. The relation between $\mathbf{x}'$ and $\mathbf{x}$ is given by

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos(\pi-\theta) & -\sin(\pi-\theta) \\ \sin(\pi-\theta) & \cos(\pi-\theta) \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}$$

where θ is the wind direction. Assuming a breathing rate β of 30 liters a minute, one can obtain an expression for the number of spores inhaled per unit time. Integrating to infinite time, one obtains the total number of spores D inhaled by a person positioned at (x, y) (or at (x', y')):

$$D = \frac{Q_T\beta}{2\pi\sigma_{x'}\sigma_{y'}\sigma_{z'}} \exp\left(-\frac{(y')^2}{2\sigma_{y'}^2}\right) \exp\left(-\frac{(H')^2}{2\sigma_{z'}^2}\right) \left(1 + \text{erf}(x')\right).$$

The dosage assigned to a given block is decided by the location of its center. If we choose Model A2 to simulate the BT attack, we use Glassman's formula to model the probability a of showing symptoms (in infinite time) given a dosage D [16]:

$$a(D) = \frac{1}{2}\left[1 + \text{erf}\left(\frac{\ln(D/D_0)}{S\sqrt{2}}\right)\right] \tag{11}$$

where $D_0 = 8600$ spores and $S = 3.44$. These correspond to a human ID_{50} of 8600 spores and a probit slope of 0.67 [15, 16]. If Model D is chosen instead, we employ Eq. 5 to determine the probability of infection given a dose D. Since the population in a block is known, we can then use the probability of infection to calculate the number of people in the block who will proceed to develop symptoms over time, per the incubation period model.

In this study, we use $P_{\text{domain}} = 3 \times 10^6$ and two plume directions, $\theta = 170^\circ$ and 125°. The two releases result in, respectively, 686,068 and 1,869,741 *exposed* individuals, i.e., individuals who have received a dose of one spore or more. The maximum doses observed in the two cases are 30,877 and 314,053 respectively. The dose range is divided into 100 equal bins and a histogram of the number of people in each bin is developed for each of the cases. The histogram is then normalized to obtain the "exposure" PDF, i.e, the PDF of the dose received by an individual in the exposed population. Given the large population ($P_{\text{domain}} = 3 \times 10^6$), the PDF developed from a histogram with 100 bins is quite smooth. Note that only a fraction of the exposed population will develop symptoms, with an individual's probability of being infected (and subsequently developing symptoms) being given by Glassman's relation (Eq. 11) or Eq. 5.

The "exposure" PDFs developed for $\theta = 170^\circ$ and 125° are then used to sample from a smaller exposed population of p_{exposed} for each of the tests. Values of p_{exposed} and θ used for the different cases are in Table VII. Each exposed individual is then allowed to become infected with a dose-dependent probability. The resulting infected sub-population yields the final dose distribution.

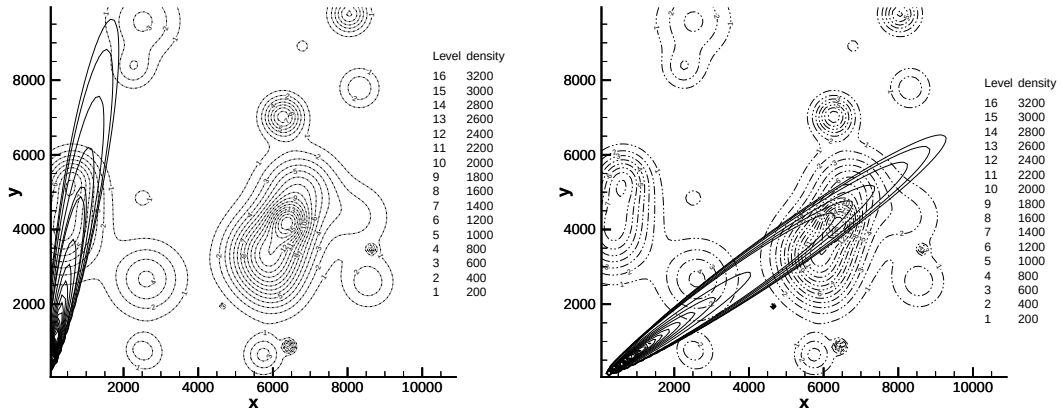

Figure 15. Dosage plumes plotted over the population distribution for $\theta = 170^\circ$ (left) and 125° (right). We see on the right that the extremities of the plume extend into a high population density region. Population density is measured in number of people per square kilometer. Thus we may expect a substantial number of high-dosage cases, resulting in a higher average dosage D.

Table VII. The wind direction, θ, and the size of the exposed population, p_{exposed}, used to generate the infected population in various attacks. For Cases I, Ia, II, and IIa, Eq. 11 is used for the probability of infection, while for Cases III, IIIa, IV, and IVa, Eq. 5 is used.

	$p_{\text{exposed}} = 10^3$	$p_{\text{exposed}} = 10^4$
$\theta = 170^\circ$	Case Ia, Case IIIa	Case I, Case III
$\theta = 125^\circ$	Case II, Case IV	Case IIa, Case IVa

Dose distributions resulting from this process, for all the cases (viz. Cases Ia, I, II, IIa, IIIa, III, IVa and IV) are depicted in Fig. 16. We plot the inverse CDF of doses—i.e., the abscissa is the fraction of the infected population which receives a dose less than or equal to the ordinate. In each inset, we also plot a histogram of the dose distribution. Note that while the doses may easily span two orders of magnitude, about 80% of the infected people lie within a one-decade range.

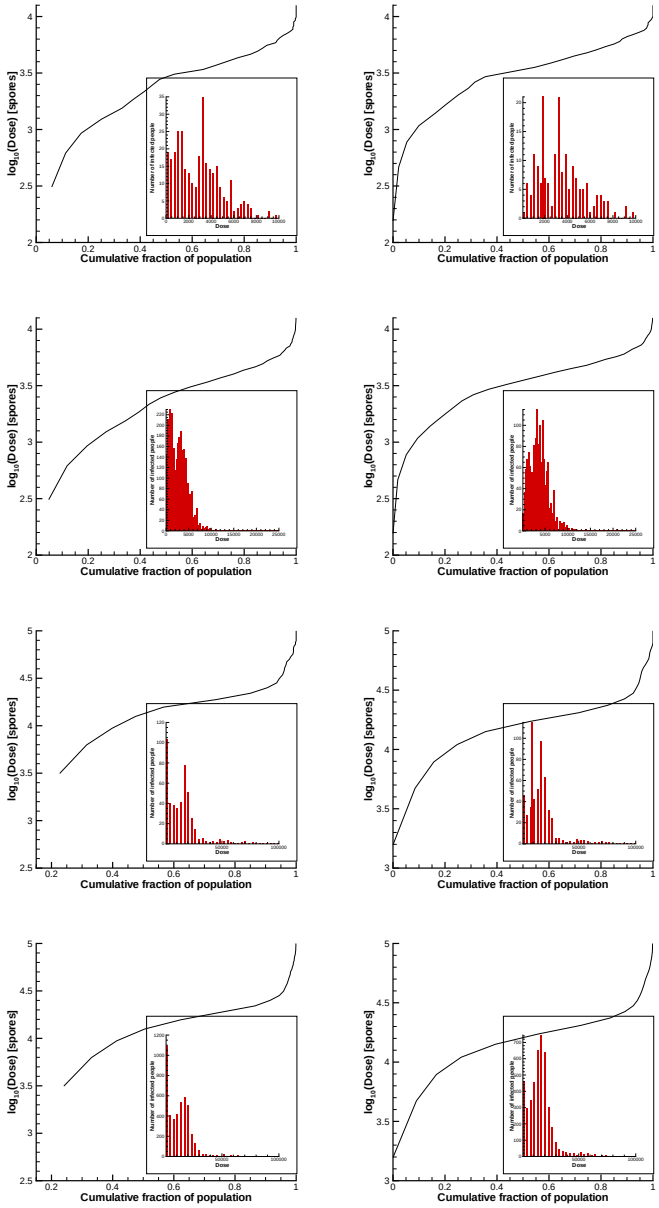

Figure 16. The inverse cumulative distribution of doses for Cases Ia, I, II, and IIa (left column) and Cases IIIa, III, IV, IVa (right column). The abscissa is the fraction of the infected population which receives a dose less than or equal to the ordinate. Inset: we plot histograms containing the number of infected people in each dose bin. While the histograms have long tails, the bulk of the population receives doses spanning one order of magnitude.

.