Published in final edited form as:

Stat Methods Med Res. 2019 December ; 28(12): 3591–3608. doi:10.1177/0962280218805780.

Accurate quantification of uncertainty in epidemic parameter estimates and predictions using stochastic compartmental models

Christoph Zimmer^{1,2}, Sequoia I Leuba¹, Ted Cohen¹, Reza Yaesoubi³

¹Epidemiology of Microbial Diseases, Yale School of Public Health, New Haven, CT, USA

²Bosch Center for Artificial Intelligence, Robert Bosch GmbH, Renningen, Germany

³Health Policy and Management, Yale School of Public Health, New Haven, CT, USA

Abstract

Stochastic transmission dynamic models are needed to quantify the uncertainty in estimates and predictions during outbreaks of infectious diseases. We previously developed a calibration method for stochastic epidemic compartmental models, called Multiple Shooting for Stochastic Systems (MSS), and demonstrated its competitive performance against a number of existing state-of-the-art calibration methods. The existing MSS method, however, lacks a mechanism against filter degeneracy, a phenomenon that results in parameter posterior distributions that are weighted heavily around a single value. As such, when filter degeneracy occurs, the posterior distributions of parameter estimates will not yield reliable credible or prediction intervals for parameter estimates and predictions. In this work, we extend the MSS method by evaluating and incorporating two resampling techniques to detect and resolve filter degeneracy. Using simulation experiments, we demonstrate that an extended MSS method produces credible and prediction intervals with desired coverage in estimating key epidemic parameters (e.g. mean duration of infectiousness and R_0) and short- and long-term predictions (e.g. one and three-week forecasts, timing and number of cases at the epidemic peak, and final epidemic size). Applying the extended MSS approach to a humidity-based stochastic compartmental influenza model, we were able to accurately predict influenza-like illness activity reported by U.S. Centers for Disease Control and Prevention from 10 regions as well as city-level influenza activity using real-time, city-specific Google search query data from 119 U.S. cities between 2003 and 2014.

Keywords

Calibration; outbreak; mathematical model; real time; influenza

Corresponding author: Christoph Zimmer, Bosch Center for Artificial Intelligence (BCAI), 60 College Street, New Haven, CT 06511, USA. christoph.zimmer@yale.edu.

Declaration of conflicting interests

The authors declare that there is no conflict of interest. However Dr. Christoph Zimmer changed affiliation during the peer review process from Yale School of Public Health to Bosch Center for Artificial Intelligence (BCAI) but there is no conflict of interest.

Supplemental Material

Supplemental material is available for this article online.

1 Introduction

The accurate prediction of epidemic behavior and the estimation of key epidemic parameters (e.g. expected number of secondary cases) are essential to inform control of emerging infectious threats. Stochastic transmission dynamic models are especially useful to characterize the uncertainty in estimates and predictions early in outbreaks given the importance of chance events when the number of infectious individuals is small. While a wide variety of methods have been developed to calibrate deterministic models,^{1–3} the calibration of stochastic models has received less attention.

We previously developed a calibration method for stochastic epidemic compartmental models called Multiple Shooting for Stochastic Systems (MSS),⁴ and demonstrated its competitive performance against a number of several existing state-of-the-art calibration methods including a likelihood approximation with an assumption of independent Poisson observations,⁵ a particle filtering method,⁶ and an ensemble Kalman filter method.⁶ The existing MSS method, however, lacks a mechanism to guard against filter degeneracy, a phenomenon that results in parameter posterior distributions that are weighted heavily around a single value. As such, when filter degeneracy occurs, these posterior distributions will not yield reliable credible and prediction intervals for parameter estimates and predictions.

In this work, we extend the MSS method by evaluating and incorporating two resampling technique to detect and resolve filter degeneracy. In simulation experiments, we use these two algorithms for resolving filter degeneracy and show that they are both effective and the extended MSS outperforms the existing MSS method in terms of coverage of the prediction intervals. To evaluate the performance of the extended MSS method in describing real-world influenza activity, we apply this method to the influenza-like-illness (ILI) surveillance data provided by U.S. Centers for Disease Control and Prevention (CDC) between 2003 and 2016 as well as to city-level Google Flu Trends (GFT) influenza data between 2003 and 2014 to retrospectively predict seasonal flu activity in 119 U.S. cities during this period.

2 Method

2.1 Calibration

We assume that epidemiological observations $y_1, \dots, y_n$ can be obtained at time points $t_1, \dots, t_n$. For example, the observations $y_1, \dots, y_n$ could be the weekly number of newly diagnosed cases. The observation times $t_1, \dots, t_n$ do not need to be equidistant. Our goal is to use the observations $y_1, \dots, y_n$ to estimate the parameters that govern disease transmission during an ongoing epidemic, such as the duration that individuals remain infectious or the number of secondary cases attributable to a single infectious case (i.e. reproduction number). To this end, we let the probability distribution $\pi_i(\theta)$ denote our belief about the true value of parameter vector θ having obtained observations $y_1, \dots, y_i$. We will use each new observation y_i to update this *prior* distribution according to

$$\pi_i(\theta | y_i, y_{i-1}, \dots, y_1) = \pi_{i-1}(\theta | y_{i-1}, \dots, y_1) \mathcal{P}(y_i | y_1, \dots, y_{i-1}; \theta) \quad (1)$$

$[Posterior\ at\ time\ t_i] = [Prior\ at\ time\ t_i] \times [probability\ to\ observe\ y_i]$ where $\mathcal{P}(y_i | y_1, \dots, y_{i-1}; \theta)$ is the probability to observe the current observation y_i given the epidemic history $y_1, \dots, y_{i-1}$. Note that the posterior at time t_{i-1} , π_{i-1} , is as well the prior distribution at time t_i . Since no observation is gathered prior to $i = 1$, we set $\pi_0(\theta | y_0) = \pi_0(\theta)$ which describes our prior knowledge about the epidemic parameters at the onset of the epidemic.

In the following subsections, we will describe how to characterize and approximate $\mathcal{P}$ using stochastic compartmental epidemic models. By knowing $\mathcal{P}$, we can then propagate the distribution π_i through the course of the epidemic using equation (1).

2.2 MSS method

As we show below, calculating the exact value of the function $\mathcal{P}(\bullet)$ is computationally prohibitive. To cope with this challenge, we use the MSS method to approximate $\mathcal{P}$ as it has proven to be computationally efficient and numerically accurate for this purpose. The MSS method has been developed in a systems biology context⁷⁻⁹ and has been extended to calibrate stochastic compartmental models of epidemics.⁴ It has demonstrated competitive performance against a number of existing calibration methods in extensive and thorough simulation experiments.⁴ We briefly describe the MSS method below and refer the reader to previous publications for additional details.⁴

Given accumulated observations $y_1, \dots, y_i$, we use $\Pi(\bullet | y_1, y_2, \dots, y_i)$ to denote the *belief state* at time t_i . A belief state is a probability distribution over all possible epidemic states which returns the probability that the epidemic is in a particular state when $y_1, \dots, y_i$ are observed. The epidemic state at time t_i is denoted by v_i and is a vector of integers representing the number of individuals in each epidemic compartment (e.g. number of susceptible, infected, or recovered individuals). We use v_0 to denote the initial epidemic state ($i = 1$) and $\Pi(v_0)$ to denote our belief about the epidemic state at time t_0 .

The probability function $\mathcal{P}(y_i | y_1, y_2, \dots, y_{i-1}; \theta)$ in equation (1) can now be calculated by conditioning on the epidemic state at time t_{i-1} , denoted by v_{i-1} , and the epidemic state v_i at time t_i

$$\mathcal{P}(y_i | y_1, y_2, \dots, y_{i-1}; \theta) = \sum_{v_i \in \Omega_i} \sum_{v_{i-1} \in \Omega_{i-1}} P(y_i | v_i, v_{i-1}) p(v_i | v_{i-1}; \theta) \times \Pi(v_{i-1} | y_1, y_2, \dots, y_{i-1}; \theta) \quad (2)$$

[Probability for current observation conditioned on history] = [Sum over all possible current states] [Sum over all possible previous states] [Probability for observation conditioned on specific current state] $\times$ [Transition probability to move from specific previous to specific current state] $\times$ [Belief state probability to be in specific previous state given history]. In equation (2), Ω_i is the support of the belief state at time t_i , P is the observation probability mapping the state v_i to the observation y_i that can also incorporate any additional

uncertainty in the data collection such as reporting errors, and p is the transition probability to move from state v_{i-1} at time t_{i-1} to state v_i at time t_i .

To use equation (2), we need to calculate probability functions $P(\bullet)$ and $p(\bullet)$, and the belief state $\Pi(\bullet)$ given observations $y_1, \dots, y_i$. The MSS method employs a linear noise approximation (LNA) to approximate the transition probability $p(\bullet)$ in equation (2).^{4,8,9} The LNA assumes that the probability distribution of $v_i | v_{i-1}$ can be approximated by a normal distribution $\mathcal{N}(x_i, \Sigma_i)$, where x_i is the solution of the ordinary differential equation (ODE) representation of the epidemic during interval $[t_{i-1}, t_i]$

$$\begin{aligned} \frac{d}{dt}x(t, v_{i-1}; \theta) &= \Gamma \Lambda(x(t, v_{i-1}; \theta), \theta), \quad (3) \\ x(0, v_{i-1}; \theta) &= v_{i-1} \end{aligned}$$

In the ODE system (3), the vector $x(t, v_{i-1}; \theta)$ is the epidemic state of the ODE model at time t given the initial state v_{i-1} , the vector $\Lambda(x(t, v_{i-1}; \theta), \theta)$ denotes the instantaneous changes in the epidemic when at state $x(t, v_{i-1}; \theta)$, and the matrix Γ describes the instantaneous changes of the epidemic state.

In using a normal distribution $\mathcal{N}(x_i, \Sigma_i)$ to approximate the probability distribution of $v_i | v_{i-1}$, the covariance matrix Σ_i will be the solution of the following ODE system on the interval $[t_{i-1}, t_i]$ ^{10,11}

$$\begin{aligned} \frac{d}{dt}\Sigma(t, v_{i-1}; \theta) &= J(x, \theta)\Sigma(t, v_{i-1}; \theta) + \Sigma(t, v_{i-1}; \theta)J(x, \theta)^t + D(x; \theta) \quad (4) \\ \Sigma(0, v_{i-1}; \theta) &= \hat{\Sigma}_{i-1} \end{aligned}$$

In the ODE system (4), $J(x, \theta) = \Gamma \frac{d}{dx} \Lambda(x, \theta)$ and D is a $K \times K$ matrix with the (i, j) entity equal to $\sum_{k=1}^K \Gamma_{jk} \Gamma_{jk} \Lambda(x, \theta)$, where K is the number of compartments in the epidemic model. The initialization of the covariance for each interval $\hat{\Sigma}_{i-1}$ will be discussed in the next subsections.

We next describe two approaches for calculating the belief state Π in equation (2). The first approach (which we refer to as MSSa) seeks to simplify the computational complexity of equation (2) by representing the belief state $\Pi(\cdot | y_1, \dots, y_i)$ as a step function that takes 1 for the most probable state (denoted by $\hat{v}_{i-1}$) and 0 elsewhere. This effectively lets us drop the second summation in equation (2). For each interval, a new state estimate $\hat{v}_i$ is determined as follows: given a previous state estimate $\hat{v}_{i-1}$ at time t_{i-1} , the probability of observing y_i at the end of the interval $[t_{i-1}, t_i]$ is equal to $P(y_i | v_i, \hat{v}_{i-1})p(v_i | \hat{v}_{i-1}; \theta)$. We now define the state at time t_i as our state estimate $\hat{v}_i$ that leads to the highest probability of observing y_i

$$\hat{v}_i = \underset{v_i \in \Omega_i}{\operatorname{argmax}} P(y_i | v_i, \hat{v}_{i-1}) p(v_i | \hat{v}_{i-1}; \theta) \quad (5)$$

[State estimate] = [Maximum over possible states] of [Probability for observation conditioned on specific current state] $\times$ [Transition probability to move from previous state estimate to specific current state]. At the beginning of each period, the ODE system to propagate the epidemic state to the next period is initialized with the state estimate $\hat{v}$. Since the MSSa approach assumes a point distribution for the belief state Π , the LNA method to characterize the distribution of the epidemic state in the next period is initialized with zero variance for $\hat{\Sigma}_{i-1}$ (see Zimmer et al.⁴ for details).

The second approach for calculating the belief state n , which we refer to as MSSb, assumes that the observation probability $P(y_i | v_i, \hat{v}_{i-1})$ is Gaussian and models the relation between the observation, y_i , and the state as $y_i = h(v_i, \hat{v}_{i-1}) + \epsilon_i$ where $h(\bullet)$ is the observation function mapping the state of the epidemic v_i to the observation y_i , and $\epsilon_i \sim \mathcal{N}(0_{\text{obs}}, \Sigma^{\text{obs}})$ is a normally distributed error term with zero mean vector 0_{obs} and an observation error covariance Σ^{obs} (see equation (14) for an example in the context of a SIRS model). In this case, the belief state n will follow a Gaussian distribution.^{12–14} That is, $v_i | y_0, \dots, y_i \sim \mathcal{N}(\hat{v}_i, \hat{\Sigma}_i)$ where the state estimate $\hat{v}_i$ and the covariance estimate $\hat{\Sigma}_i$ are calculated as the follows^{12–14}

$$\begin{aligned} \hat{v}_i &= x_i + K_i(y_i - \hat{y}_i) \\ \hat{\Sigma}_i &= \Sigma_i - K_i R_i K_i^T \end{aligned} \quad (6)$$

In the equations above, R_i is the observation variance, $\hat{y}_i$ is the predicted outcome, and K_i is the so-called Kalman gain:

$$\begin{aligned} \hat{y}_i &= h(x_i) \\ R_i &= C_i \Sigma_i C_i^T + \Sigma^{\text{obs}} \\ K_i &= \Sigma_i C_i^T R_i^{-1} \end{aligned}$$

where $\bullet^T$ stands for matrix transpose, $\bullet^{-1}$ for the matrix inverse and $C_i = \frac{\partial h}{\partial x} h(x)$. In this case, the ODE system (equation (3)) is initialized with the state estimate $\hat{v}_{i-1}$ and the LNA system (equation (4)) is initialized with the covariance estimate $\hat{\Sigma}_{i-1}$.

If the measurement error ϵ_i follows a Gaussian distribution, MSSb is expected to outperform MSSa because MSSb accounts for the Gaussian error when updating estimates of the epidemic state and its variance.

2.3 Resolving filter degeneracy

To initialize the recursion (1), we sample (with replacement) M parameter vectors $\theta_1, \theta_2, \dots, \theta_M$ from a prior parameter distribution $\pi_0(\theta)$, and M initial state vectors $v_{0,1}, v_{0,2}, \dots, v_{0,M}$ from a prior state distribution $\Pi(v_0)$. We consider the initial states as an additional dimension of the parameter vector and denote $\vartheta_m = (\theta_m, v_{0,m})$, $m = 1, \dots, M$. With any new observation y_i , we calculate the probabilities $\pi_i(\vartheta_m)$ for each vectors ϑ_m , $m = 1, \dots, M$, according to the recursion (1). This allows us to assign a weight, $w_m = \pi_i(\vartheta_m)$, to each sampled parameter vector ϑ_m , $m = 1, \dots, M$.

Since we are using a finite set of samples $\vartheta_1, \dots, \vartheta_M$ from the prior distributions π_0 and $\Pi(v_0)$, it is possible that after a few observations, the recursion (1) results in probability weights that are significant for only a few parameter vectors and are several orders of magnitude lower for the rest. In an extreme scenario, only one parameter vector is assigned a non-trivial weight and all others are assigned weights that are almost zero. This phenomenon is commonly referred to as *filter degeneracy*.^{15–17}

Filter degeneracy can cause serious problems when calculating credible or prediction intervals for model parameters or when attempting to use the model to make predictions since such degeneracy may result in a posterior distribution that is essentially a point distribution. We note that if we have unlimited computational resources, filter degeneracy can be avoided since we could theoretically test the entire space of the parameters to accurately calculate the posterior distributions. However, as we are always bound by computation and time limitations, we need to use a sample of parameter vectors to estimate the posteriors. Next, we describe two approaches to detect and resolve filter degeneracy when the recursion (1) is used to calculate the likelihood of accumulated observations during an epidemic.

2.3.1 Detecting and resolving filter degeneracy using effective sample size—

Detection of filter degeneracy is relatively simple and can be achieved by measuring the relative contribution of each sampled parameter vector. To this end, we first normalize weights using¹⁷

$$\tilde{w}_m = \frac{w_m}{\sum_{j=1}^M w_j} \quad (7)$$

and then calculate the effective sample size N_{eff} as a measure of (non-)degeneracy

$$\hat{N}_{\text{eff}} = \frac{1}{\sum_{m=1}^M \tilde{w}_m^2} \quad (8)$$

Before any observation is obtained (i.e. $i = 0$), we assign equal weights to sampled parameter vectors, and hence, N_{eff} is equal to M . In the other extreme, if all the weight is on one particle, N_{eff} is equal to 1. Therefore, one way of detecting degeneracy is to determine a threshold value N_C with $1 < N_C < M$ below which we trigger the algorithm described next to resolve filter degeneracy. The calculation of N_{eff} is not time-consuming, and thus, one can easily check if filter degeneracy has occurred in each iteration. Choosing a higher N_C implies a stricter mechanism against degeneracy but also imposes a larger computational cost. We found that $N_C = M/2$ works well in our study.

To resolve filter degeneracy once it is detected, we re-sample with replacement $\tilde{M}$ parameter vectors $\vartheta'_1, \dots, \vartheta'_{\tilde{M}}$ from $\vartheta_1, \dots, \vartheta_M$ according to current weights $w_1, \dots, w_M$. Using these samples and the kernel density estimation defined in the equation below, we construct a density estimate of the parameter distribution

$$\hat{f}(\vartheta) = \frac{1}{\tilde{M}} \det(H)^{-\frac{1}{2}} \sum_{l=1}^{\tilde{M}} K\left(\det(H)^{-\frac{1}{2}}(\vartheta - \vartheta'_l)\right) \quad (9)$$

[Estimated density] = [Normalization by number of points and bandwidth] [Sum over all points] [Kernel probability for each point]. In equation (8), $\det(H)$ is the determinant of matrix H , the (multivariate) kernel $K(\bullet)$ is a (multivariate) probability density (e.g. Gaussian), and H is the bandwidth matrix which controls the smoothness of the density estimate $\hat{f}(\cdot)$.

In a one-dimensional case, the density estimate (9) reduces to $\hat{f}(\vartheta) = \frac{1}{\tilde{M}\sqrt{H}} \sum_{l=1}^{\tilde{M}} K\left(\frac{\vartheta - \vartheta'_l}{\sqrt{H}}\right)$,

where K can be a Gaussian kernel $K(u) = 1/\sqrt{2\pi} e^{-\frac{u^2}{2}}$. Small values for H (the so-called bandwidth or smoothing parameter) place higher probability around each sample and lower probability elsewhere, whereas high values for H lead to a smoother density function. See Supporting Information “Technical details” for details and our choice of $K(\bullet)$ and H .

The re-sample size $\tilde{M}$ can be larger or smaller than M , but since the procedure to correct filter degeneracy is computationally inexpensive, one would rather choose $\tilde{M} \geq M$ to improve the quality of density estimation. For the results presented here, we chose $M = \tilde{M} = 1000$ as in Zimmer et al.⁴

Next, M new parameter vectors $\vartheta'_1, \dots, \vartheta'_M$ are sampled based on density function $\hat{f}$. As no state estimates and weights have yet been calculated for these new parameter vectors, we calculate the new weight with the following

$$\bar{w}_m = \prod_{k=1}^i \mathcal{P}(y_k | y_1, \dots, y_{k-1}, \vartheta'_m) \quad (10)$$

$[New\ weight] = [Product\ over\ all\ previous\ time\ intervals] [Probability\ for\ observation\ given\ its\ history]$ with i being the current iteration, θ'_m the parameter vector component of θ'_m , and the initial state estimate $v'_{0,m}$ determined by the state vector component of θ'_m . Whenever distinction is needed, we refer to this approach by adding “-Neff”, e.g. MSSb-Neff for the MSSb version with effective sample sized based approach against filter degeneracy.

2.3.2 Detecting and resolving filter degeneracy using incremental mixture importance sampling—The Incremental Mixture Importance Sampling (IMIS) algorithm is an alternative to detect and resolve filter degeneracy. It is an iterative algorithm that can be applied in each iteration of the filter. This section is a brief review of the algorithm described in Raftery and Bao.¹⁶ The IMIS algorithm checks each of its iterations for degeneracy and suggests new samples until filter degeneracy is resolved. The IMIS algorithm is initialized ($k = 0$) by normalizing the weights of the parameter sample (equation 6). The k th iteration of IMIS ($k \in \{1, 2, \dots\}$) checks whether the samples are degenerated by calculating the expected fraction of unique points after re-sampling:

$$Q(\tilde{w}) = \frac{1}{B_k} \sum_{j=1}^{B_k} \left(1 - (1 - \tilde{w}_j)^M\right)$$

where $B_k = M + k * B$ and B is the number of re-samples in each iteration. If the expected number of unique points is too small

$$Q(\tilde{w}) < Q \quad (11)$$

then the IMIS algorithm is carried out for an additional iteration.

In each iteration k , the IMIS algorithm samples new inputs (with hopefully higher weights) until the expected number of points passes the desired threshold (equation 10). The details can be found in the Supporting Information “IMIS details” or in the original publication.¹⁶

Once the termination criterion is satisfied ($Q(\tilde{w}) \geq Q$), M inputs are re-sampled with replacement from $\vartheta_1, \dots, \vartheta_{B_k}$ and their corresponding state estimates $\hat{v}_{i,1}, \dots, \hat{v}_{i,B_k}$ with the corresponding weights $w_1^{(k)}, \dots, w_{B_k}^{(k)}$. We refer to this approach with MSSb-IMIS.

Figure 1 shows the pseudo code of our MSS extension being able to resolve filter degeneracy, and Figure 2 lists the free parameters of the methods studied here.

2.4 Quantifying uncertainty in parameter estimates and epidemic predictions

We use credible intervals to characterize the uncertainty in parameter estimates. The $100(1 - \kappa)\%$ credible interval for the parameter vector θ is defined as the interval between $\kappa/2$ and $(1 - \kappa)/2$ -quantiles of the probability distribution $\pi_i(\theta | y_1, \dots, y_i)$. Accordingly, the $100(1 - \kappa)\%$ prediction interval for the desired prediction target Z (e.g. the number of cases diagnosed in

the next week or the time of the peak of the epidemic) is calculated as the interval between $\kappa/2$ and $(1 - \kappa/2)$ -quantiles of the probability distribution $P(Z|y_1, \dots, y_i)$ with

$$P(Z|y_1, \dots, y_i) = \int_{\theta \in \Theta} \int_{v_i \in \Omega_i} P(Z|v_i; \theta) \Pi(v_i|y_1, \dots, y_i; \theta) \pi_i(\theta|y_1, \dots, y_i) dv_i d\theta \quad (12)$$

[Probability to observe target conditioned on history] = [Integration over parameter space] [Integration over state space] [Probability to observe target conditioned on specific state and specific parameter] [Probability for specific state conditioned on history and parameter] [Probability for specific parameter conditioned on history]. Here Θ is the parameter space and $P(Z|v_i; \theta)$ is the probability to observe Z given the epidemic state v_i and the parameter vector θ . Due to the Markovian property of compartmental models, predictions Z are independent given v_i, θ —the model can be fully described by knowing current state v_i and parameter values θ .

This probability can be calculated with the help of a simulation model. We evaluate equation (11) by obtaining M' new realizations of parameter vectors according to their current weights and then carrying out M' simulations initialized with the state estimate $\hat{v}'_{i,1}, \dots, \hat{v}'_{i,M'}$ until the time of the target Z . For this study, we set $M' = 100$.

2.5 Humidity-based SIRS model

To evaluate the effectiveness of the approach proposed above to resolve filter degeneracy, we designed a simulation study using a humidity-based SIRS (Susceptible $\rightarrow$ Infected $\rightarrow$ Recovered $\rightarrow$ Susceptible) model. We later use this model to predict influenza activities in cities in the United States using GFT data.

The humidity-based SIRS model divides the population into three disjoint subgroups.^{6,19} Susceptible members may be infected at a rate proportional to the prevalence of infectious individuals in the population and will immediately become infectious to others. Infectious individuals recover at a constant rate and develop temporary immunity against the circulating pathogen. Once the immunity wanes, the individual becomes susceptible to infection. These transitions are represented as

where γ is the mean infectious period, α is the average duration of immunity, and $\beta(t)$ is the transmission rate defined as $\beta(t) = R_0(t)/\gamma$. Here

$R_0(t) = \exp\left(-180q(t) + \log\left(R_{0_{\max}} - R_{0_{\min}}\right)\right) + R_{0_{\min}}$ with a maximal and minimal daily reproductive number $R_{0_{\max}}$ and $R_{0_{\min}}$, and the function $q(t)$ describes the absolute humidity

over time⁶ (see Supporting Information “Interpolation of humidity data” for details on how to interpolate the city-level humidity data for the United States).

The model can be also written in ODE form as

$$\begin{aligned}\frac{dS}{dt} &= \frac{R}{\alpha} - \frac{\beta(t)IS}{N} \\ \frac{dI}{dt} &= \frac{\beta(t)IS}{N} - \frac{I}{\gamma} \\ \frac{dR}{dt} &= \frac{I}{\gamma} - \frac{R}{\alpha}\end{aligned}\quad (14)$$

We assume that we can observe the number of new infections during each week, which corresponds to the number of population members that move from the compartment S to I in each week. We also assume that our population size is constant, and hence, we can replace R with $R = N_{pop} - S - I$ in ODE system (14) and eliminate the third equation.

To use the weekly number of cases to calibrate this SIRS model, we need to specify the observation function $P(\bullet)$ (equation (2)). We note that this time-series does not directly correspond to any of the model states and instead represents the *flow* from S to I . If there were no recoveries in the time interval $[t_{i-1}, t_i]$, we could calculate this flow by taking the difference between the number of infected people at time t_i and t_{i-1} : $v_i^{(I)} - \hat{v}_{i-1}^{(I)}$, where $v_j^{(I)}$ denotes the number of people in compartment I at time point t_j . But if there are recoveries, we need to add them to this difference. We therefore define the observation function $h(\bullet)$ as

$$h(v_i, v_{i-1}) = v_i^{(I)} - \hat{v}_{i-1}^{(I)} + R_i^* \quad (15)$$

where R_i^* denotes the number of recoveries within the time interval $[t_{i-1}, t_i]$ (see Supporting Information “Technical details”).

2.6 Simulation study

To evaluate the ability of the MSS approach to produce reliable credible and prediction intervals for epidemic parameters and predictions, we used simulated trajectories generated by the humidity-based SIRS model described above (Figure 3). In these simulation experiments, the true parameter values and the corresponding epidemic trajectories are known, and hence, we can evaluate whether the intervals characterized by our proposed approach cover the true values of parameters and prediction targets.

To generate the simulated trajectories shown in Figure 3, we used the Gillespie algorithm²⁰ (implemented in COPASI²¹) with epidemic parameters and initial states drawn from the following uniform distribution: $(\alpha, \gamma, R_{0_{max}}, R_{0_{min}}) \sim U([365, 3650] \times [1.5, 7] \times [1.3, 4] \times [0.8, 1.2])$, $S_0, I_0 \sim U([50,000, 70,000] \times [10, 100])$ and $N_{pop} = 100,000$. We use the New York humidity profile for this simulation experiment.

To ensure that we only use trajectories that display reasonable behavior, we included simulated trajectories with an attack rate of at least 10% (to avoid premature die out) and a peak week no earlier than epidemic week 49 (consistent with real-world influenza epidemics). We let the weekly number of cases be observed with some measurement error that follows a normal distribution with mean zero and standard deviation 100. We will assume that this distribution is also known to the calibration algorithm which allows us to employ the MSSb method. As expected and confirmed by our numerical experiments, in the presence of such measurement error, MSSb outperforms MSSa, and hence, we only report the results from the MSSb method. The re-sampling kernel density estimation (equation (8)) is calculated using the Mathematica²² function `KernelMixtureDistribution` with a Gaussian kernel, a maximum of 5 mixture kernels, and bandwidths of 100, 0.1, 0.1, 0.025, 1000, 10 that correspond to parameters α , γ , $R_{0_{max}}$ and $R_{0_{min}}$ as well as initial states $v_0^{(S)}$ and $v_0^{(I)}$. These bandwidths are determined based on the range that each model parameter may vary over.

3 Results

We first present the results of the simulation experiment described above. We then employ the MSS method to predict regional-level ILI activities in the United States using CDC ILI surveillance data, and city-level influenza activities using GFT data.

3.1 Simulation study

Figure 4 displays how the effective sample size N_{eff} and the fraction of unique points ($Q(\tilde{w})$ in equation (10)) change as the model is re-calibrated using the latest observations accumulating during epidemics. N_{eff} approaching 1 or $Q(\tilde{w})$ approaching 0 implies that the probability weight of one of the parameter vectors is approaching one and all the others zero, signaling the occurrence of filter degeneracy (blue curve). Figure 4 demonstrates that filter degeneracy clearly occurs in the old MSS version and both methods described here to rectify filter degeneracy are effective (red and green curve).

Figure 5 compares the predictions made by the MSSb-Neff method (red line) with the true value of several epidemic targets (black crosses) for two simulated trajectories. These key epidemic targets include infection prevalence, number of cases at the epidemic peak (number of cases in the worst week), the number of new cases over the next week, and the final epidemic size. These two simulated trajectories present different realistic epidemic behaviors; Trajectory 1 spreads faster, peaks sooner, and results in more infections compared to Trajectory 2. The MSSb method is able to capture the behaviors of all 50 simulated trajectories (Figure 3) and to quickly distinguish these trajectories from each other (Figure 6 and Supporting Information Figures A5 to A8).

Figure 6(a) displays the proportion of 90% prediction intervals characterized by all three MSS versions for the simulated trajectories shown in Figure 3 that include the true value of the prediction target. While these simulated trajectories present markedly different epidemic behaviors, the prediction intervals produced by the new MSS versions have the expected level of coverage (Figure 6(a)). This figure also demonstrates that the old version of MSS without proper mechanism against filter degeneracy clearly fails to produce the desired

coverage for infection prevalence, or to a smaller extent, final epidemic size. Figure 6(b) shows that the point predictions produced by all MSS versions have relatively low relative error with respect to the true values of the prediction targets. This is expected as degenerated posteriors can still result in accurate point predictions, as seen in the old MSS method that lacked a mechanism to resolve filter degeneracy.

The MSS method can also be used to estimate key epidemic parameters such as the humidity driven quantity $R_0(t)$, mean duration of immunity (α), mean duration of infectiousness (γ), and number of susceptibles at the beginning of the flu season. Figure 7 compares the estimates provided by our method with the true value of these parameters for two simulated trajectories. The MSS method reached the desired coverage for the parameter estimate (Figure 8(a)) with small relative error (Figure 8(b)). We note that the width of the credible intervals is influenced by the sensitivity of the epidemic dynamics to the parameter. Tighter intervals suggest greater sensitivity, as small changes in these parameters are expected to lead to substantial changes in the system's dynamics. On the other hand, less sensitive parameters, such as mean duration of immunity, can take a larger range of values without strongly influencing the dynamics, and hence, can be less precisely determined (i.e. larger width for their credible intervals). We also clearly see that the old version of MSS, which was not equipped with mechanisms against filter degeneracy, fails to produce the desired level of coverage for credible or prediction intervals (Figures 6(a) and 8(a)) further targets can be found in Figure A9.

3.1.1 Sensitivity analyses—For the analyses presented above, we assumed that the weekly number of cases are observed with some measurement error that follows a normal distribution with mean zero and a standard deviation that remains constant over time. Supporting Information Figure A4 shows that the performance of the methods studied here to address filter degeneracy is not affected if the standard deviation of measurement error is assumed to be proportional to the number of cases observed weekly. We also note that another possible approach to address filter degeneracy is to simply increase the number of parameter samples M . While increasing M from 1000 to 10,000 improved the coverage of some calibration targets (Supporting Information Figure A1), it does not properly address the issue of filter degeneracy (Supporting Information Figure A3), and also led to a significant increase in the computation time (Supporting Information Figure A2). Finally, in our simulation experiments, while both methods we studied resolved filter degeneracy, the IMIS-based method showed a slight performance gain. However, its computational time was almost 10-fold of that of the approach based on effective sample size. Therefore, we only used the latter method to analyze the CDC ILI and GFT influenza data (see below).

3.2 Prediction and parameter estimation using CDC ILI data

Next, we use the CDC ILI data²³ to evaluate the ability of our method to produce reliable forecasts. This data set comprises 143 ILI time-series from 10 U.S. Department of Health & Human Services (HHS) regions between 2003 and 2016 including data from the first and second waves of the 2009 pandemic. Consistent with previous studies,²⁴ we modeled ILIs as a single disease. This is a simplifying assumption but has shown to be sufficient to accurately describe ILI activities.

For the results presented below, the re-sampling kernel density estimation (equation (8)) is calculated using the Mathematica function `KernelMixtureDistribution` with a Gaussian kernel, a maximum of 5 mixture kernels, and bandwidths of 100, 0.1, 0.1, 0.025 for the parameters α , γ , $R_{0_{max}}$ and $R_{0_{min}}$, and 1000 for $v_0^{(S)}$ (i.e. initial size of the susceptible population) and the median of the current samples for $v_{0,m}^{(I)}$.

3.2.1 Humidity data—Following the approach described in Text Si of Yang et al.,⁶ we independently constructed a humidity dataset from the original Phase 2 of the North American Land Data Assimilation System (NLDAS-2) dataset.²⁵ These data were derived through the National Centers for Environmental Prediction North American Regional Reanalysis. The hourly data are available on a 0.125° grid from 1979 to present. We extracted the specific humidity data for 119 cities and then averaged the hourly data to develop a daily climatology for each city from 1979 to 2016. For regional levels and national level, humidity is averaged over the cities being part of the respective area.

3.2.2 Accuracy and reliability of predictions using CDC ILI data—Figure 9 displays one-week forecasts for the 10 HHS health regions. As shown in this figure, our method can produce reliable predictions for outbreaks with different characteristics: large or small epidemic size (Region 10 vs. Region 1), pronounced or less pronounced peaks (Region 4 vs. Region 2). The 2009 influenza pandemic presents distinct behavior compared with seasonal epidemics; in 2009, most regions experienced a minor peak followed by a major one (Figure 10). As evident in Figure 10, the method proposed here can describe such behavior and provide reliable one-week predictions for the 2009 influenza pandemic.

While Figures 9 and 10 display predictions for specific seasons, similar accuracy is obtained for the other seasons (Supporting Information Figure A12). Our analysis suggests that the prediction intervals displayed in Figures 9 and 10 and those in other seasons exhibit the desired level of coverage (Supporting Information Figure A10).

3.3 Prediction and parameter estimation using GFT data

Next, we used GFT data²⁷ to evaluate the ability of our method to produce reliable city-level influenza forecasts and epidemic parameter estimates. Internet search queries have gained attention as a surveillance source for monitoring influenza activity since 2009.^{28,29} GFT data have demonstrated competitive performance in matching influenza-like illness (ILI) data in various countries and regions such as the United States,³⁰ South Korea,³¹ Southern China,³² Canada,³³ Latin America,³⁴ and New Zealand.³⁵

3.3.1 Weekly influenza-positive cases—GFT previously published municipal-level estimates of weekly ILI incidence per 100,000 people based on internet search query activity.²⁹ The U.S. CDC publishes influenza metrics through the U.S. World Health Organization Collaborating Laboratories System and the National Respiratory and Enteric Virus Surveillance System (WHO/NREVSS). WHO/NREVSS records data on the total number of respiratory specimens tested and the number positive for influenza viruses in clinical laboratories on a national, regional, and census division level.³⁶

By multiplying the GFT ILI incidence data by the CDC data on the proportion of ILI cases who were test-positive for the same epidemiological week, we constructed a dataset of ILI+ for 121 cities from 2003 to 2014. This approach is consistent with previous studies to construct ILI+ time-series that estimate the municipal influenza infections per 100,000 patient visits.¹⁹ We selected cities within the continental U.S. with complete observations for all epidemic weeks between 2003 and 2014 leading to a total of 1168 time-series for analysis.

3.3.2 Accuracy and reliability of predictions using GFT data—To fit the humidity-based SIRS model to the ILI+ time-series, we assumed the same parameter prior distributions as used in Yang et al.⁶ with one exception: if the number of accumulated cases passed 50,000 (signaling a more severe epidemic), we changed the prior distribution of α to $U[50,3650]$ (instead of using $U[365,3650]$). We found that this change was important for fitting models to these relatively severe epidemics. This highlights the important effect that the selection of prior distributions might have on the overall fit of the model. Prior distributions should ideally be determined based on domain expert knowledge. In the absence of such information, the effect of each parameter on the model behavior should be investigated visually to characterize appropriate prior distributions.

Figure 11 displays one-week forecasts for a select number of U.S. cities. As shown in this figure, our method can produce reliable predictions for outbreaks with different characteristics: large or small epidemic size (New Orleans vs. Miami), early or late peak timing (Los Angeles vs. Miami), and single or double peaks (San Francisco vs. New York). There seems to be a slight overshooting of the peaks in some of the cities (such as Houston and Chicago) which might be due to effects not covered in a relatively simple SIRS model such as behavioral changes over the course of the epidemics, school vacations, or vaccinations. The one-week predictions for other U.S. cities are provided in the Supporting Information Figure A13.

The 2009 influenza pandemic presents distinct behavior compared with seasonal epidemics; in 2009, most cities experienced a minor peak followed by a major one (Figure 12). A third peak was also observed in several cities (e.g. New Orleans). As evident in Figure 12, the method proposed here can describe such behavior and provide reliable one-week predictions for the 2009 influenza pandemic.

While Figures 11 and 12 display predictions for major cities, similar accuracy is obtained for smaller cities (Supporting Information Figure A13). Our analysis suggests that the prediction intervals displayed in Figures 11 and 12 and those in smaller cities exhibit the desired level of coverage (Supporting Information Figure A11).

Figure 13 displays the distribution of R_0 estimates calculated for the cities included in our study for season 2003/04 to 2013/14. Adopting the same approach used by Yang et al.,³⁷ we calculate R_0 by evaluating $R_0(t)$ of the humidity-based SIRS model in the week of the maximal effective R (effective R at time t is defined as $R_e(t) = R_0 \frac{S(t)}{N(t)}$ (where $\frac{S(t)}{N(t)}$ is the proportion of the population that is susceptible at time t). The R_0 estimates shown in Figure 13 are consistent with the estimates reported in past studies.^{37–40}

4 Discussion

The accurate prediction of epidemic behaviors and the estimation of epidemic parameters are challenged by the partial observability of and the intrinsic randomness in the process governing the disease spread. Characterizing the uncertainty in these prediction and parameter estimates is essential for making meaningful inference about the true parameter values and for drawing conclusions on the epidemic behavior under different control strategies.

In this work, we extend an existing calibration method called MSS and equipped it with a mechanism to produce reliable credible and prediction intervals for the estimates of epidemic parameters and predictions for the epidemic behavior. The MSS method is a calibration approach for stochastic compartmental epidemic models that has competitive performance to several existing state-of-the-art calibration methods including a particle filter⁶ and an ensemble Kalman filter.⁶ The original MSS method was not capable of handling filter degeneracy (Figure 4), a phenomenon where particles or members of an ensemble or filter have a very small likelihood that varies over orders of magnitude even for the best members or particles. Filter degeneracy can result in degenerate posterior distributions and hinder the ability to properly characterize predictions and credible intervals for predictions and parameter estimates.

Using simulation experiments on stylized models of influenza outbreaks (where parameter values were known with certainty), we demonstrated that the extended MSS method can effectively detect and resolve the occurrence of filter degeneracy (Figure 4), and—in contrast to the previous version—produce prediction and credible intervals with desired coverage level (Figures 6(a) and Figure 8(a)) for key epidemic parameters such as duration of infectiousness and the reproductive number. We note that while we use influenza as an illustrative example here, the proposed method can be readily applied to other stochastic compartmental models.

Applying the extended MSS approach with method to address filter degeneracy to a humidity-based stochastic compartmental influenza model, we found that we could accurately predict CDC ILI data for the 10 US regions and national level as well as city-level influenza activity using real-time, city-specific Google search query data from 119 U.S. cities between 2003 and 2014 (Figure 11). The proposed method was robust to distinct patterns observed over this period, such as variable peak times and variable magnitudes of these seasonal epidemics. The extended MSS method addressing filter degeneracy performs well even for exceptional seasons, such as the 2009 influenza pandemic and those with multiple peaks (Figure 12). Analyzing the coverage of prediction intervals shows that the expected fraction of prediction intervals covers the observed number of cases (Supporting Information Figure A11).

Figures 11 and 12 indicate that the humidity-based SIRS model tends to overestimate the peak of the epidemic in several cities. Such overestimation was not observed when predicting the peak of simulated epidemics (Figure 5 and Supporting Information Figures A5 and A8). This suggests that our simplified humidity-based SIRS model does not

accurately capture all complex characteristics of influenza epidemics, such as changes in contact rates or health-seeking behavior of population members through the course of epidemics.

All versions of the MSS method rely on the LNA to capture the correlations between epidemic compartments throughout the outbreak. The LNA has been successfully applied to describe intrinsic stochastic fluctuations in various systems^{4,8,9} and has a strong theoretical basis⁴¹ which suggests that LNA accurately approximates the dynamic behavior of large populations, especially when fluctuations are small compared with the steady state of the system. As demonstrated by our results, we find that LNA also performs well when the intervals over which the LNA projections are made are sufficiently small (here, the epidemic observations were accrued weekly). We also note that LNA accuracy can be improved by using higher order terms for correction, if this is needed.⁴²

In this paper, we focused on two methods for resolving filter degeneracy, both of which proved effective in calibrating our humidity-based SIRS model against simulated data. We then used one method for resolving filter degeneracy to accurately determine uncertainty for GFT ILI data. We, however, note that there exist other methods to address filter degeneracy that we did not study here.^{6,17}

While we used the ILI+ metric to evaluate the performance of the MSS method, we recognize that this metric serves only as a proxy for influenza incidence.^{43,44} As ILI+ is proportional to GFT ILI rates, any observational error may be exacerbated by abnormal web-search activity (e.g. increased media coverage often leads to higher levels of web-search activity, and in consequence, higher GFT ILI rates). This further emphasizes the importance of conventional surveillance systems for accurate monitoring of influenza activity.⁴³

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The authors wish to thank an anonymous reviewer of a previous manuscript for pointing out the formulas for state updating in case of Gaussian noise. The authors also thank A Kunkel for her advice on the R code to fetch the humidity data.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship and/or publication of this article: CZ was funded by the German research foundation (DFG) and National Institutes of Health (NIH) U54GM088558. SIL and TC were supported by NIH U54GM088558. RY was supported by grant 1K01AI119603 from the National Institute of Allergy and Infectious Diseases. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Allergy and Infectious Diseases or the National Institute of General Medical Sciences.

References

1. Alkema L, Raftery AE and Clark SJ. Probabilistic projections of HIV prevalence using Bayesian melding. *Ann Appl Stat* 2007; 1: 229–248.

2. Elderd BD, Dukic VM and Dwyer G. Uncertainty in predictions of disease spread and public health responses to bioterrorism and emerging diseases. *Proc Natl Acad Sci* 2006; 103: 15693–15697. [PubMed: 17030819]
3. Birrell PJ, Ketsetzis G, Gay NJ, et al. Bayesian modeling to unmask and predict influenza A/ H1N1pdm dynamics in London. *Proc Natl Acad Sci* 2011; 108: 18238–18243. [PubMed: 22042838]
4. Zimmer C, Yaesoubi R and Cohen T. A likelihood approach for real-time calibration of stochastic compartmental epidemic models. *PLOS Comput Biol* 2017, 10.1371/journal.pcbi.1005257.
5. Riley S, Fraser C, Donnelly CA, et al. Transmission dynamics of the etiological agent of SARS in Hong Kong: impact of public health interventions. *Science* 2003; 300: 1961–1966. [PubMed: 12766206]
6. Yang W, Karspeck A and Shaman J. Comparison of filtering methods for the modeling and retrospective forecasting of influenza epidemics. *PLOS Comput Biol* 2014; 10: e1003583. [PubMed: 24762780]
7. Zimmer C and Sahle S. Parameter estimation for stochastic models of biochemical reactions. *J Comput Sci Syst Biol* 2012; 6: 11–21.
8. Zimmer C and Sahle S. Deterministic inference for stochastic systems using multiple shooting and a linear noise approximation for the transition probabilities. *IET Syst Biol* 2015; 9: 181–192. [PubMed: 26405142]
9. Zimmer C Reconstructing the hidden states in time course data of stochastic models. *Math BioSci* 2015; 269: 117–129. [PubMed: 26363082]
10. Thomas P, Matuschek H and Grima R. Intrinsic noise analyzer: a software package for the exploration of stochastic biochemical kinetics using the system size expansion. *PLOS one* 2012; 7(6): e38518. [PubMed: 22723865]
11. Van Kampen NG. *Stochastic processes in physics and chemistry*. Vol. 1, Amsterdam, TheNetherlands: Elsevier, 1992.
12. Finkenstadt B, Woodcock DJ, Komorowski M, et al. quantifying intrinsic and extrinsic noise in gene transcription using the linear noise approximation: an application to single cell data. *Ann Appl Stat* 2013; 7: 1960–1982.
13. Leander J, Lundh T and Jirstrand M. Stochastic differential equations as a tool to regularize the parameter estimation problem for continuous time dynamical systems given discrete time measurements. *Math Biosci* 2014; 251: 54–62. [PubMed: 24631177]
14. Fearnhead P, Giagos V and Sherlock C. Inference for reaction networks using the linear noise approximation. *Biometrics* 2014; 70: 457–466. [PubMed: 24467590]
15. Doucet A, Godsill S and Andrieu C. On sequential Monte Carlo sampling methods for Bayesian filtering. *Stat Comput* 2000; 10: 197–208.
16. Raftery AE and Bao L. Estimating and projecting trends in HIV/AIDS generalized epidemics using incremental mixture importance sampling. *Biometrics* 2010; 66: 1162–1173. [PubMed: 20222935]
17. Arulampalam MS, Maskell S, Gordon N, et al. A tutorial on particle filters for online nonlinear/ non-Gaussian Bayesian tracking. *IEEE Trans Signal Process* 2002; 50: 174–188.
18. Silverman BW. *Density estimation for statistics and data analysis* In: *Monographs on statistics and applied probability*. London: Chapman and Hall; 1986..
19. Shaman J, Karspeck A, Yang W, et al. Real-time influenza forecasts during the 2012–2013 season. *Nat Commun* 2013; 4: 2837. [PubMed: 24302074]
20. Gillespie DT. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J Comput Phys* 1976; 22: 403–434.
21. Hoops S, Sahle S, Gauges R, et al. COPASI - a complex pathway simulator. *Bioinformatics* 2006; 22: 3067–3074. [PubMed: 17032683]
22. Mathematica, Version 10.4, Wolfram Research, Inc., Champaign, IL (2016).
23. Centers for Disease Control and Prevention. Flu activity and surveillance, www.cdc.gov/flu/weekly/fluactivitysurv.htm (accessed 1 October 2018).
24. Hickmann KS, Fairchild G, Priedhorsky R, et al. Forecasting the 2013–2014 influenza season using wikipedia. *PLOS Comput Biol* 2015, 10.1371/journal.pcbi.1004239.

25. Xia Y, et al. *NLDAS primary forcing data L4 hourly 0.125 × 0.125 degree V002*. NCEP/EMC (2009). Greenbelt, MD: Goddard Earth Sciences Data and Information Services Center, 2016, <https://commons.wikimedia.org/wiki/File:PNGedUSoutline.png>.
26. Codomo D U.S. map: Wikipedia 2015..
27. *Data source: Google Flu Trends*, <http://www.google.org/flutrends> (accessed 1 October 2018).
28. Carneiro HA and Mylonakis E. Google trends: a web-based tool for real-time surveillance of disease outbreaks. *Clin Infect Dis* 2009; 49: 1557–1564. [PubMed: 19845471]
29. Ginsberg J, Mohebbi MH, Patel RS, et al. Detecting influenza epidemics using search engine query data. *Nat Lett* 2009; 457: 1012–1014.
30. Preis T and Moat HS. Adaptive nowcasting of influenza outbreaks using Google searches. *R Soc Open Sci* 2014, DOI: 10.1098/rsos.140095.
31. Cho S, Sohn CH, Jo MW, et al. Correlation between national influenza surveillance data and Google trends in South Korea. *PLOS One* 2013, 10.1371/journal.pone.0081422.
32. Kang M, Zhong H, He J, et al. Using Google trends for influenza surveillance in South China. *PLOS One* 2013, 10.1371/journal.pone.0055205.
33. Martin LJ, Lee BE and Yasui Y. Google Flu Trends in Canada: a comparison of digital disease surveillance data with physician consultations and respiratory virus surveillance data, 2010–2014. *Epidemiol Infect* 2016; 144: 325–332. [PubMed: 26135239]
34. Pollett S, Boscardin WJ, Azziz-Baumgartner E, et al. Evaluating Google Flu Trends in Latin America: important lessons for the next phase of digital disease detection. *Clin Infect Dis* 2017; 64: 34–41. [PubMed: 27678084]
35. Wilson N, Mason K, Tobias M, et al. Interpreting “Google Flu Trends” data for pandemic H1N1 influenza: the New Zealand experience. *Euro Surveill* 2009; 14: pii = 19386, 10.2807/ese.14.44.19386-en.
36. Centers for Disease Control and Prevention. Overview of influenza surveillance in the United States. Washington, DC: Centers for Disease Control and Prevention, 2017.
37. Yang W, Lipsitch M and Shaman J. Inference of seasonal and pandemic influenza transmission dynamics. *Proc Natl Acad Sci* 2015; 112: 2723–2728. [PubMed: 25730851]
38. Boëlle PY, Ansart S, Cori A, et al. Transmission parameters of the A/H1N1 (2009) influenza virus pandemic: a review. *Influenza Other Resp Virus* 2011; 5: 306–316.
39. Fraser C, Donnelly CA, Cauchemez S, et al. Pandemic potential of a strain of influenza A (H1N1): early findings. *Science* 2009; 324: 1557–1561. [PubMed: 19433588]
40. Yang Y, Sugimoto JD, Halloran ME, et al. The transmissibility and control of pandemic influenza A (H1N1) virus. *Science* 2009; 326: 729–733. [PubMed: 19745114]
41. Grima R An effective rate equation approach to reaction kinetics in small volumes: theory and application to biochemical reactions in nonequilibrium steady-state conditions. *J Chem Phys* 2010; 133: 035101. [PubMed: 20649359]
42. Thomas P, Matuschek H and Grima R. Intrinsic noise analyzer: a software package for the exploration of stochastic biochemical kinetics using the system size expansion. *PLOS One* 2012; 7: e38518. [PubMed: 22723865]
43. Olson DR, Konty KJ, Paladini M, et al. Reassessing Google Flu Trends data for detection of seasonal and pandemic influenza: a comparative epidemiological study at three geographic scales. *PLOS Comput Biol* 2013, 10.1371/journal.pcbi.1003256.
44. Cook S, Conrad C, Fowlkes AL, et al. Assessing Google Flu Trends performance in the United States during the 2009 influenza virus A (H1N1) pandemic. *PLOS one* 2011, 10.1371/journal.pone.0023610.

1. Initialization
 - (a) Choose an initial prior probability function $\pi_0((\theta, \nu_0))$ for the parameter θ and the initial states ν_0 .
 - (b) Choose an initial set of parameters $\{\theta_1, \dots, \theta_M\}$ and initial states $\{\nu_{0,1}, \dots, \nu_{0,M}\}$.
2. Calibration: For each observation y_i , $i \in \{1, 2, \dots\}$,
 - (a) Calculate the probability $\mathcal{P}(y_i|y_1, \dots, y_{i-1}; \theta_m)$ for every $m = 1, \dots, M$ (using Eq. (2) and the method we described).
 - (b) Update the parameter posterior distribution:

$$\pi_i(\theta|y_1, \dots, y_i) \leftarrow \mathcal{P}(y_i|y_1, \dots, y_{i-1}; \theta_m) \pi_{i-1}(\theta|y_1, \dots, y_{i-1})$$
 for every $m = 1, \dots, M$.
 - (c) Update the belief state $\Pi(\nu_i|y_1, \dots, y_i; \theta_m)$ for every $m = 1, \dots, M$ (through Eq. (5)).
 - (d) **Check whether degeneracy occurs by i) comparing the effective sample size N_{eff} (Eq. (8)) with a threshold N_C or ii) by comparing the expected fraction of unique points after re-sampling with a threshold Q in Eq. (11).**
 - (e) **If step (d) detects filter degeneracy: i) use a density function $\hat{f}$ (Eq. 9) to re-sample new parameters or ii) use the IMIS algorithm.**

Figure 1.

An algorithm for real-time calibration of stochastic compartmental epidemic models including mechanisms against filter degeneracy.

IMIS: Incremental Mixture Importance Sampling.

- M : number of parameter vectors to sample to describe the parameter distribution π . In the absence of mechanism to resolve filter degeneracy, M should be set to a large number so that the parameter posterior distribution is not degenerate. When a mechanism against filter degeneracy is implemented, M could be set to a smaller number and thus be less computationally expensive. We chose $M = 1000$ as in [4].
- M' : number of forward simulations to estimate and form prediction intervals for the epidemic prediction targets. M' should be chosen such that increasing it any further will not meaningfully change the prediction intervals. We chose $M' = 100$ as in [4].
- Observation variance for observation probability P in Eq. (2): set to 100 for the simulation experiments, and set to 10 when applying the MSS method to ILI and GFT data. This number was chosen based on visual investigation of the model fit against the ILI and GFT data in a select number of cities.

Using effective sample size to resolve filter degeneracy:

- H : bandwidth for density estimation. It was determined according to the Silverman's method [18] implemented in Mathematica or the approach described in the text.
- N_C : critical value of N_{eff} to trigger filter degeneracy ($1 < N_C < M$). A higher N_C implies a stricter mechanism against degeneracy at a larger computational cost. We chose $N_C = M/2$.
- $\tilde{M}$: number of parameter vectors being re-sampled with replacement to resolve degeneracy. We chose $\tilde{M} = M$.

Using IMIS algorithm to resolve filter degeneracy:

- Q : critical threshold to trigger the IMIS algorithm. We chose $Q = 1 - 1/\exp(1)$ as suggested by [16].
- B : number of parameter vectors re-sampled by the IMIS algorithm. We chose $B = 1/10M = 100$ as suggested by [16].

Figure 2.

List of free parameters of the MSS method and the two algorithms to resolve filter degeneracy.

MSS: Multiple Shooting for Stochastic Systems; ILI: influenza-like illness; GFT: Google Flu Trends; IMIS: Incremental Mixture Importance Sampling.

Figure 3.
Fifty stochastic trajectories produced by simulating the humidity-based SIRS model.

Figure 4.

Effectiveness of the proposed mechanisms to address filter degeneracy: For the original MSS method (represented by the blue curve), the effective sample size (N_{eff}) and the expected fraction of unique points $Q(\tilde{w})$ approaches zero as the model is recalibrated based on the new epidemic observations. This suggests the occurrence of filter degeneracy. This phenomenon is avoided in both new versions of the MSS method (represented by the red and green curves). Solid lines are created by calculating the average of N_{eff} and $Q(\tilde{w})$ over 50 simulated trajectories, and shaded areas represent the 90% confidence region.

MSS: Multiple Shooting for Stochastic Systems; IMIS: Incremental Mixture Importance Sampling.

Figure 5.

Predicting epidemic targets in two simulated trajectories. Each row represents one simulated epidemic. Black crosses represent the true value of the prediction targets observed in the simulated trajectory and red lines represent the mean posterior estimate proposed by the MSSb-Neff method using the epidemic observations until the week displayed on the x-axis. The red region shows the 90% prediction intervals.

Figure 6.

Coverage of the 90% prediction intervals and RE produced by the MSS versions for several prediction targets at different times relative to peak for the simulated trajectories shown in Figure 3. (a) Proportion of 90% prediction intervals that included the true value of the prediction target. If this fraction (which is a random variable itself) falls in the region between the two dotted lines, it suggests that the estimated prediction intervals have the appropriate level of coverage. The dotted lines are calculated as the 95% percentile interval of a binomial distribution with the probability 0.9 and sample size 50 (number of simulated trajectories in Figure 3) and then divided by the sample size 50. (b) RE for the mean of the posterior distribution for predicting “Infection prevalence,” “Number of cases at the epidemic peak,” “Number of new cases over the next week,” and “Final epidemic size.”

MSS: Multiple Shooting for Stochastic Systems; RE: relative error.

Figure 7.

Estimating epidemic parameters of two simulated trajectories. Each row represents one simulated epidemic. Black crosses represent the true value of the parameter in the simulated trajectory and red lines represent the mean posterior estimate proposed by the MSSb-Neff method using the cumulative epidemic observations up to the epidemic week displayed on the x-axis. The red region shows the 90% credible intervals.

Figure 8.

Coverage of the 90% credible intervals and RE produced by the MSS versions for estimates of model parameters at different times relative to peak for simulated trajectories shown in Figure 3. (a) Proportion of 90% credible intervals that included the true value of the prediction target. If this fraction (which is a random variable itself) falls in the region between the two dotted lines, it suggests that the estimated credible intervals have the appropriate level of coverage. The dotted lines are calculated as the 95% percentile interval of a binomial distribution with the probability 0.9 and sample size 50 (number of simulated trajectories in Figure 3) and then divided by the sample size 50. (b) RE for the mean of the posterior distribution for estimating “Duration of Immunity,” “Duration of Infectiousness,” number of “Initial Infected,” and “ $R_0(t)$.”

MSS: Multiple Shooting for Stochastic Systems; RE: relative error.

Figure 9.

One week predictions of ILI and the corresponding prediction intervals for the 2013 season. The x-axes represent epidemic weeks and the y-axes show the percent ILI cases amongst total patient visits during each week. Prediction for a given epidemic week is carried out with the knowledge accumulated up to the past week. Black crosses represent the observed values, the red line the mean posterior, the dark red area the 50% prediction region, and the light red area the 90% prediction region. The underlying U.S. map's source is Codomo.²⁶

Figure 10.

One week predictions of ILI and the corresponding prediction intervals for the 2009 pandemic. The x-axes represent epidemic weeks and the y-axes show the percent ILI cases amongst total patient visits during each week. Prediction for a given epidemic week is carried out with the knowledge accumulated up to the past week. Black crosses represent the observed values, the red line the mean posterior, the dark red area the 50% prediction region, and the light red area the 90% prediction region. The underlying U.S. map's source is Codomo.²⁶

Figure 11.

One week predictions of ILI+ and the corresponding prediction intervals for the 2013 influenza season. The x-axes represent epidemic weeks and the y-axes show the number of new influenza cases arising during each week. Prediction for a given epidemic week is carried out with the knowledge accumulated up to the past week. Black crosses represent the observed values, the red line the mean posterior, the dark red area the 50% prediction region, and the light red area the 90% prediction region. The underlying U.S. map's source is Codomo.²⁶

Figure 12.

One week predictions of ILI+ and the corresponding prediction intervals for the 2009 influenza pandemic. The x-axes represent epidemic weeks and the y-axes show the number of new influenza cases arising during each week. Prediction for a given epidemic week is carried out with the knowledge accumulated up to the past week. Black crosses represent the observed values, the red line the mean posterior, the dark red area the 50% prediction region, and the light red area the 90% prediction region. The underlying U.S. map's source is Codomo.²⁵

Figure 13.

Distribution of R_0 estimates calculated for the included cities between the season 2003/04 and 2013/14. Here, w1 and w2 stands for the first and the second pandemic waves of 2009 and w1w2 represents the full pandemic period which includes both waves of 2009.