

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

Christoph Zimmer<sup>(1,3)</sup>, Sequoia I. Leuba<sup>(1)</sup>, Ted Cohen<sup>(1)</sup> and Reza Yaesoubi<sup>(2)</sup>

<sup>1</sup> Epidemiology of Microbial Diseases, Yale School of Public Health, New Haven, CT, USA

<sup>2</sup> Health Policy and Management, Yale School of Public Health, New Haven, CT, USA

<sup>3</sup> Bosch Center for Artificial Intelligence, Robert Bosch GmbH, Renningen, Germany

Email: christoph.zimmer@yale.edu; christoph.zimmer@de.bosch.com

\*Corresponding author

**Statistical Methods in Medical Research**

## Supporting Information:

### Interpolation of humidity data

Visual inspection of the humidity indicates a sinusoidal shape, and hence, we chose a sinusoidal interpolation:

$$\bar{q}(t) = a \sin(b(t - c)) + d.$$

As the daily data is averaged over multiple years, the desired period of the sine function is 365 days, and therefore,  $b = 2\pi/365$ . The coefficients  $a$ ,  $c$ , and  $d$  are determined by fitting the above function to the humidity data using a least squares method.

### Technical details

To calculate the number of recoveries, we introduce an auxiliary equation

$$\frac{dR^*}{dt} = I/\gamma, R^*(0) = 0$$

which we initialize with 0 on each interval. This “additional” compartment  $R^*$  counts the number of people recovering in each week.

When we analyzed the GFT data, we chose a Gaussian distribution to derive the state estimates for the MSSa approach. As we wanted to minimize the impact of the Gaussian distribution assumption, we did not use any covariance update (as in MSSb). The MSSa approach is in this manuscript always combined with the effective sample size based mechanism against filter degeneracy.

### IMIS details

In each iteration  $k$  the IMIS algorithm

1. chooses the parameter with the maximal weight from the weights  $w_1, \dots, w_{B_k}$  as center  $\vartheta^{(k)}$ . Next, a covariance  $\Sigma_{IMIS}^{(k)}$  is estimated from the  $1 \leq B \leq M$  inputs with the smallest Mahalanobis distance to the center  $\vartheta^{(k)}$ . The Mahalanobis distance is calculated as the squared difference of the parameters weighted with the prior variance.

2. Next,  $B$  new inputs are suggested by sampling from a normal distribution  $H_k$  with mean  $\vartheta^{(k)}$  and covariance  $\Sigma_{IMIS}^{(k)}$ . The weights for these parameters are determined by calculating the new weights

$$w_j^{new} = \prod_{s=1}^i \mathcal{P}(y_s | y_1, \dots, y_{i-1}, \theta_j), \text{ for } j = B_{k-1} + 1, \dots, B_k \quad (16)$$

3. The new importance weights are then determined as

$$w_j^{(k)} = c \cdot w_j^{new} \frac{\pi_0(\theta_j)}{q^{(k)}(\theta_j)}, \text{ for } j = 1, \dots, B_k$$

where  $c$  is a normalizing constant and  $q^{(k)} = \frac{M}{B_k} \pi_0 + \frac{B}{B_k} \sum_{s=1}^k H_s$ .

## Resolving filter degeneracy versus using a finer grid

As pointed out in the results section, it is possible to overcome the effects of filter degeneracy by using a finer grid. Here, we use the MSS method without filter degeneracy with 10000 parameter samples (calling it MSSold-10000) instead of 1000 (MSSold) before:

Figure A1 shows that MSSb-Neff and MSSold-10000 are almost equally good in terms of coverage. That seems to show that the effects of filter degeneracy have been reduced by using more parameter samples. However, Figure A3 shows that filter degeneracy has not been overcome and is still strongly present. Having 10000 parameter samples increases the number of parameters that have significant weight (on average 6 parameters) which overcomes the effect of filter degeneracy on coverage, but not the effect of filter degeneracy itself. In addition, computational time increased as shown in Figure A2. In larger models, we expect the advantage of MSSb-Neff to be even stronger as so few parameters with non-negligible weight might not be enough to capture the dynamics.

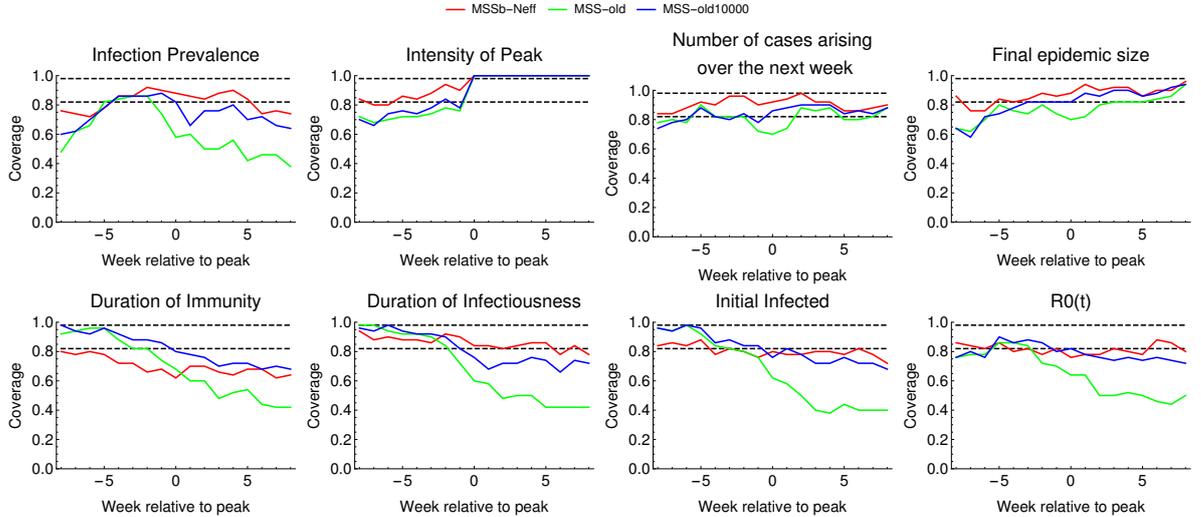


Figure A1: Coverage of the 90% credible intervals produced by each MSS method for prediction of prediction targets and estimates of model parameters at different times relative to peak for simulated trajectories shown in Fig. 3

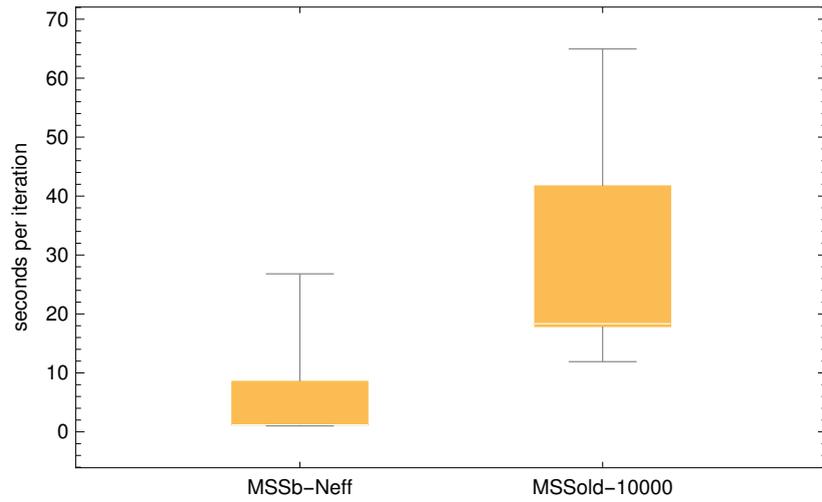


Figure A2: **Comparison of computational time: computational time of MSSb-Neff and MSSold-10000 in seconds per iteration.**

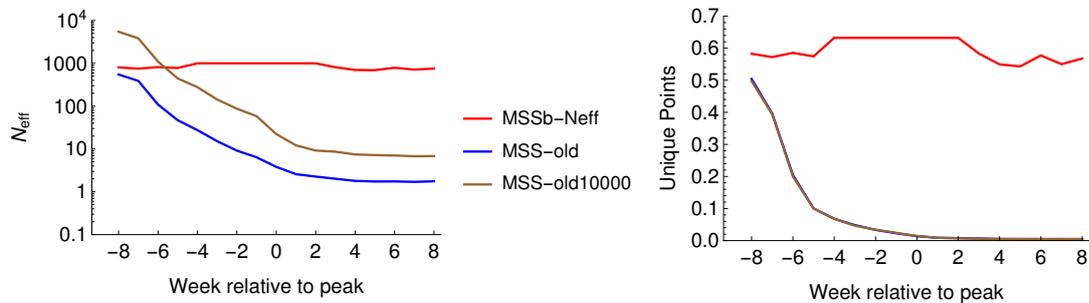


Figure A3: **Demonstrating the effects of increasing the size of parameter samples on filter degeneracy.** Average value of  $N_{eff}$  (left) and fraction of unique points,  $Q(\tilde{w})$ , (right) over 50 runs.

## Supporting Figures

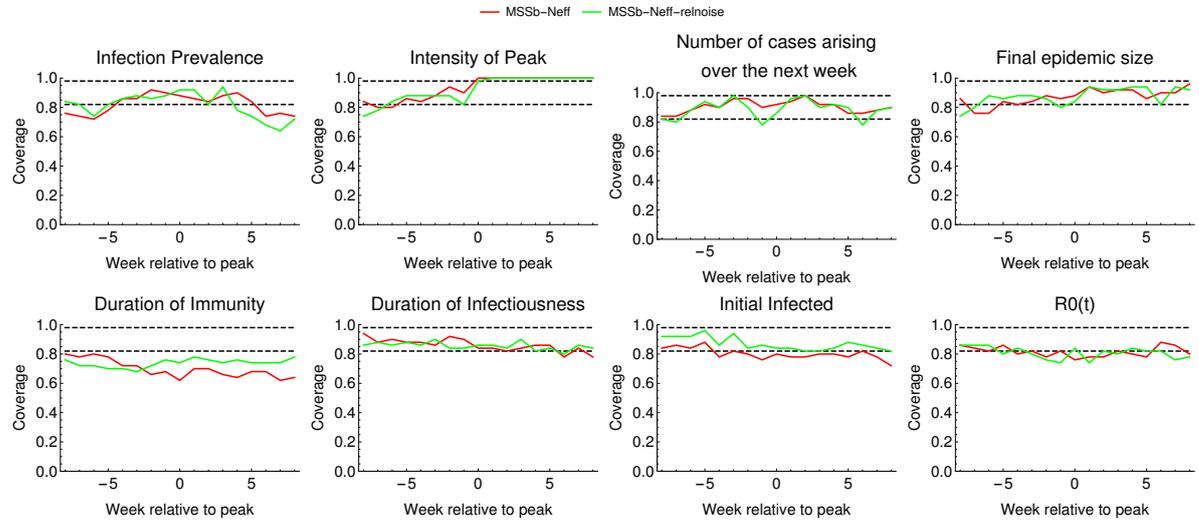


Figure A4: Coverage of the 90% credible and prediction intervals produced by the MSSb-Neff method for prediction of prediction targets and estimates of model parameters at different times relative to peak for simulated trajectories with relative noise of 10% in green (MSSb-Neff-relnoise). The results for additive noise are added for comparison in red (MSSb-Neff) as in Figure 6 and 8.

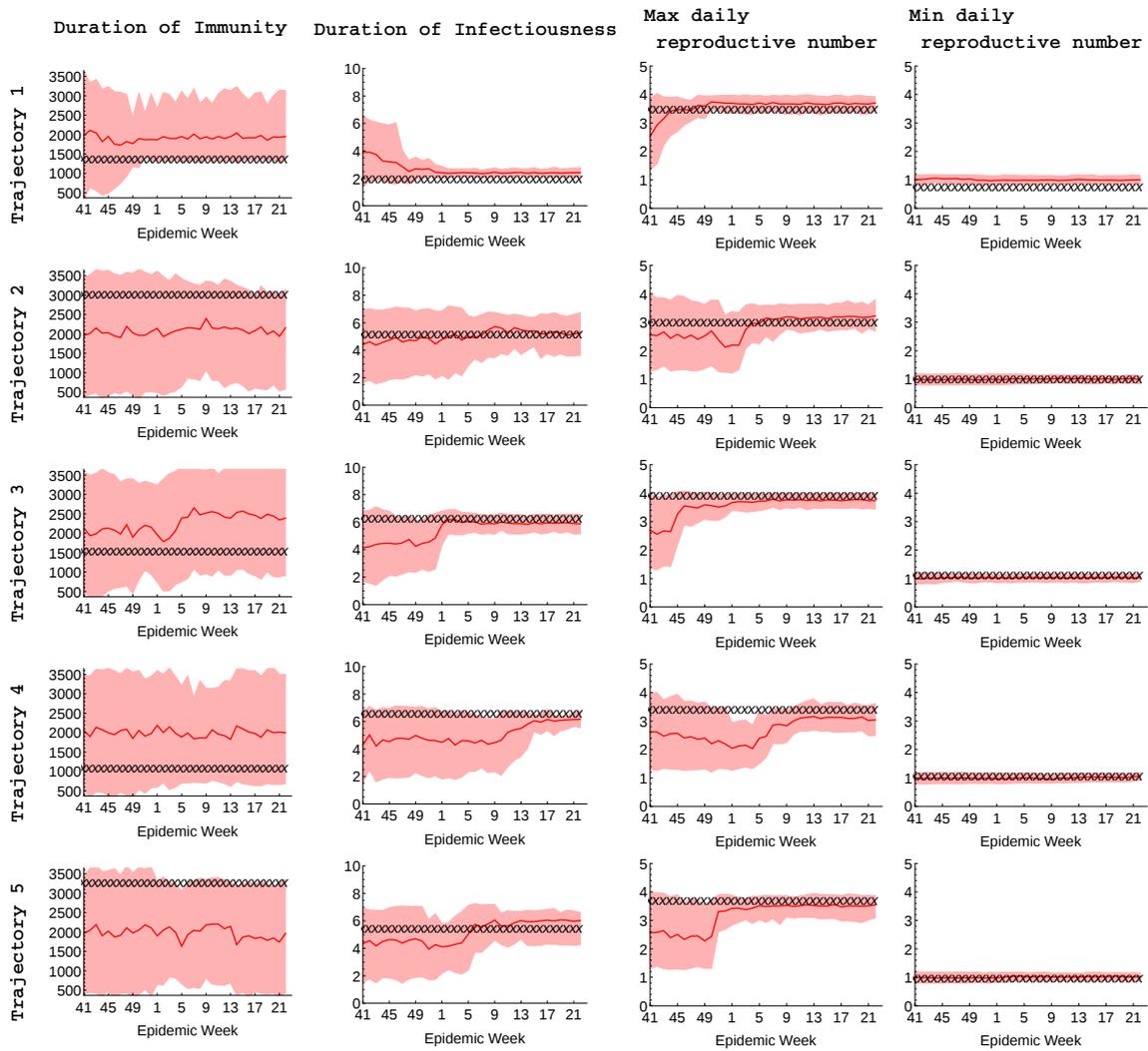


Figure A5: **Estimates and credible intervals for key epidemic parameters for 5 simulated trajectories:** The black crosses show the true value for the target and the red line shows the mean posterior estimate given the information until the time point displayed on the x-axis. The red region shows the 90% credible intervals.

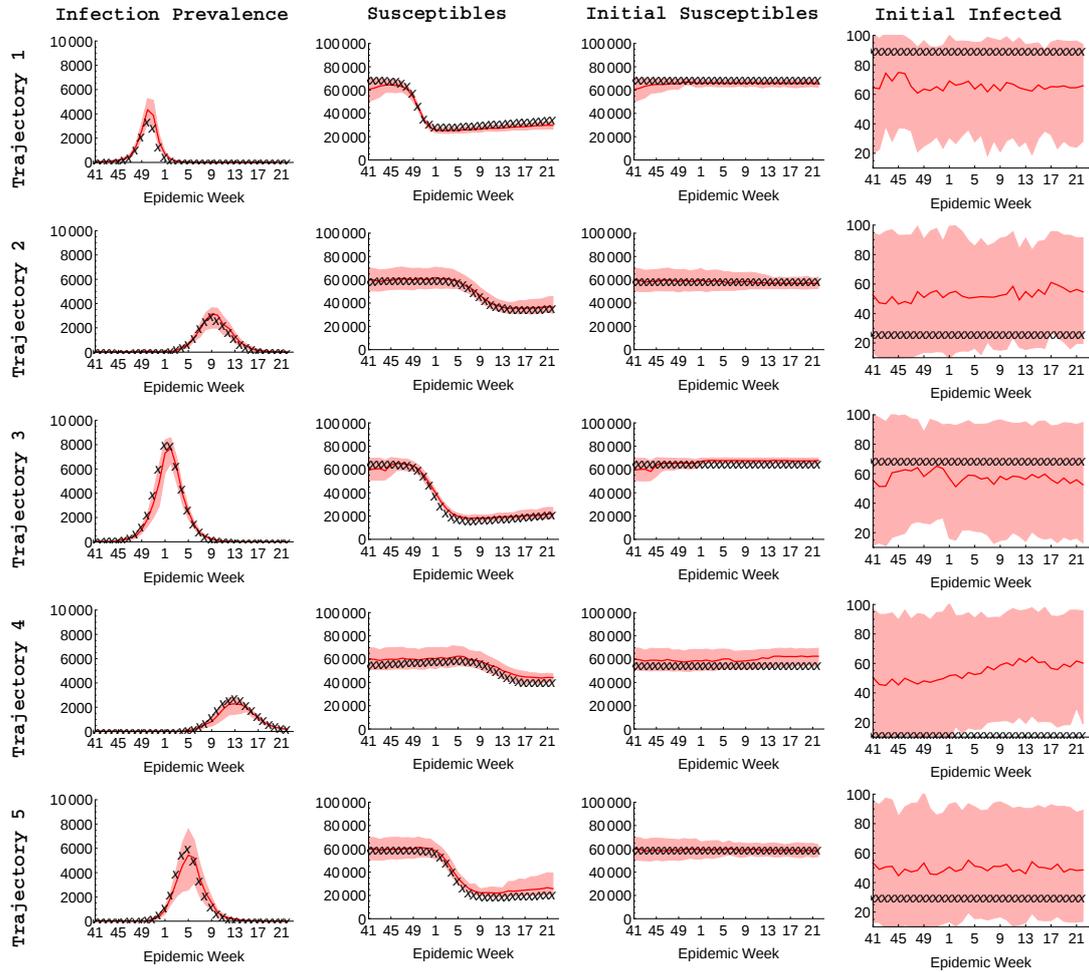


Figure A6: **Estimates and credible intervals for epidemic states using 5 simulated trajectories:** The black crosses show the true value for the target and the red line shows the mean posterior estimate given the information until the time point displayed on the x-axis. The red region shows the 90% credible intervals.

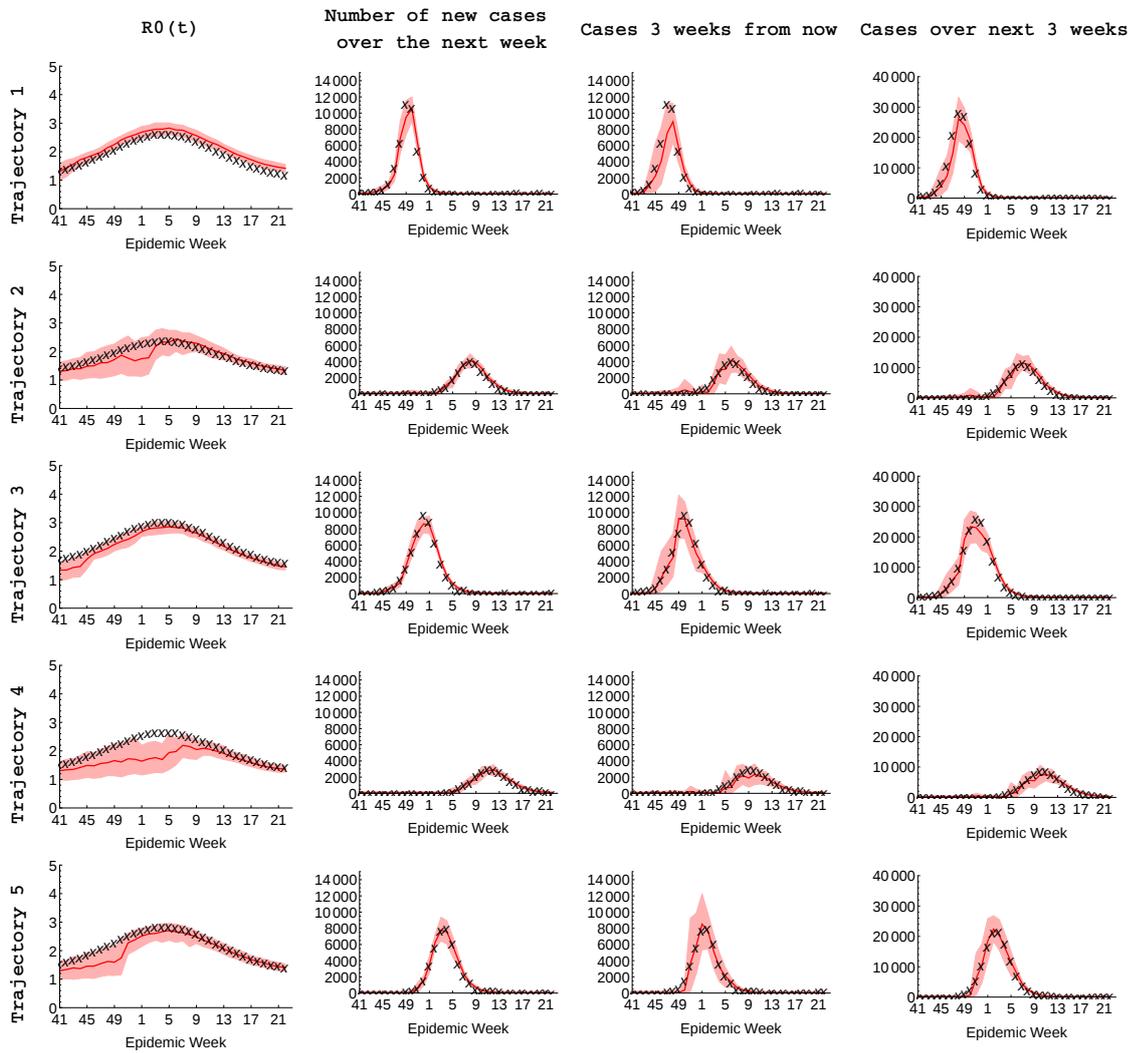


Figure A7: **Short-term predictions and the corresponding prediction intervals for 5 simulated trajectories:** The black crosses show the true value for the target and the red line shows the mean posterior estimate given the information until the time point displayed on the x-axis. The red region shows the 90% posterior prediction interval.

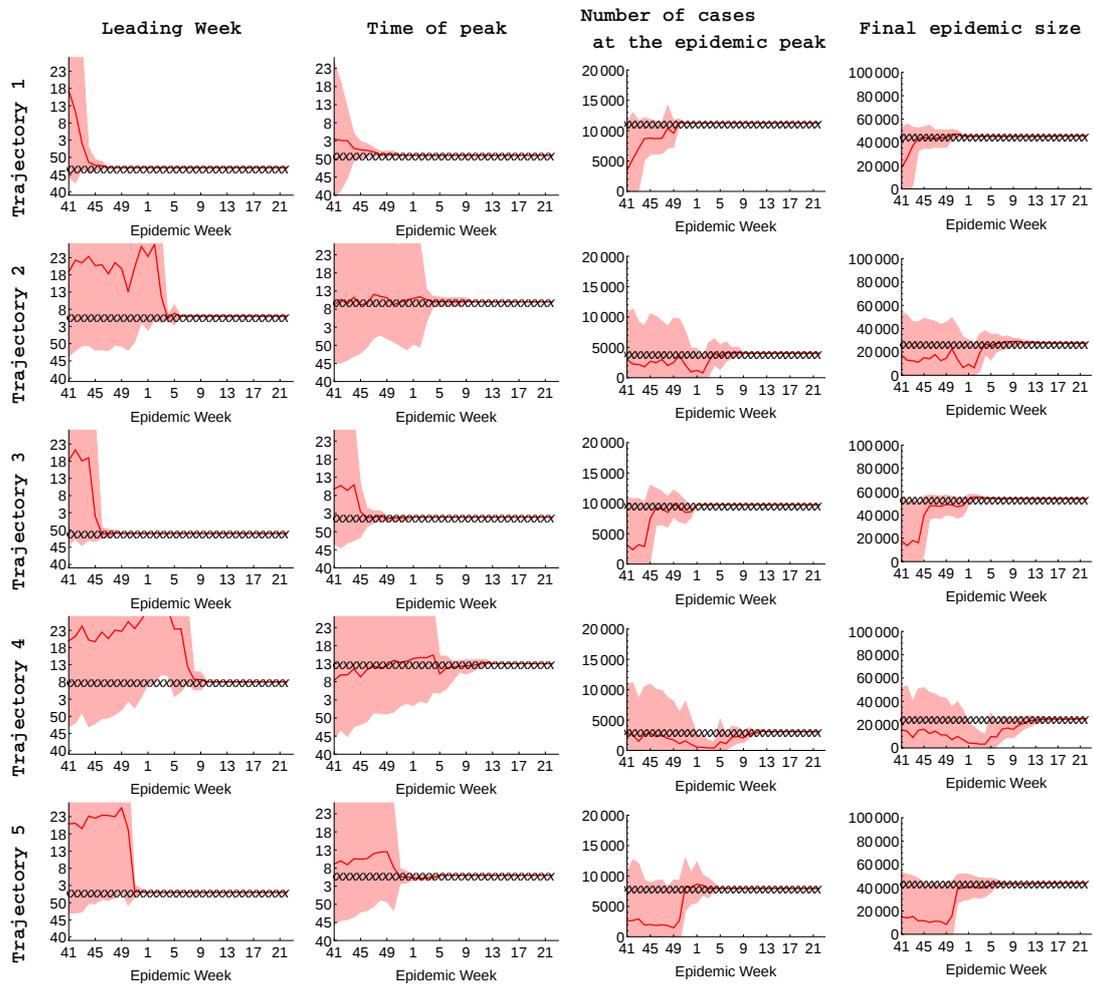


Figure A8: **Mid- and long-term predictions and the corresponding prediction intervals for 5 simulated trajectories:** The black crosses show the true value for the target and the red line shows the mean posterior prediction given the information until the time point displayed on the x-axis. The red region shows the 90% posterior prediction interval. Leading week is defined as the first of three weeks with number of cases larger than 1,000 per 100,000 population.

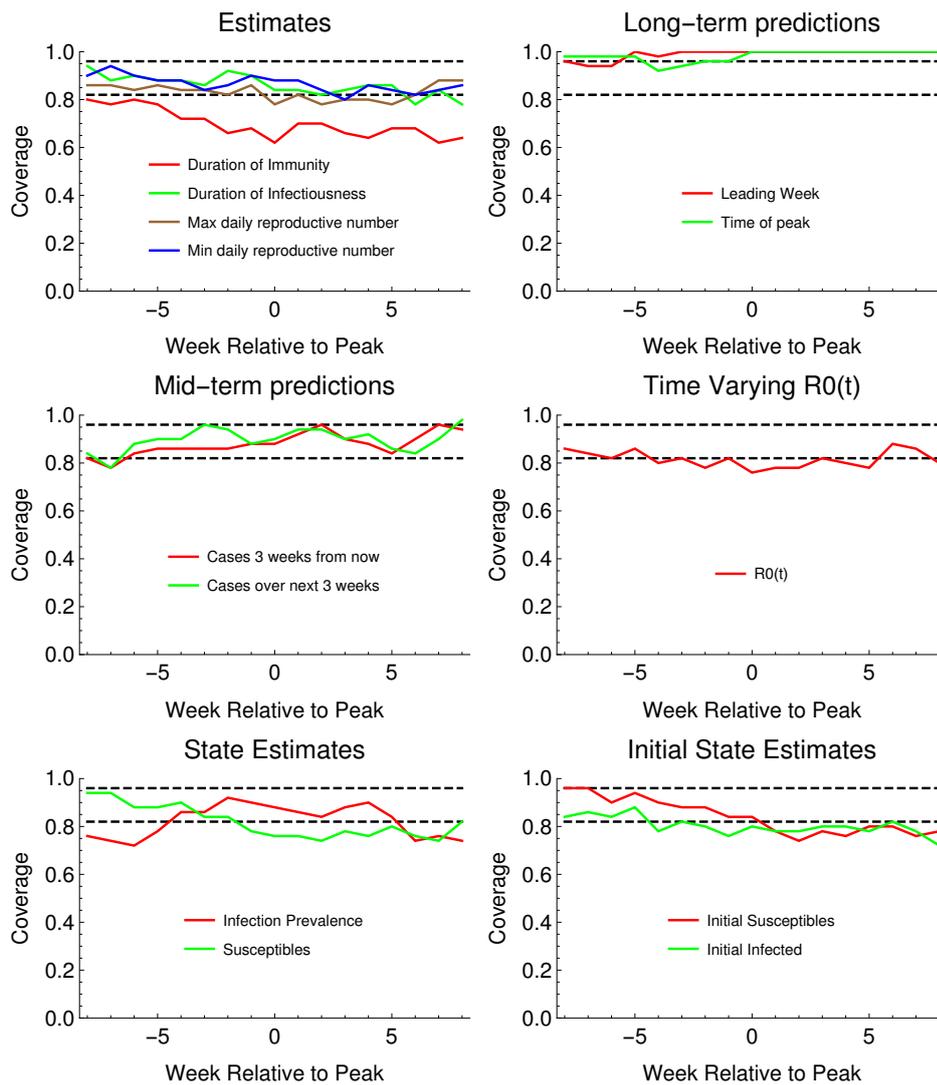


Figure A9: Coverage of the 90% credible and prediction intervals produced by the MSSb-Neff method for several prediction targets at different times to peak for the simulated trajectories shown in Fig. (3): Proportion of 90% credible and prediction intervals that included the true value of the prediction target. If this fraction (which is a random variable itself) falls in the region between the two dotted lines, it suggests that the produced credible or prediction intervals present a desired 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, and then the percentile interval is divided by 50.

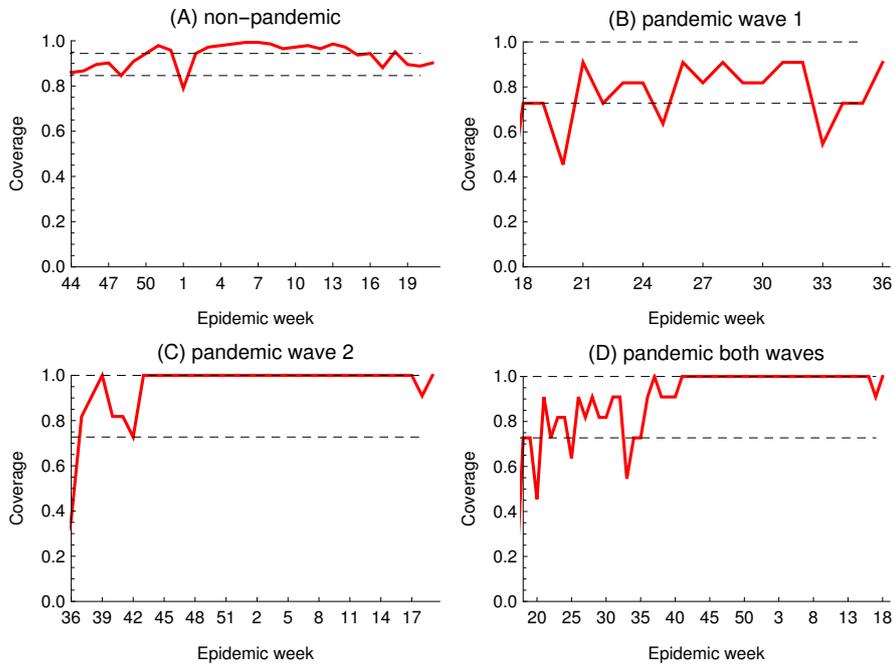


Figure A10: **Coverage of the 90% prediction intervals produced by the MSSa method for predicting the next weeks ILI over all regions and years:** Proportion of 90% posterior prediction intervals that included the true value of the prediction target. If this fraction (which is a random variable itself) falls in the region between the two dotted lines, it suggests that the produced prediction intervals present a desired 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, and then the percentile interval is divided by 50. The

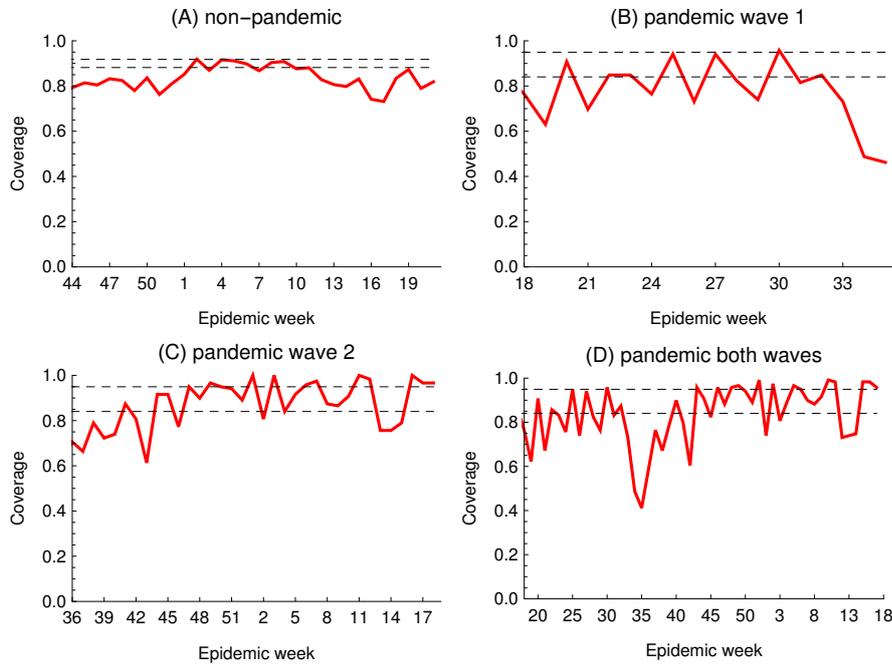


Figure A11: **Coverage of the 90% prediction intervals produced by the MSSa method for predicting the next week cases over all cities in the U.S.:** Proportion of 90% posterior prediction intervals that included the true value of the prediction target. If this fraction (which is a random variable itself) falls in the region between the two dotted lines, it suggests that the produced prediction intervals present a desired 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, and then the percentile interval is divided by 50. Panel (B) and (D) also suggest that coverage decreases around EW 35-36 of the pandemic year which is likely an effect of the unusual pandemic behavior not fully captured in the SIRS model.

## Supporting Files

Figure A12: **One week predictions of ILI and corresponding prediction intervals for all regions and national level:** The x-axes represent epidemic weeks and y-axes shows the percent of ILI amongst total patient visits. Prediction for a given epidemic week is carried out with the knowledge up to the past week. Black crosses represent the observed values, the red line is the mean posterior, the dark red area is the 50% prediction region, and the light red area is the 90% prediction region. FILE: indiv\_all-ili.pdf

Figure A13: **One week predictions of ILI+ and corresponding prediction intervals for all seasons and U.S. cities:** The x-axes represent epidemic weeks and y-axes show the number of new influenza cases during each week. Prediction for a given epidemic week is carried out with the knowledge up to the past week. Black crosses represent the observed values, the red line is the mean posterior, the dark red area is the 50% prediction region, and the light red area is the 90% prediction region. FILE: indiv\_all.pdf

Reproducible computer code:

- File A: Simulation-Study.tar.gz, reproducible computer code for the results in the simulation study section
- File B: GFT-CDC.tar.gz, reproducible computer code for the results in the GFT data section and the CDC data section. This code contains components of GFT data [26], NASA humidity data [24], CDC data [22], and a Wikipedia map [25]. Re-use only permitted according to their policies.