Forecasting Frameworks

Dexter Barrows
March 7, 2016

1 Forecasting Setup

This section will focus on taking the stochastic SIR model from the previous section, truncating the synthetic data output from realizations of that model, and seeing how well IF2 and HMC can reconstruct out-of-sample forecasts.

An example of a simulated system with truncated data can be seen in Figure [1] below.

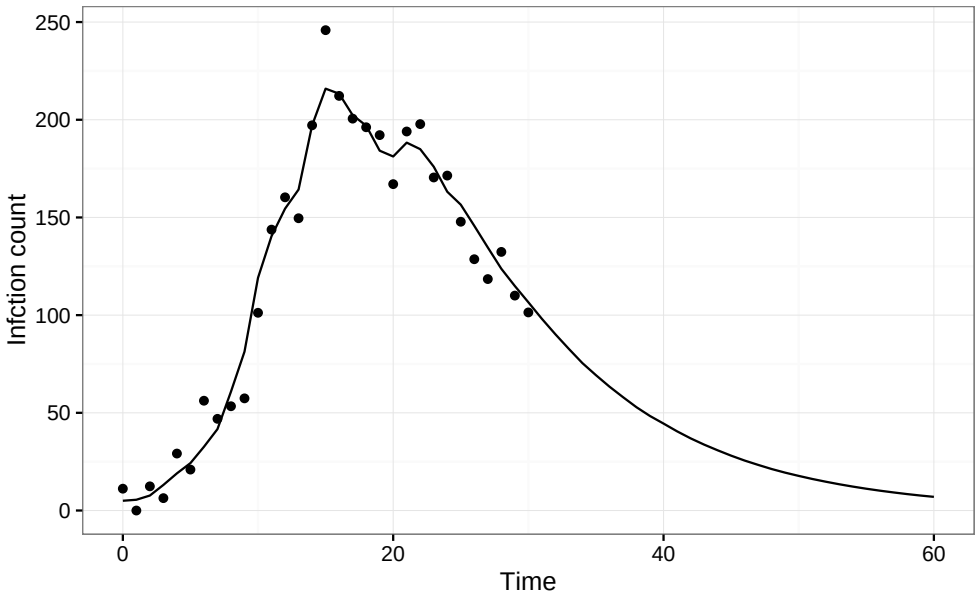

Figure 1: Infection count data truncated at $T = 30$. The solid line shows the true underlying system states, and the dots show those states with added observation noise. Parameters used were $R_0 = 3.0$, $r = 0.1$, $\eta = .05$, $\sigma_{proc} = 0.5$, and additive observation noise was draw from $\mathcal{N}(0, 10)$

As before, the solid line shows the true system states, while the dots show the data. In essence we want to be able to give either IF2 of HMC only the data points and have it reconstruct the entirety of the true system states.

2 IF2

For IF2, we will take advantage of the fact that the particle filter will produce state estimates for every datum in the time series given to it, as well as producing parameter maximum likelihood point estimates. Both of these sources of information will be used to produce forecasts by parametric bootstrapping using the final parameter estimates from the particle swarm after the last IF2 pass, then using the newly generated parameter sets along with the system state point estimates from the first fitting to perform normal bootstrapping into the future of the time series.

We will truncate the data at half the original time series length (to $T = 30$), and fit the model as previously described.

First, we can see the state estimates for each time point produced by the last IF2 pass in Figure [2] below.

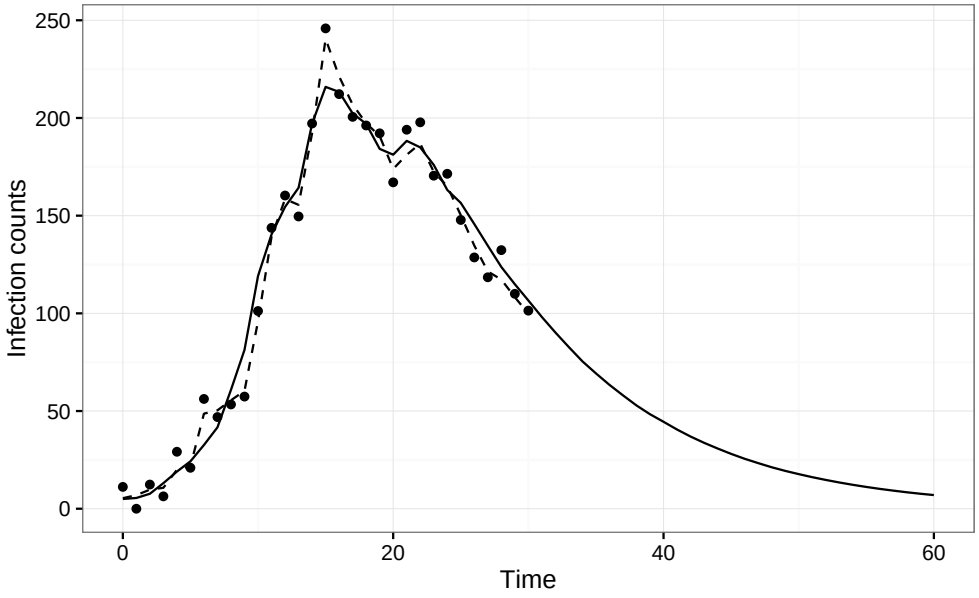

Figure 2: Infection count data truncated at $T = 30$ from Figure [1] above. The dashed line shows IF2's attempt to reconstruct the true underlying state from the observed data points.

Recall that IF2 is not trying to generate parameter estimation densities, but rather produce a point estimate. Since we wish to determine the approximate distribution of each of the parameters in addition to the point estimate, we must turn to another method, parametric bootstrapping.

2.1 Parametric Bootstrapping

The goal of the parametric bootstrap is use an initial density sample θ^* to generate further samples $\theta_1, \theta_2, ..., theta_M$. It works by using θ to generate artificial data sets $D_1, D_2, ..., D_M$ to which we can refit our model of interest and generate new parameter sets.

An algorithm for parametric bootstrapping using IF2 and our stochastic SIR model is shown in Algorithm [1].

Algorithm 1: Parametric Bootstrap

Input : Forward simulator $S(\theta)$, data set D

/* Initial fit */

$\theta^* \leftarrow IF2(D)$

/* Generate artificial data sets */

for $i = 1 : M$ **do**

  $D_i \leftarrow S(\theta^*)$

/* Fit to new data sets */

for $i = 1 : M$ **do**

  $\theta_i \leftarrow IF2(D_i)$

Output: Distribution samples $\theta_1, \theta_2, ..., \theta_M$

2.2 IF2 Forecasts

Using the parameter sets $\theta_1, \theta_2, ..., \theta_M$ and the point estimate of the state provided by the initial IF2 fit, we can use a normal bootstrap to produce estimates of the future state. A plot showing a projection of the data from the previous plots can be seen in Figure [3].

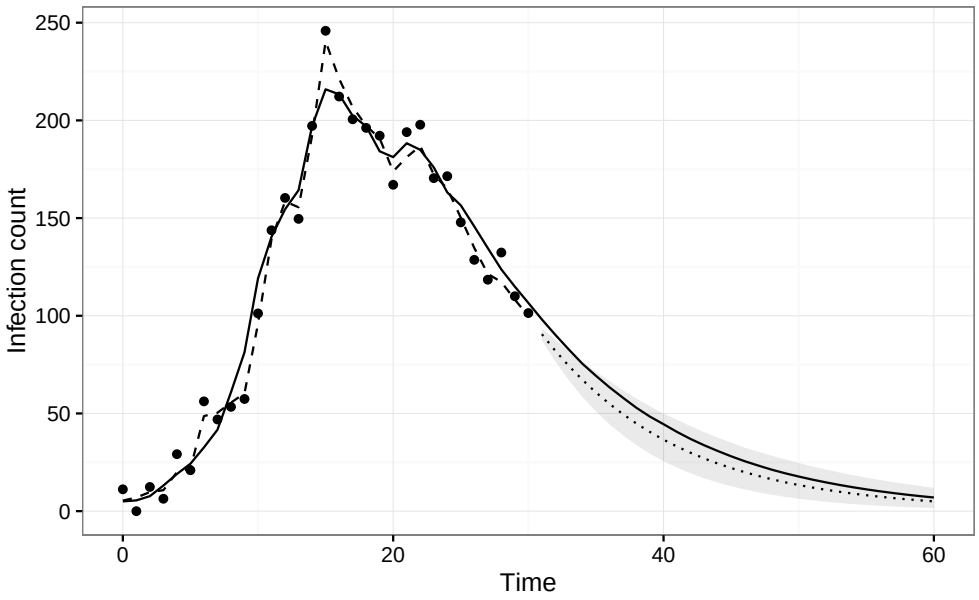

Figure 3: Forecast produced by the IF2 / parametric bootstrapping framework with $M = 200$ trajectories. The dotted line shows the mean estimate of the forecasts, and the grey ribbon shows the centre 95th quantile.

It should be noted that the estimates produced are without added observation noise. We can see a reproduction of the data Figure [3], but with observation noise in Figure [4].

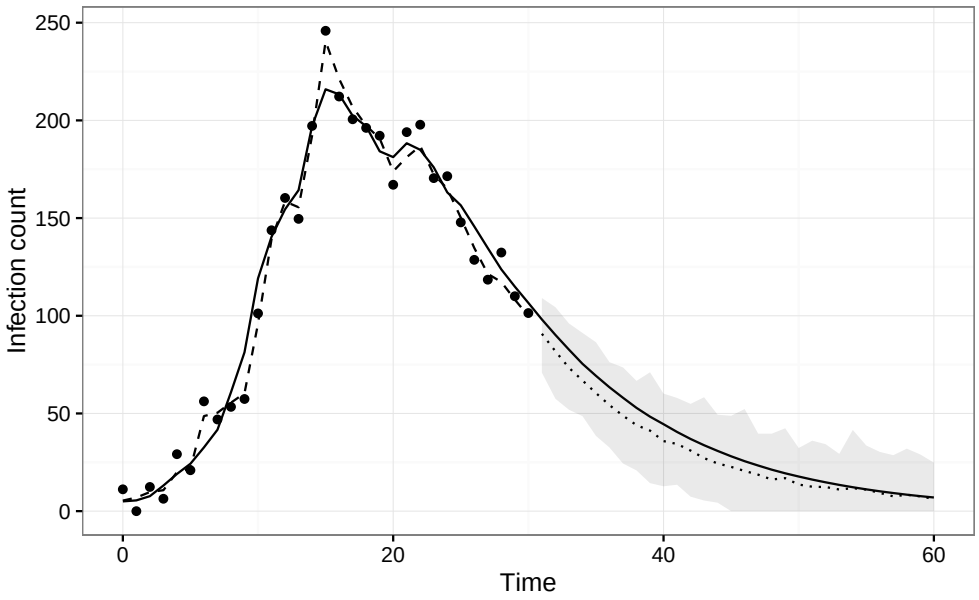

Figure 4: Forecast produced by the IF2 / parametric bootstrapping framework, with added observation noise drawn from the fitted parameter sets. The dotted line shows the mean estimate of the forecasts, and the grey ribbon shows the centre 95th quantile.

We can define a metric to gauge forecast effectiveness by calculating the SSE and dividing that value by the number of values predicted to get the average squared error per point. For the data in Figure [3] the value was $\overline{SSE} = 37.005$.

3 HMCMC

For HMC we cannot use final state estimates as in IF2 as we only have the parameter density estimates to work with. Instead we have to rely on a pure bootstrapping approach. We are, however, not just limited to the primary parameter set samples produced by Stan – the samples also contain estimates of the process noise present in the system, which we can use to help inform the forward simulation in portions where we have fitting data.

As before we fit the stochastic SIR model to the partial data, but now perform bootstrapping as described above, and obtain the plot in Figure [5].

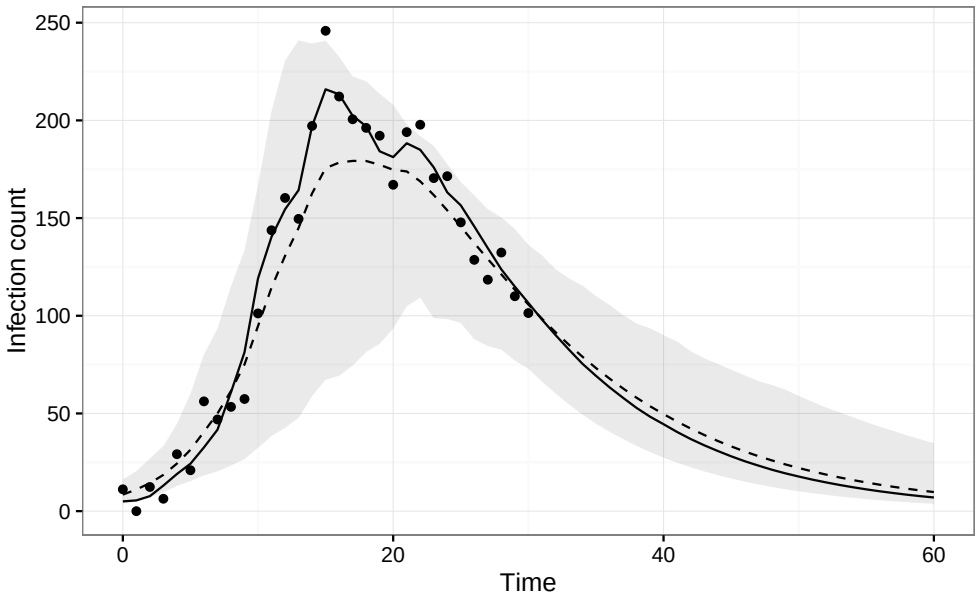

Figure 5: Forecast produced by the HMCMC / bootstrapping framework with $M = 200$ trajectories. The dotted line shows the mean estimate of the forecasts, and the grey ribbon shows the centre 95th quantile.

As before we can evaluate the averaged SSE of the forecast for the data shown, giving $\overline{SSE} = 17.74223$.

4 Truncation vs. Error

Of course the above mini-comparison only shows one truncation value for one trajectory. Really, we need to know how each method performs on average given different trajectories and truncation amounts. In effect we wish to "starve" each method of data and see how poor the estimates become with each successive data point loss.

Using each method, we can fit the stochastic SIR model to successively smaller time series to see the effect of truncation on forecast averaged SSE. This was performed with 10 new trajectories drawn for each of the desired lengths. The results are shown in Figure [6].

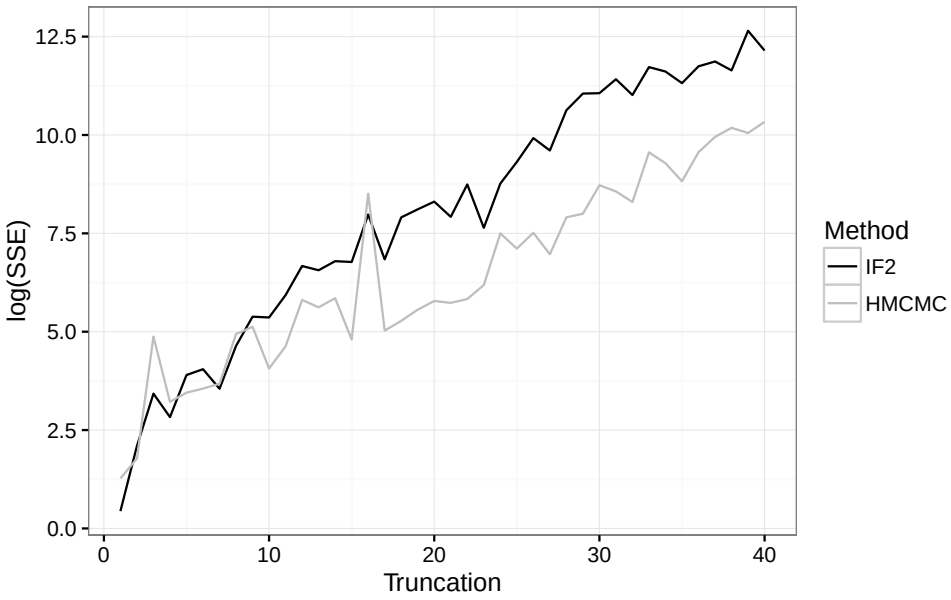

Figure 6: Error growth as a function of data truncation amount. Note that the y-axis shows the natural log of the averaged SSE, not the total SSE.

Here we have out first real result in this paper. HMCMC consistently outperforms IF2 for truncation levels more than about 8 or 9, and does so by an increasing margin as the truncation level increases.