Ebola Forecasting - Error Analysis

Andy Shen

Data Input and Cleaning

We assume the most accurate dataset is the most recent dataset of the outbreak. We tally the cases such that there is a running total of infections at each date.

We then import the projections from the Hawkes and Recursive models. For these predictions, the date preceding the forecasts is the last date of that dataset with at least one case. The forecasted numbers then predict the additional number of infections 7, 14, and 21 days, respectively, after that date.

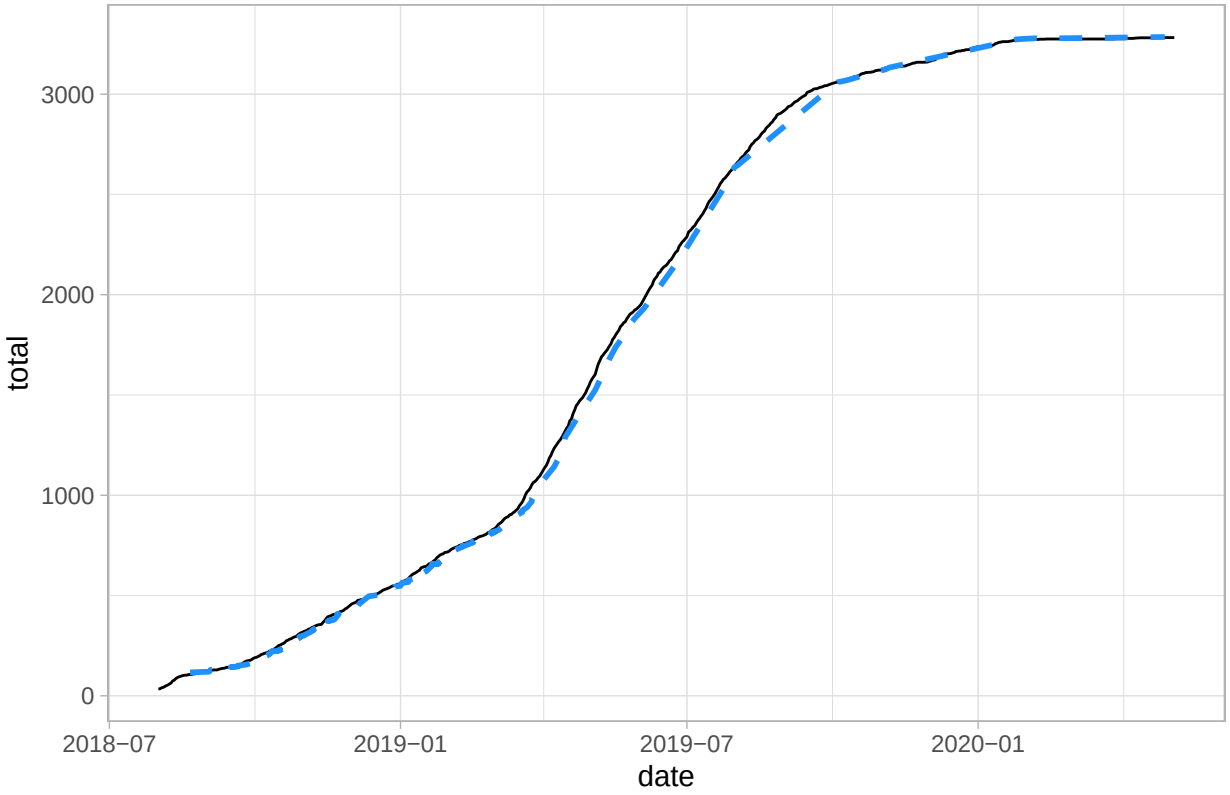
The code has been omitted for readability.

Hawkes Complete Outbreak Analysis

7-Day Forecast Analysis

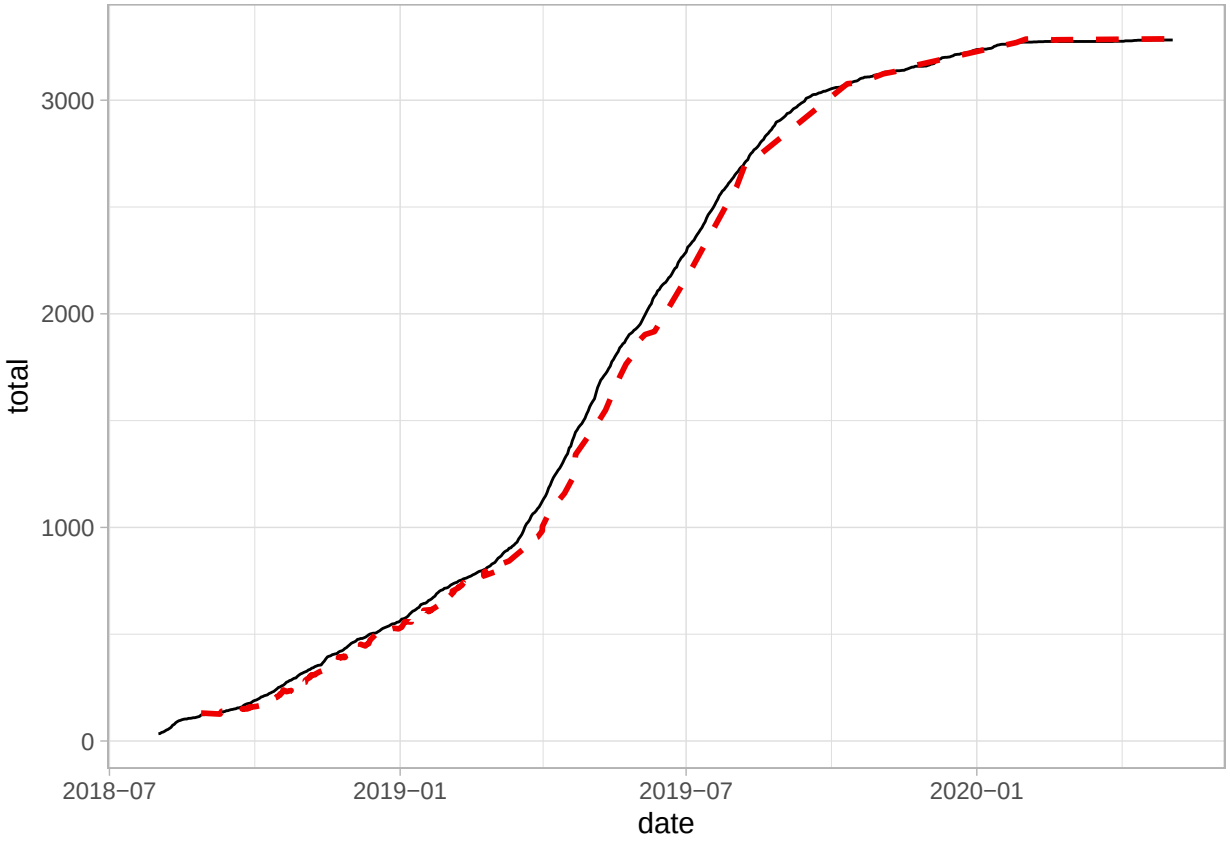
Figure 1 below shows the Hawkes 7-Day Forecasts for all simulations with respect to the recorded number of infections. The RMSE values for all forecasts are included in Table 1.

Figure 1: Hawkes 7−Day Forecasts for All Simulations



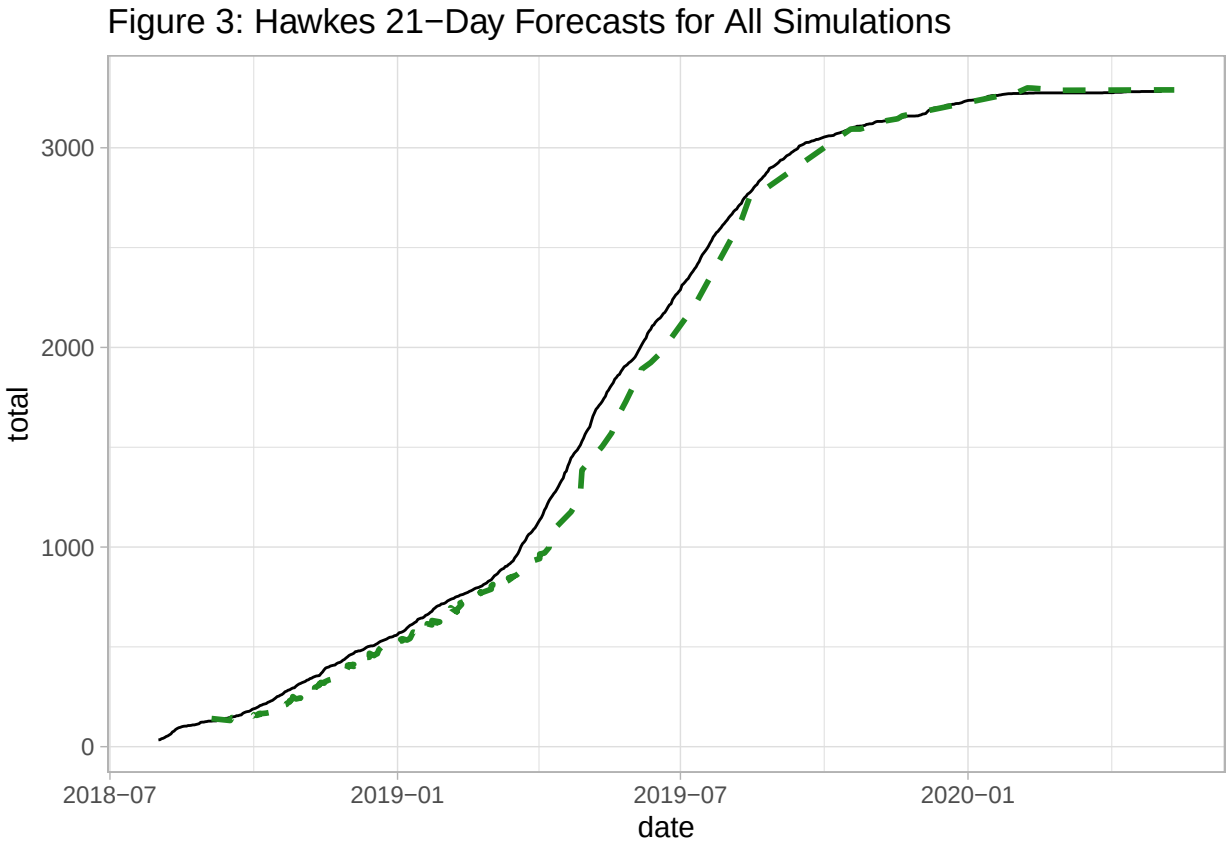
14-Day Forecast Analysis

Figure 2 below shows the Hawkes 14-Day Forecasts for all simulations with respect to the recorded number of infections.



21-Day Forecast Analysis

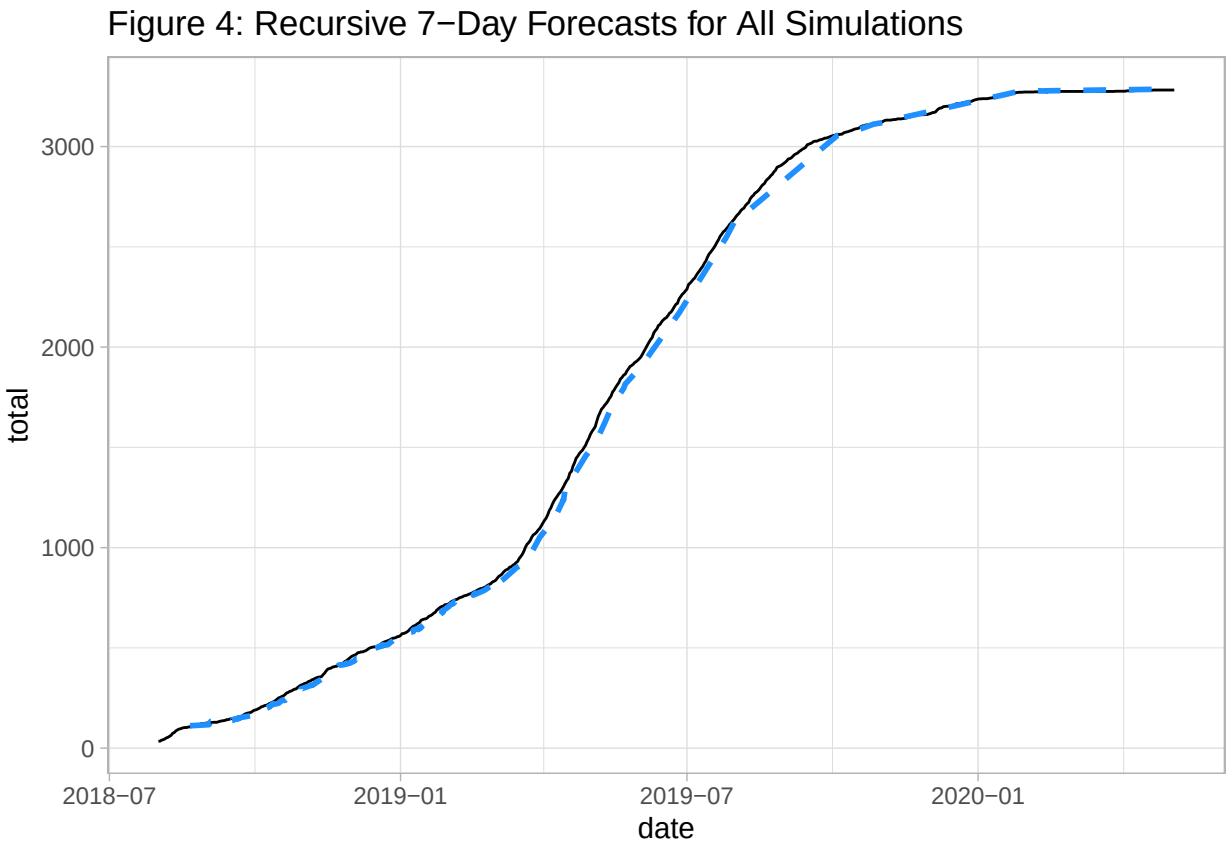
Figure 3 below shows the Hawkes 21-Day Forecasts for all simulations with respect to the recorded number of infections.



Recursive Complete Outbreak Analysis

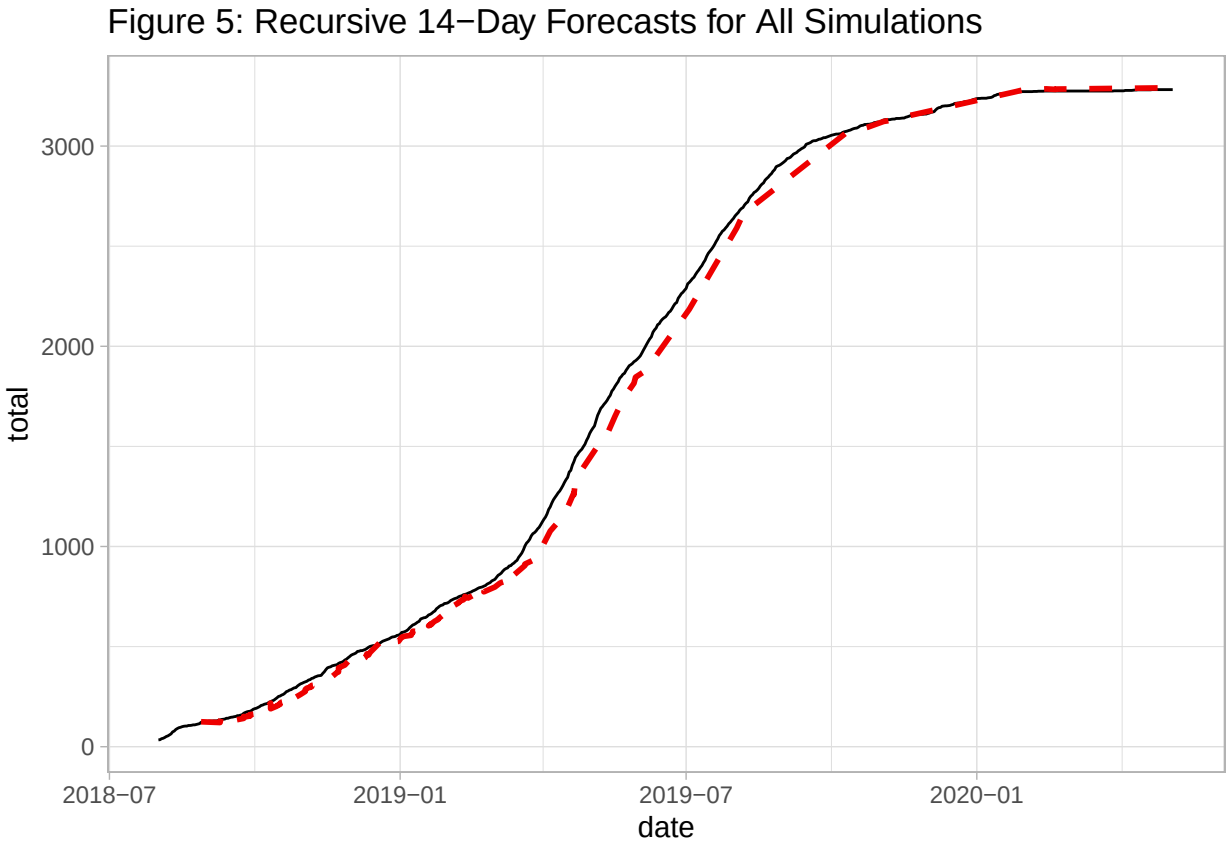
7-Day Forecast Analysis

Figure 4 below shows the Recursive 7-Day Forecasts for all simulations with respect to the recorded number of infections.



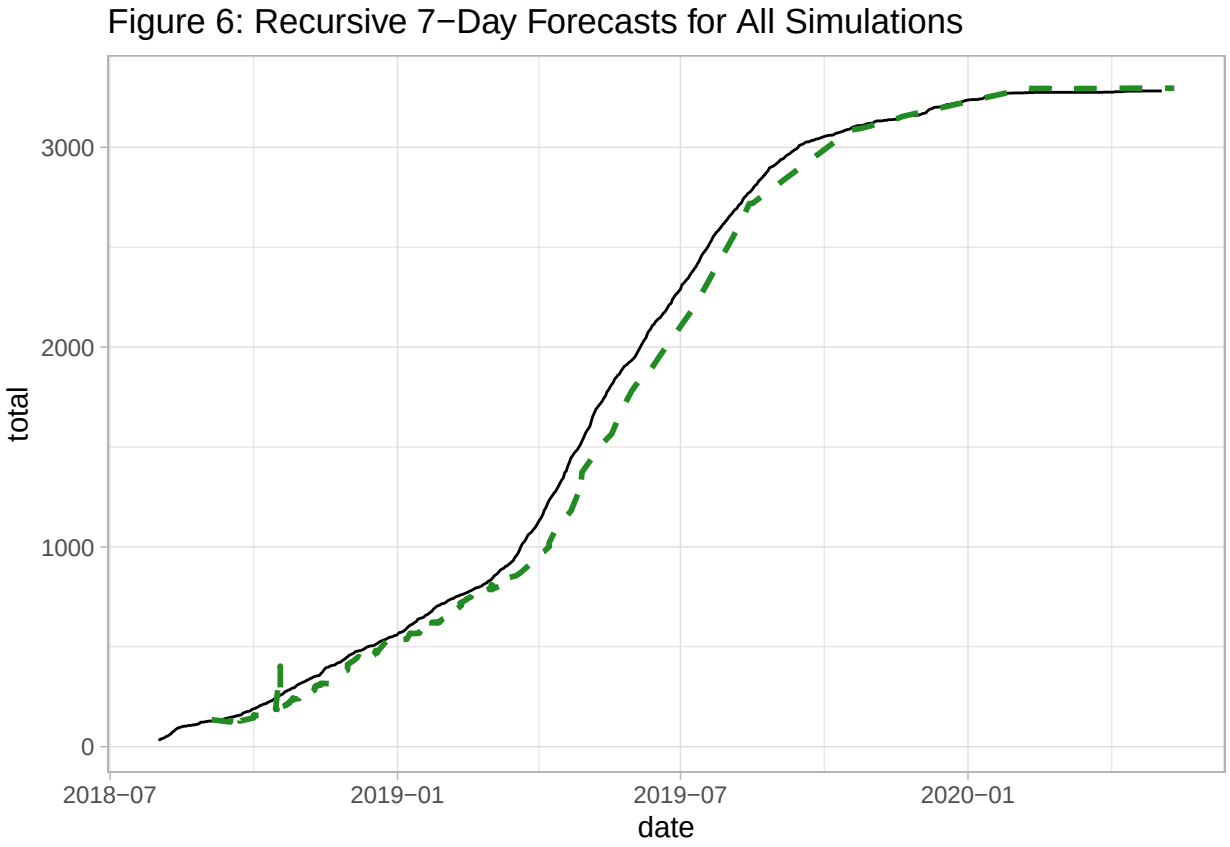
14-Day Forecast Analysis

Figure 5 below shows the Recursive 14-Day Forecasts for all simulations with respect to the recorded number of infections.



21-Day Forecast Analysis

Figure 6 below shows the Recursive 21-Day Forecasts for all simulations with respect to the recorded number of infections.



RMSE for Full Hawkes and Recursive Forecasts

We compute the Root Mean Square Error (RMSE) of the 7, 14, and 21-day forecasts for both the Hawkes and Recursive models. The RMSEcan be computed as

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)}{N}}$$

where N is the total number of observations.

The table below (Table 1) shows the RMSE values for the Hawkes and Recursive models, with respect to every simulated forecast during the outbreak.

Table 1: RMSE values for Hawkes and Recursive Models for all simulations.

	Hawkes	Recursive
7-day	27.90	29.67
14-day	58.68	60.70
21-day	88.44	92.22

We see that the Hawkes model forecasts have a consistently lower RMSE than those of the Recursive model for all three prediction days, when looking at all simulations.

Hawkes Partial Outbreak Analysis

After analyzing the entire model, we then refine our analysis to a select number of forecasts.