Ebola Forecasting - Error Analysis

Andy Shen

8/15/2020

Data Input and Cleaning

Outbreak Dataset

```
rm(list=ls())
true <- read.csv("/Volumes/GoogleDrive/.shortcut-targets-by-id/15UGkfREtfqH3LdfHmCsSpFJ5SrTnSeyt/ebola/1
source("outbreak_vis.R") #script with functions
```

```
## -- Attaching packages ---------------------------------------------------------------------------
```

```
## v ggplot2 3.3.2     v purrr   0.3.4
## v tibble  3.0.1     v dplyr   1.0.0
## v tidyr   1.1.0     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.5.0
```

```
## -- Conflicts ------------------------------------------------------------------------------------
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
##
## Attaching package: 'lubridate'
```

```
## The following objects are masked from 'package:base':
##
##     date, intersect, setdiff, union
```

```
rgx <- "\\d{1,2}\\/\\d{1,2}\\/\\d{4}" #date structure regex
true <- true[str_detect(true$Date, rgx),] #omits rows without a date
colnames(true) <- c("date", "cases")
true$date <- mdy(true$date)
true$cases[is.na(true$cases)] <- 0
true <- true %>% mutate(total = cumsum(cases))
last_date <- true$date[length(true$date)]
last_case <- true$total[length(true$total)]
```

Hawkes Projections Dataset

```
hproj <- read.csv("/Volumes/GoogleDrive/.shortcut-targets-by-id/1LaD1nL_OAOposW2fr2XDcs6BLHVBC-jA/2019 N
hproj <- hproj %>% select(date_last_case, pred.7, pred.14, pred.21)
hproj$date_last_case <- mdy(hproj$date_last_case)
dates <- as.character(hproj$date_last_case)
preds <- t(hproj %>% select(pred.7,pred.14,pred.21))
```

Hawkes Complete Outbreak Analysis

7-Day Forecast Analysis

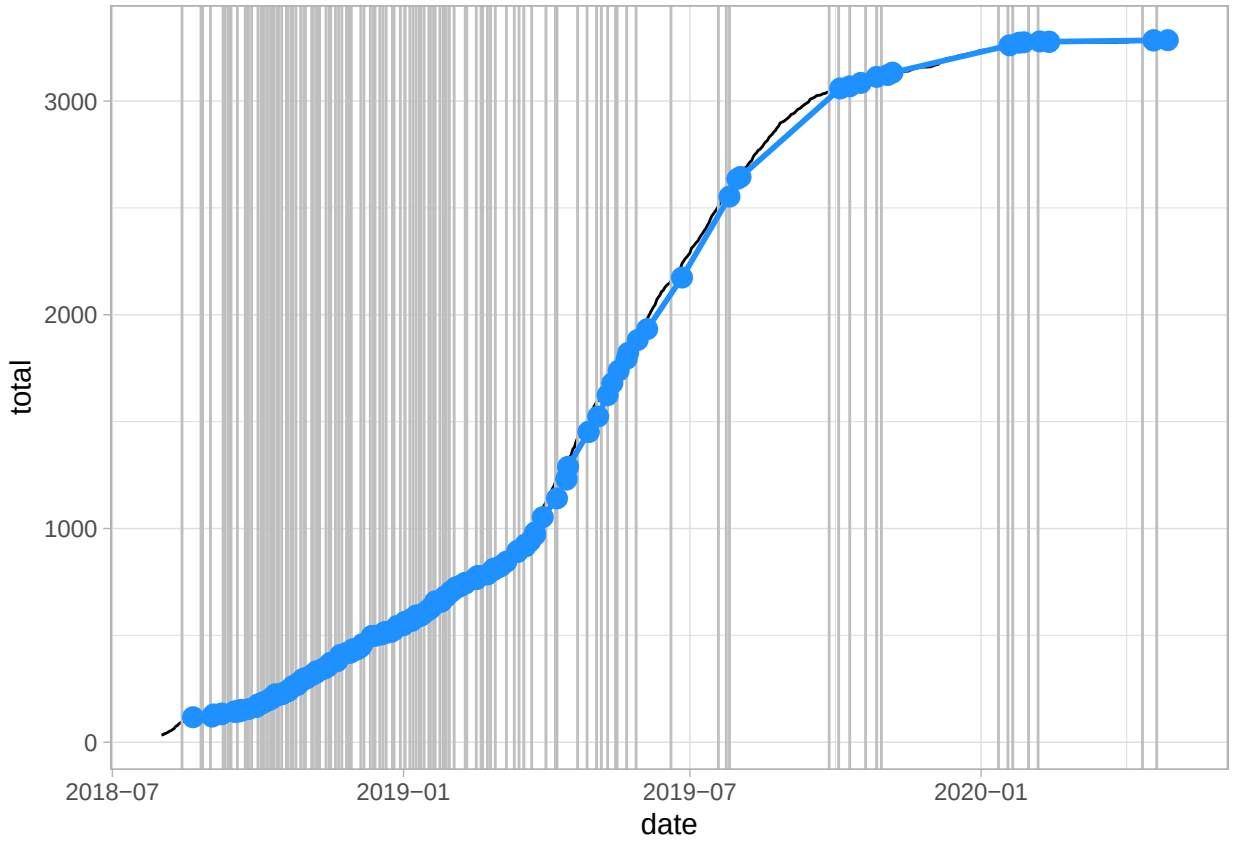
```
mod <- single_forecast(dates, preds, days = 7, res = TRUE)
head(mod$results, 10)
```

```
##    prior.date prior.total forecast.date actual.total forecast.total resids
## 1  2018-08-14          95    2018-08-21          106            117    -11
## 2  2018-08-26         113    2018-09-02          128            120      8
## 3  2018-08-27         115    2018-09-03          128            130     -2
## 4  2018-09-01         126    2018-09-08          132            131      1
## 5  2018-09-01         126    2018-09-08          132            133     -1
## 6  2018-09-01         126    2018-09-08          132            134     -2
## 7  2018-09-10         136    2018-09-17          148            143      5
## 8  2018-09-09         134    2018-09-16          147            143      4
## 9  2018-09-10         136    2018-09-17          148            143      5
## 10 2018-09-10         136    2018-09-17          148            144      4
```

```
mod$rmse
```

```
##       RMSE
## 1 27.83955
```

```
mod$plot
```



14-Day Forecast Analysis

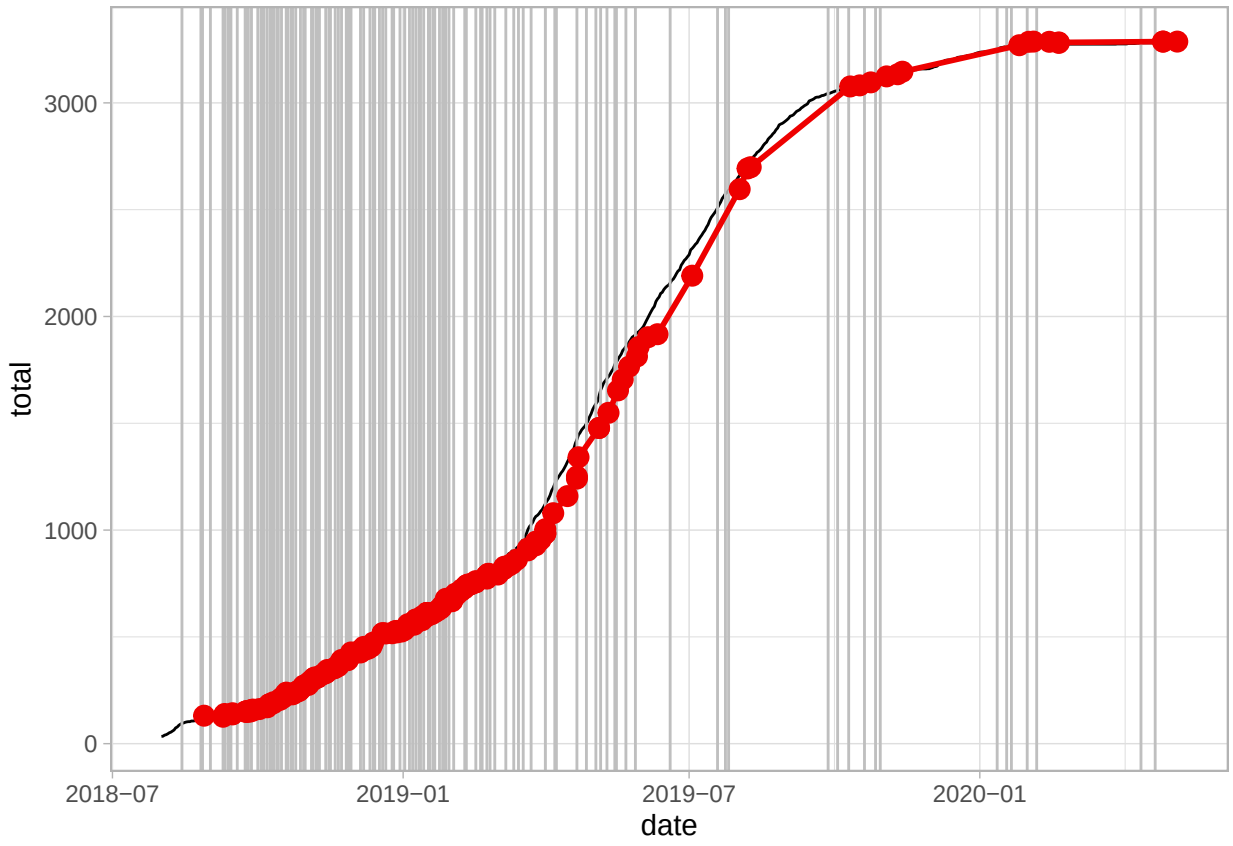
```
mod <- single_forecast(dates, preds, days = 14, res = TRUE)
results <- mod$results
head(mod$results, 10)
```

```
##    prior.date prior.total forecast.date actual.total forecast.total resids
## 1  2018-08-14          95    2018-08-28          122          131.0   -9.0
## 2  2018-08-26         113    2018-09-09          134          126.0    8.0
## 3  2018-08-27         115    2018-09-10          136          139.0   -3.0
## 4  2018-09-01         126    2018-09-15          145          137.0    8.0
## 5  2018-09-01         126    2018-09-15          145          139.0    6.0
## 6  2018-09-01         126    2018-09-15          145          142.5    2.5
## 7  2018-09-10         136    2018-09-24          167          149.0   18.0
## 8  2018-09-09         134    2018-09-23          159          150.0    9.0
## 9  2018-09-10         136    2018-09-24          167          150.0   17.0
## 10 2018-09-10         136    2018-09-24          167          152.0   15.0
```

```
mod$rmse
```

```
##       RMSE
## 1 58.68716
```

```
mod$plot
```



21-Day Forecast Analysis

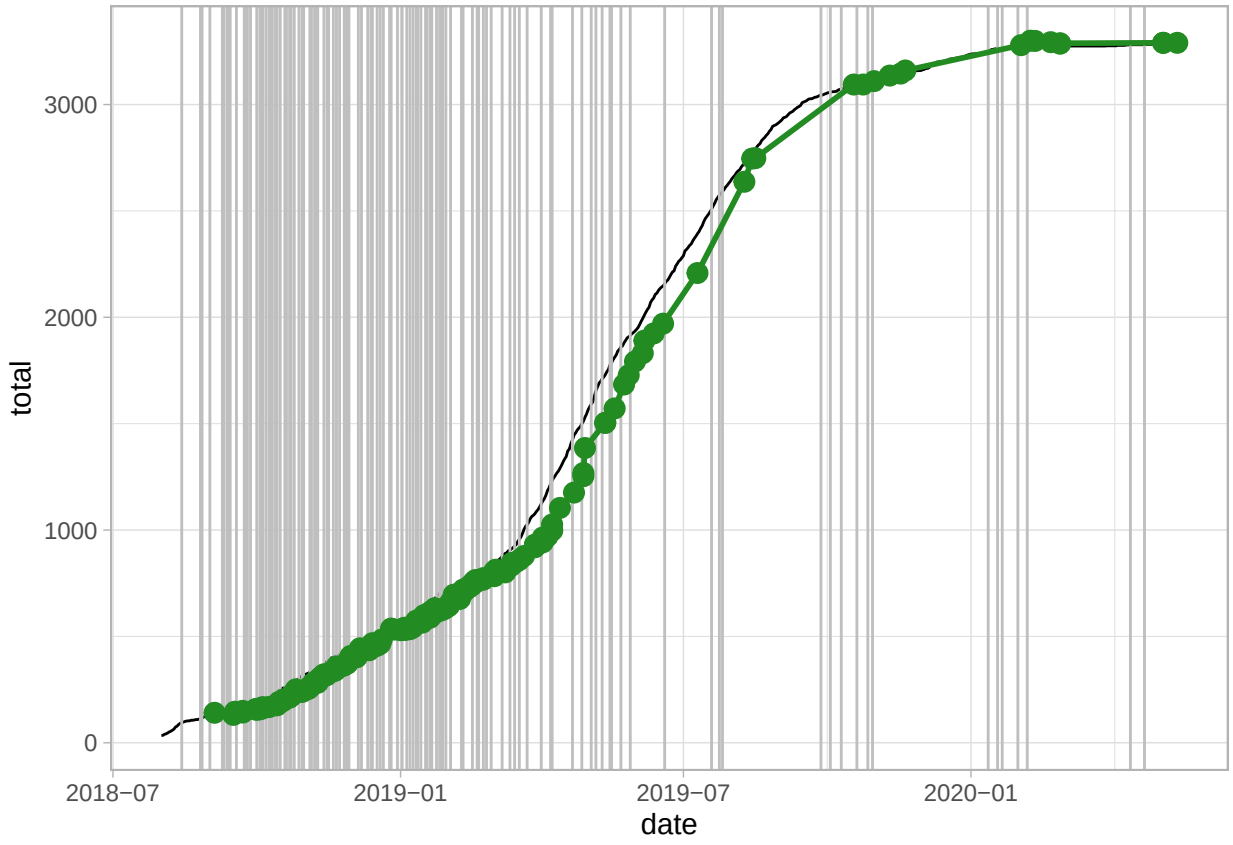
```
mod <- single_forecast(dates, preds, days = 21, res = TRUE)
head(mod$results, 10)
```

```
##    prior.date prior.total forecast.date actual.total forecast.total resids
## 1  2018-08-14          95    2018-09-04          128            141    -13
## 2  2018-08-26         113    2018-09-16          147            131     16
## 3  2018-08-27         115    2018-09-17          148            146      2
## 4  2018-09-01         126    2018-09-22          158            142     16
## 5  2018-09-01         126    2018-09-22          158            146     12
## 6  2018-09-01         126    2018-09-22          158            150      8
## 7  2018-09-10         136    2018-10-01          190            156     34
## 8  2018-09-09         134    2018-09-30          187            157     30
## 9  2018-09-10         136    2018-10-01          190            156     34
## 10 2018-09-10         136    2018-10-01          190            159     31
```

```
mod$rmse
```

```
##       RMSE
## 1 88.41671
```

```
mod$plot
```



Hawkes Partial Outbreak Analysis

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