IHME calibration

Jouni Tuomisto

4/18/2020

IHME calibration

This document used IHME predictions about COVID-19 epidemic and hospital capacity to calibrate these estimates.

```
Sys.setlocale("LC_TIME", "C")
```

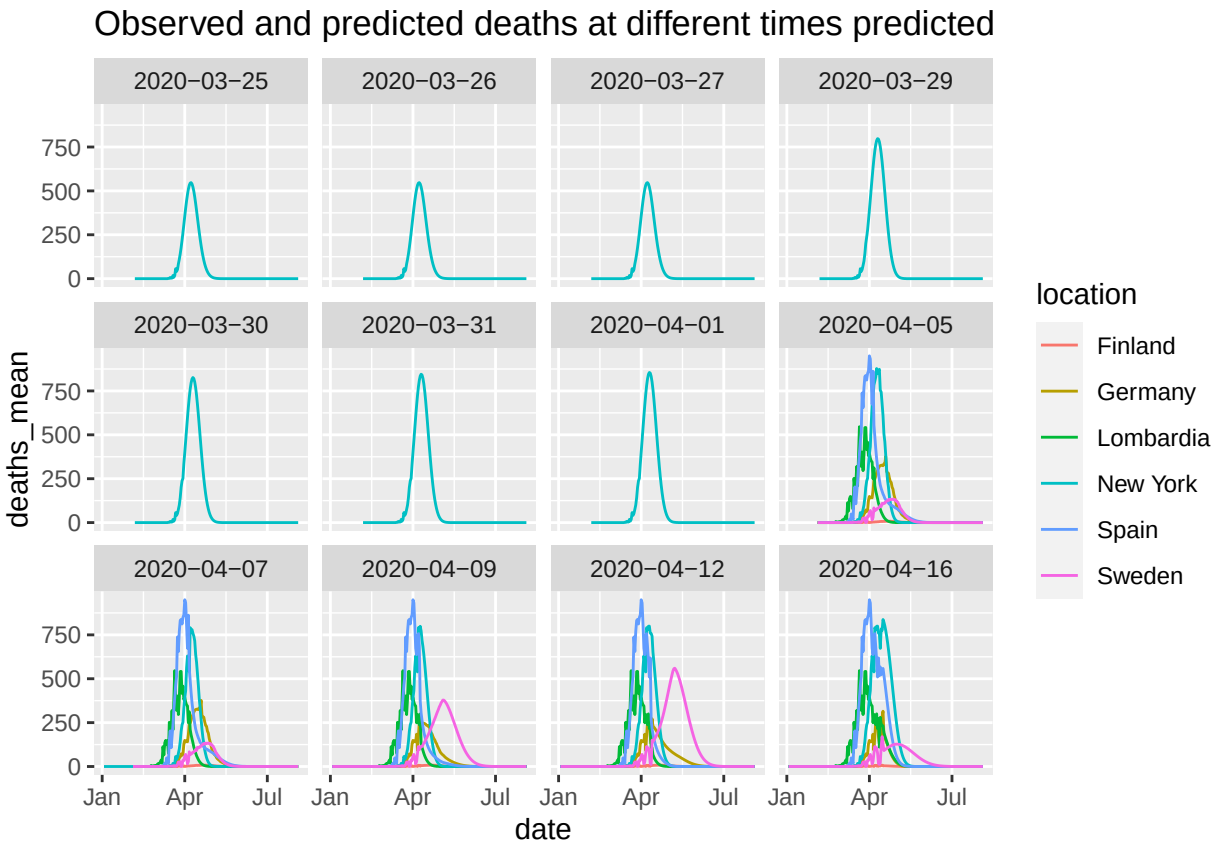
```
## [1] "C"
```

```
data_date <- c("2020-03-25","2020_03_26","2020_03_27","2020_03_29","2020_03_30",
               "2020_03_31.1","2020_04_01.2",#"2020_04_05.05.us",
               "2020_04_05.08.all", "2020_04_07.06.all","2020_04_09.04","2020_04_12.02",
               "2020_04_16.05")
filename <- c("ihme-covid19_all_locs", rep("hospitalization_all_locs_corrected",3),
  rep("Hospitalization_all_locs",8))

dat <- data.frame()
for(i in 1:length(data_date)) {
  tmp <- read.csv(paste0("~/Downloads/ihme-covid19/", data_date[i], "/", filename[i],".csv"))
  tmp$V1 <- NULL
  tmp$X <- NULL
  colnames(tmp)[colnames(tmp)=="date_reported"] <- "date"
  if(all(c("location", "location_name") %in% colnames(tmp))) tmp$location_name <- NULL
  colnames(tmp)[colnames(tmp)=="location_name"] <- "location"
  tmp$date <- as.Date(tmp$date)
  tmp <- tmp[tmp$location %in% c("Finland","New York","Lombardia","Germany",
                                 "Sweden","Spain"),]
  dat <- rbind(cbind(
    dataset=as.Date(gsub("_","-",substr(data_date[i],1,10))),tmp),
    dat)
}

comparison <- merge(
  dat[dat$date >= "2020-04-05" & dat$date <= "2020-04-16" & dat$dataset < dat$date,],
  dat[dat$dataset=="2020-04-16",colnames(dat)!="dataset"],
  by=c("date","location")) # x is prediction, y is observation
comparison <- comparison[c("date","location","dataset","deaths_mean.y","deaths_mean.x","deaths_lower.x"
comparison$deaths_prediction <- (comparison$deaths_mean.y - comparison$deaths_lower.x) / (comparison$dea
```

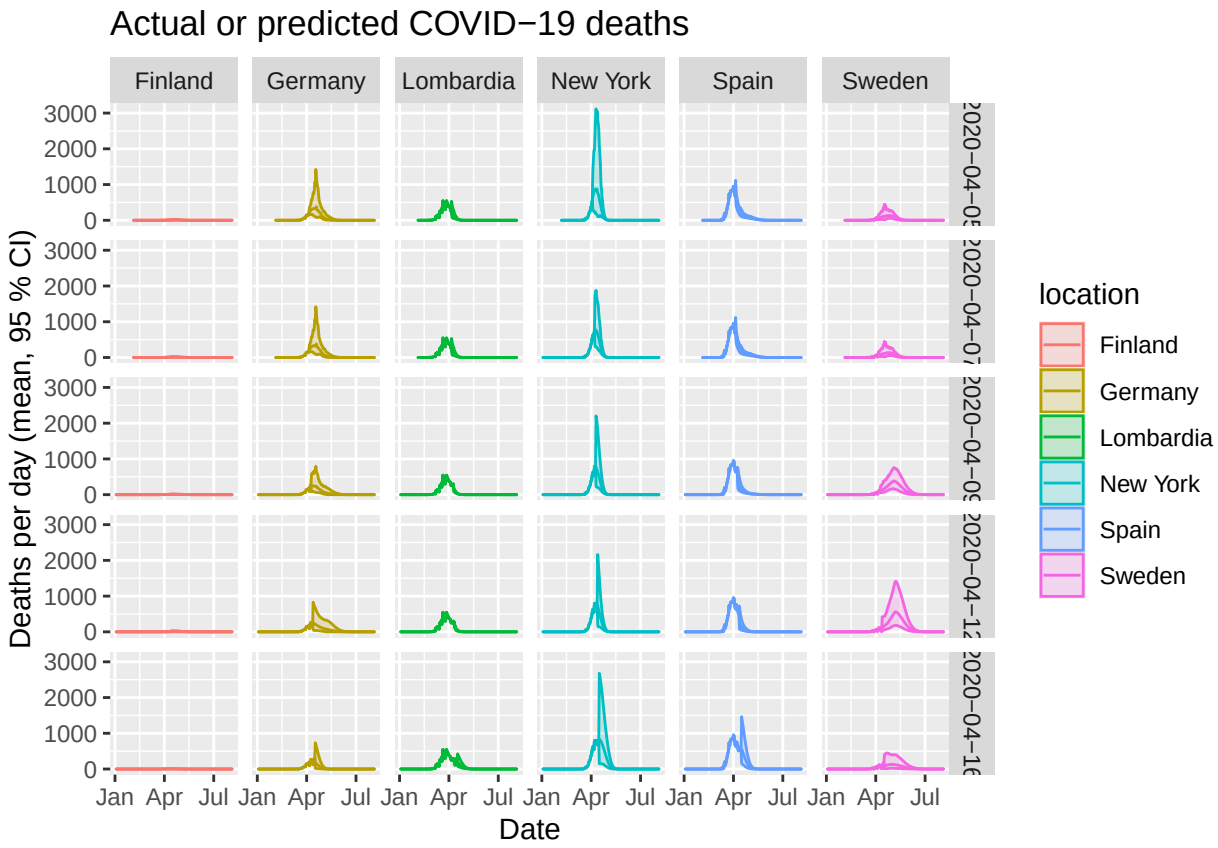
Including Plots

```
library(ggplot2)
```

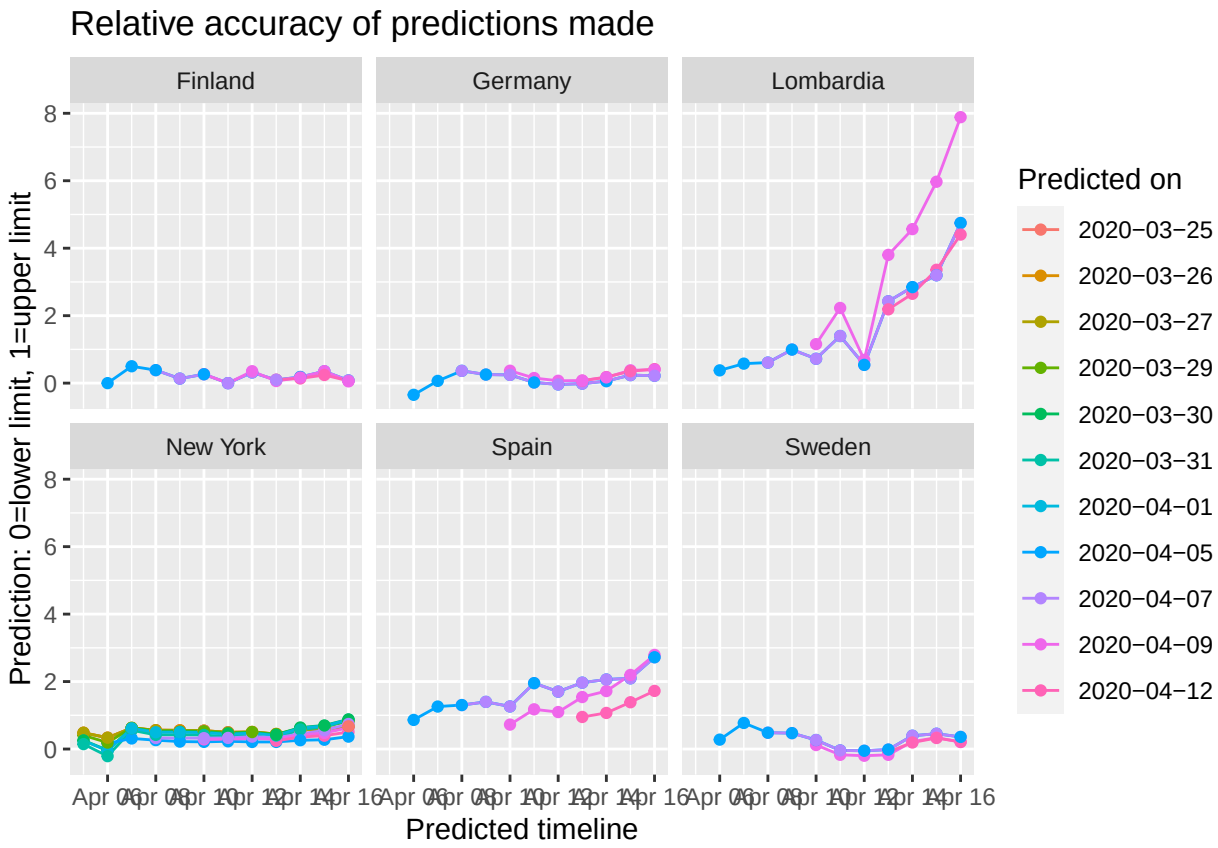
```
ggplot(dat, aes(x=date, y=deaths_mean, colour=location))+geom_line()+
  facet_wrap(~dataset)+
  labs(title="Observed and predicted deaths at different times predicted")
```




```
ggplot(dat[dat$dataset>="2020-04-05",], aes(x=date, y=deaths_mean, colour=location))+geom_line()+
  geom_ribbon(aes(ymin=deaths_lower, ymax=deaths_upper, fill=location), alpha=0.2)+
  facet_grid(dataset~location)+
  labs(
    title="Actual or predicted COVID-19 deaths",
    y="Deaths per day (mean, 95 % CI)",
    x="Date"
  )
```




```
#tmp <- comparison
#tmp$deaths_prediction[tmp$date > tmp$dataset] <- NA
ggplot(comparison, aes(x=date, y=deaths_prediction, colour=as.character(dataset)))+geom_line()+geom_poi
  facet_wrap(~location)+
  labs(
    title="Relative accuracy of predictions made",
    x="Predicted timeline",
    y="Prediction: 0=lower limit, 1=upper limit"
  )+
  guides(colour=guide_legend(title="Predicted on"))
```


Relative accuracy of predictions made
Finland
Germany
Lombardia
New York
Spain
Sweden
Prediction: 0=lower limit, 1=upper limit
Predicted timeline
Predicted on
2020−03−25
2020−03−26
2020−03−27
2020−03−29
2020−03−30
2020−03−31
2020−04−01
2020−04−05
2020−04−07
2020−04−09
2020−04−12