COVID-19 Forecast Similarity Analysis

Nutcha Wattanachit, Johannes Bracher, Evan Ray, Nick Reich

10/10/2022

Concept: Examples of forecasts and Cramer distances - CA

The plots show the distances between the ensemble and UMass-MechBayes at two different forecast dates. We can see that lower approximated CD values indicates more similarity between the two forecasts - as the plots of the predictive quantiles on forecast date 2021-04-05 shows a lot of overlap. On the other hand, the plots of the predictive quantiles on forecast date 2021-08-02 shows less overlap accompanying by higher approximated CD values.

Model Type - Concept

Overall Similarities

- Targets: all wk ahead inc deaths
- End dates: May 02 2020-Dec 18 2021
- Model: The same set as in the evaluation paper (28 models) that have 4 horizons on each submission and that have more than 10000k submissions
- Probability levels: All
- Locations: National and states (excluding US territory and DC), so 51 locations

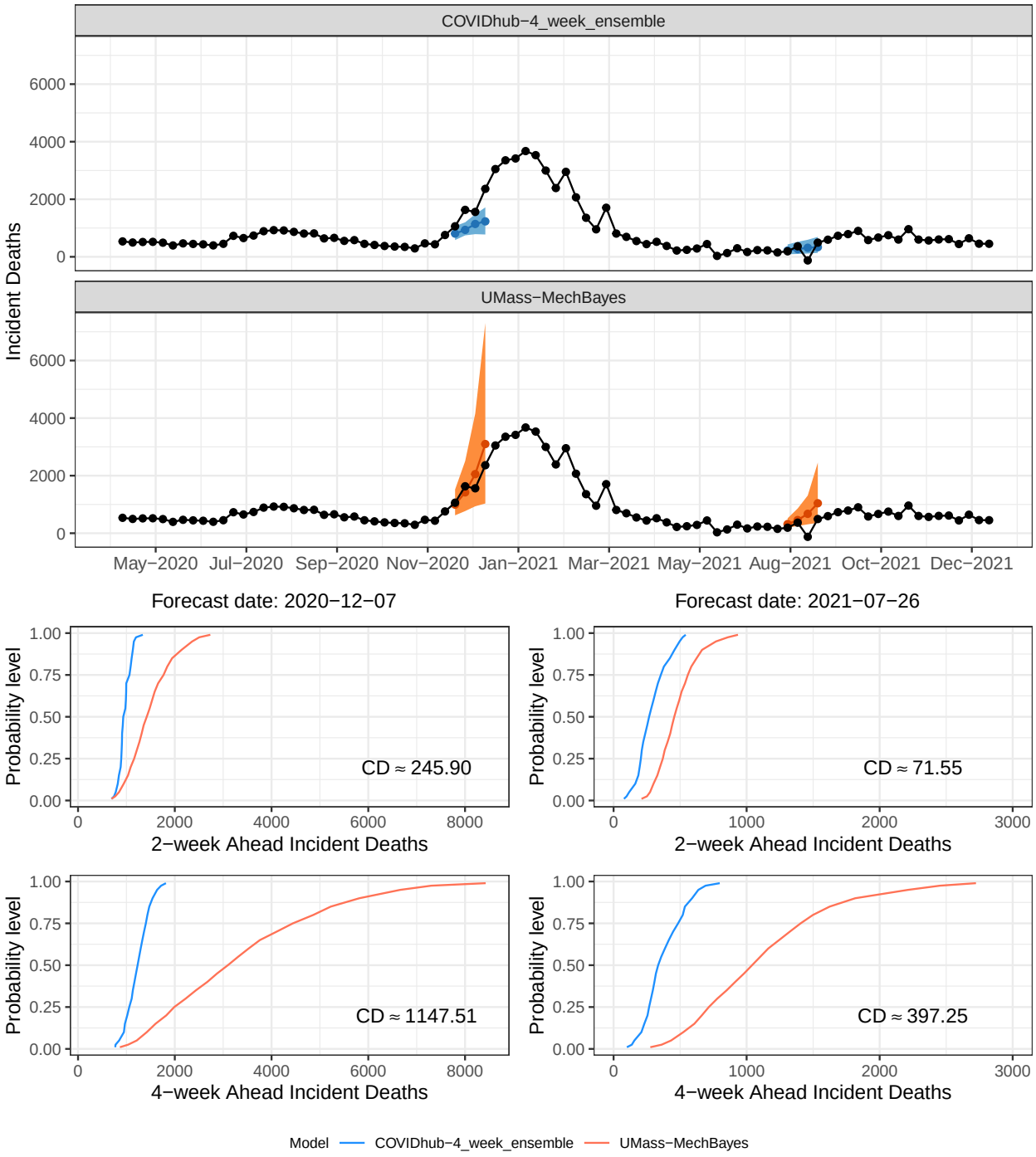


Figure 1: COVID-19 deaths forecasts (left) and all predictive quantiles (right)

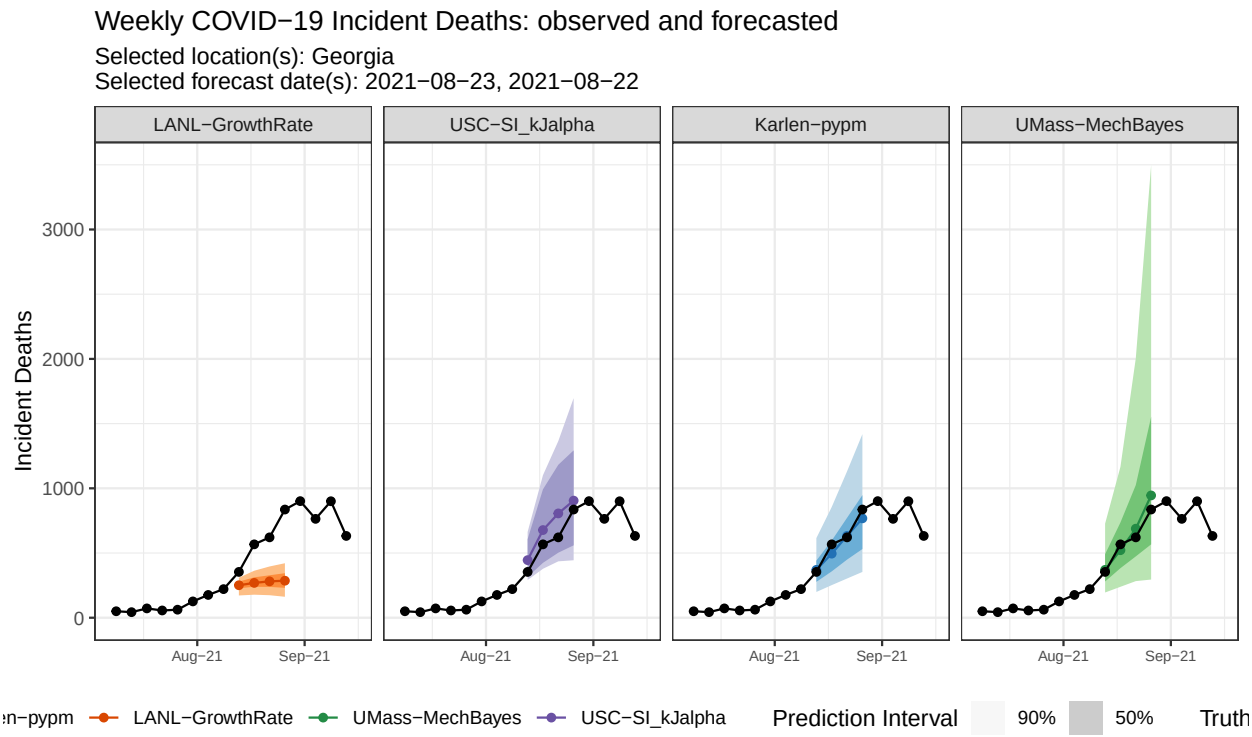


Figure 2: COVID-19 deaths forecasts (left) and all predictive quantiles (right)

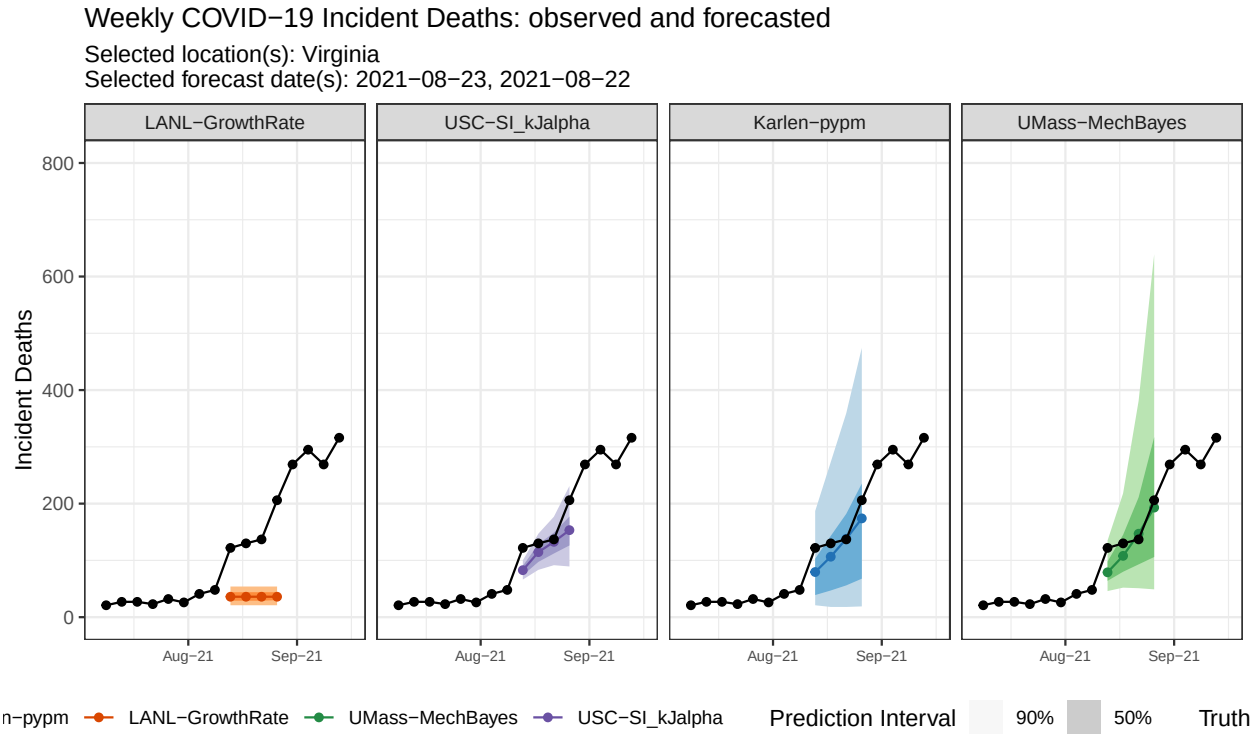


Figure 3: COVID-19 deaths forecasts (left) and all predictive quantiles (right)

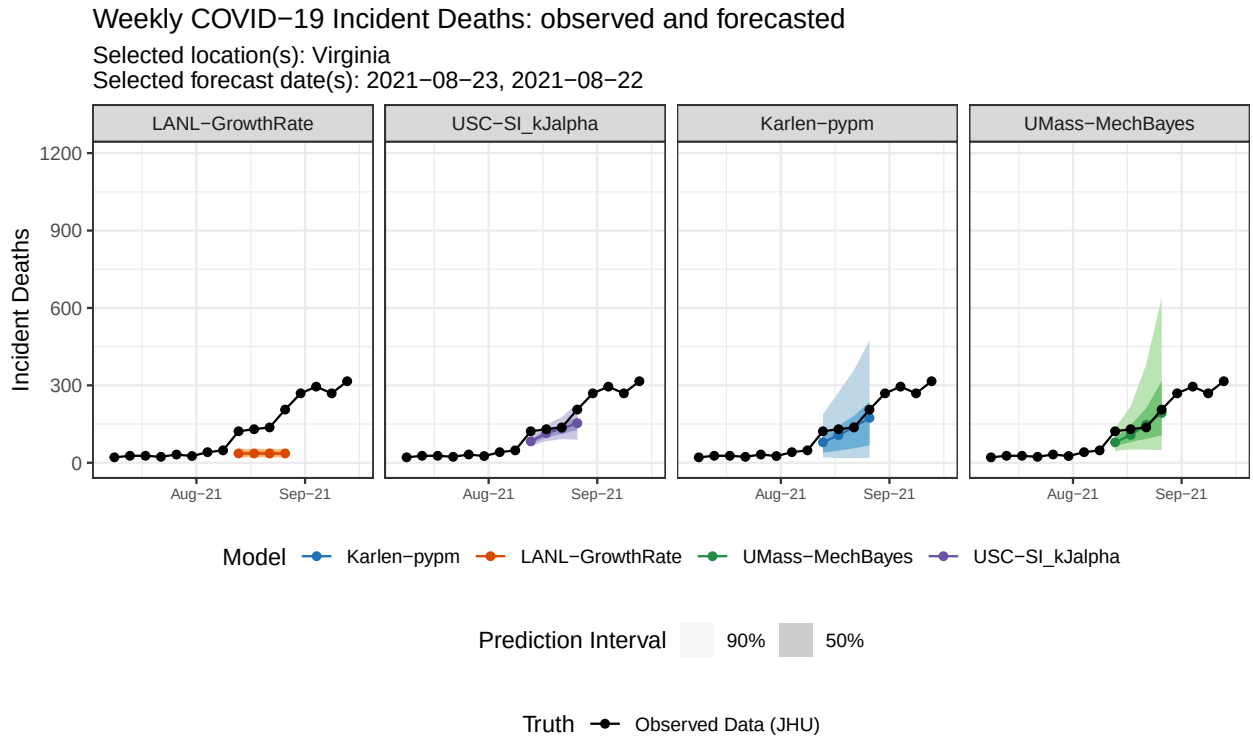
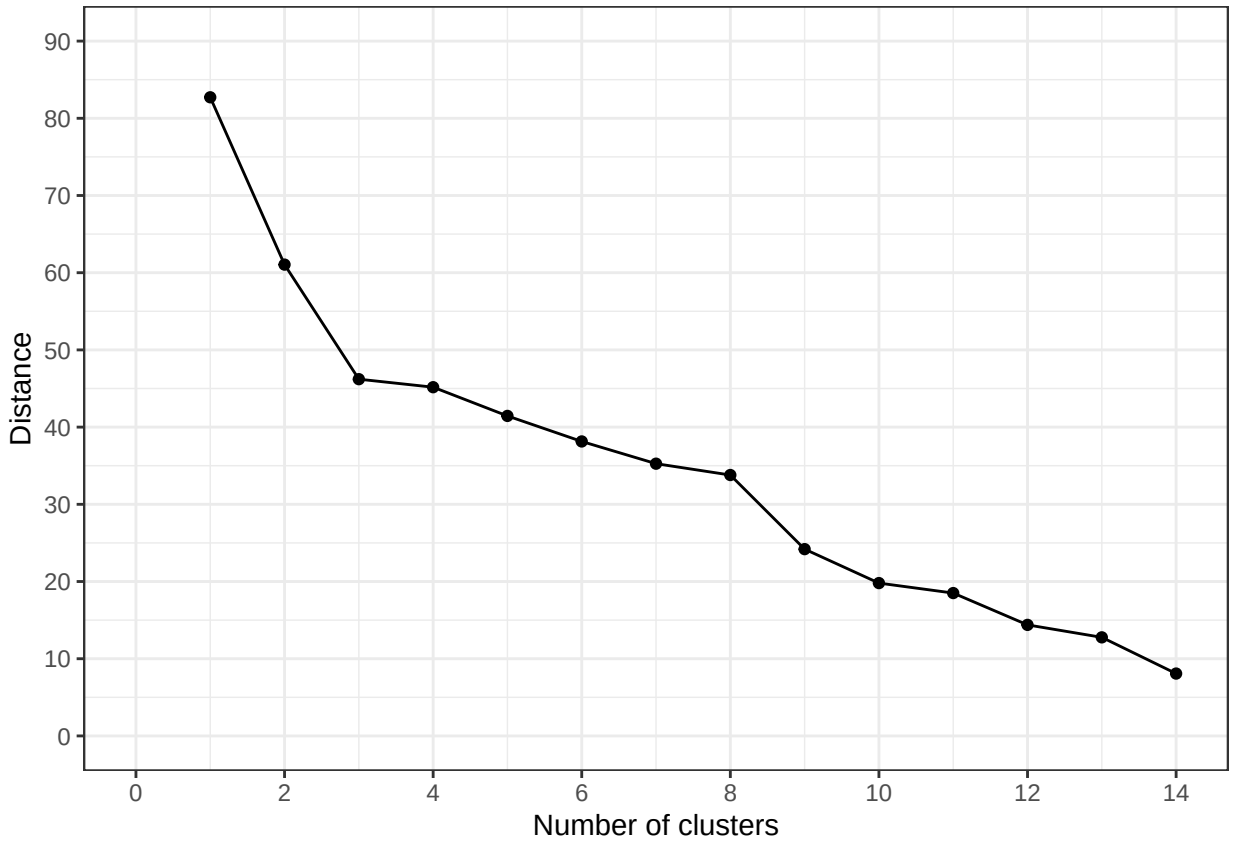
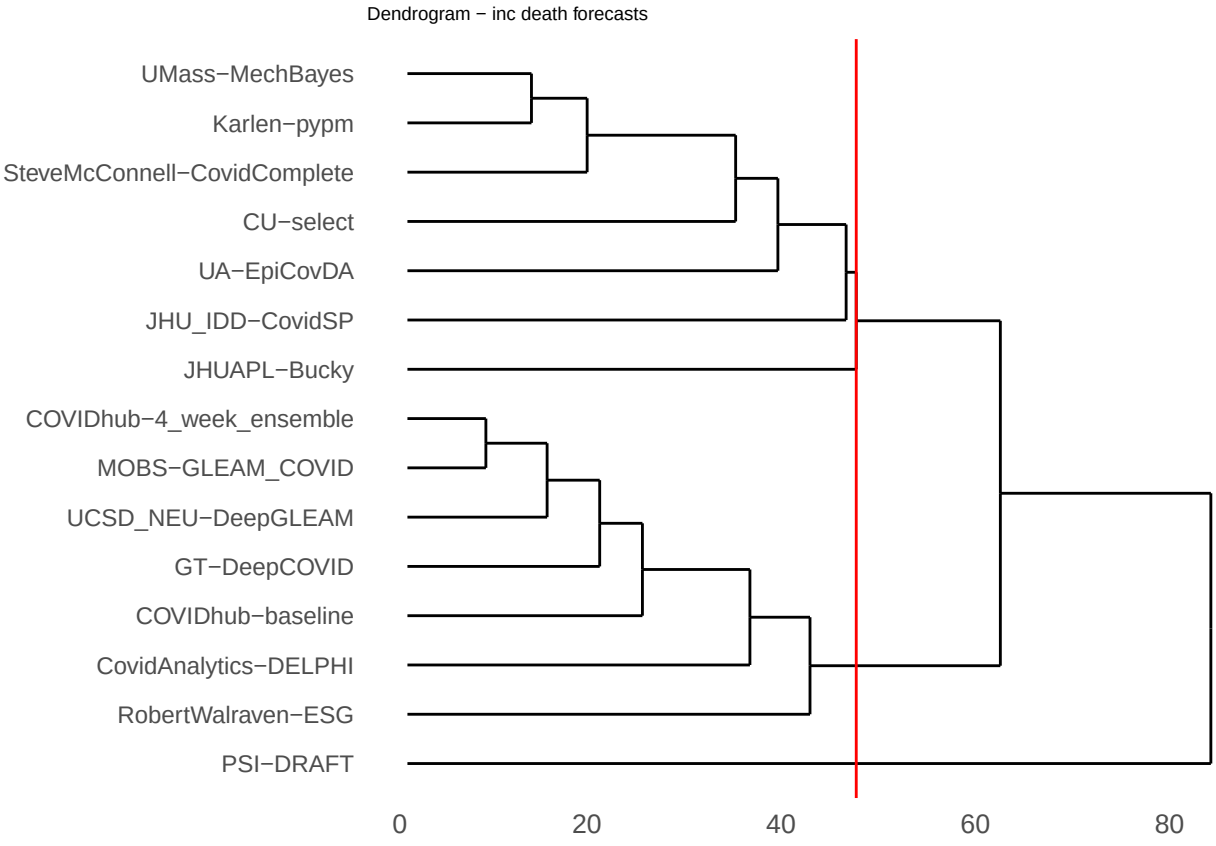
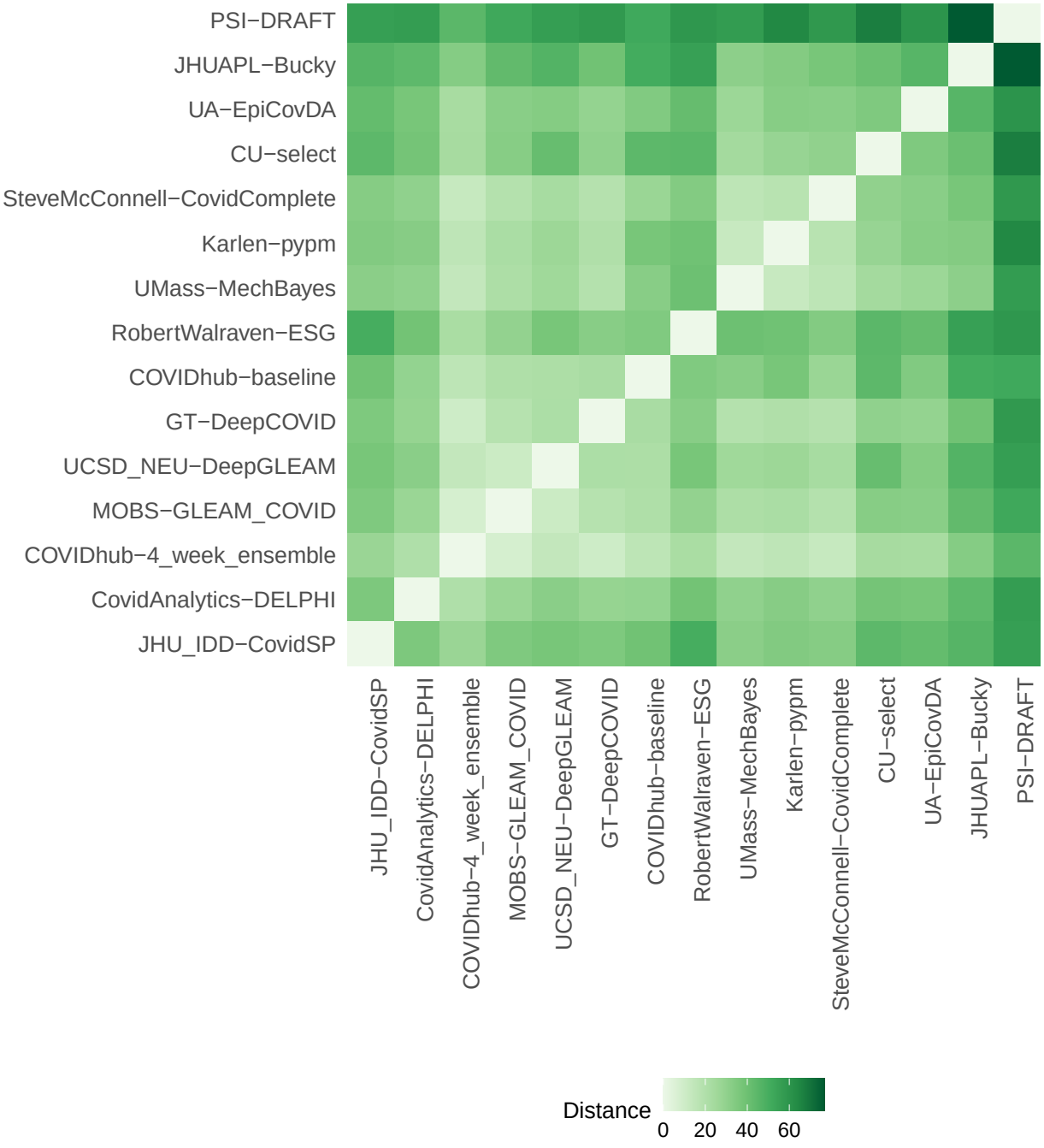


Figure 4: COVID-19 deaths forecasts (left) and all predictive quantiles (right)




Dendrogram − inc death forecasts
UMass−MechBayes
Karlen−pypm
SteveMcConnell−CovidComplete
CU−select
UA−EpiCovDA
JHU_IDD−CovidSP
JHUAPL−Bucky
COVIDhub−4_week_ensemble
MOBS−GLEAM_COVID
UCSD_NEU−DeepGLEAM
GT−DeepCOVID
COVIDhub−baseline
CovidAnalytics−DELPHI
RobertWalraven−ESG
PSI−DRAFT
0
20
40
60
80



Distance/dissimilarity as a signal/leading indicator of an increase in incident deaths

Cross-correlations of median approx. CDs and incident deaths are calculated for 2 "up" periods, the winter 2020 and the fall 2021, are calculated for 34 selected states based on the criteria below. The maximum lags and leads are 9 weeks before and after each target dates during 9 weeks of up periods.

Forecast inclusion criteria

- Targets: 2 and 4 wk ahead inc death
- Target End Dates: Jan 2021-Jan 2022
- Model: all models submitted on a particular date
- Probability levels: All
- Locations: All 34 states (continental US) except HI, AK and DC. We also exclude DE, ME,MT,OK,OR,VT for low or highly variable incidents. We also exclude AZ, CA, CT, MD, MO, NE, NJ, RI since they have only one peak.

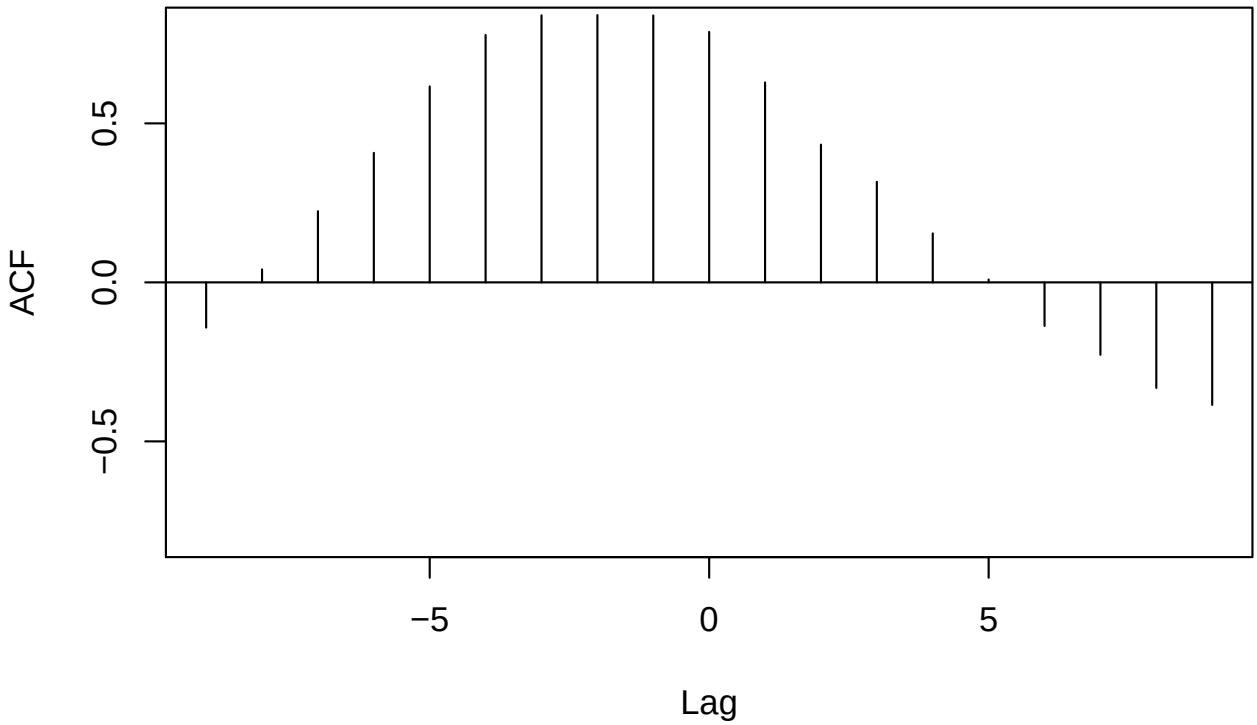
Cross correlations

The distributions of cross-correlations across the two up periods of winter 2020/2021 and fall 2021 waves (and 9 weeks before the first week and 9 after the last week) show that there is no evidence that the median CDs of 2 week ahead forecasts consistently lead or lag the increase in incident deaths, however there is some evidence that the median CDs of 4 week ahead forecasts lead the increase in incident deaths.

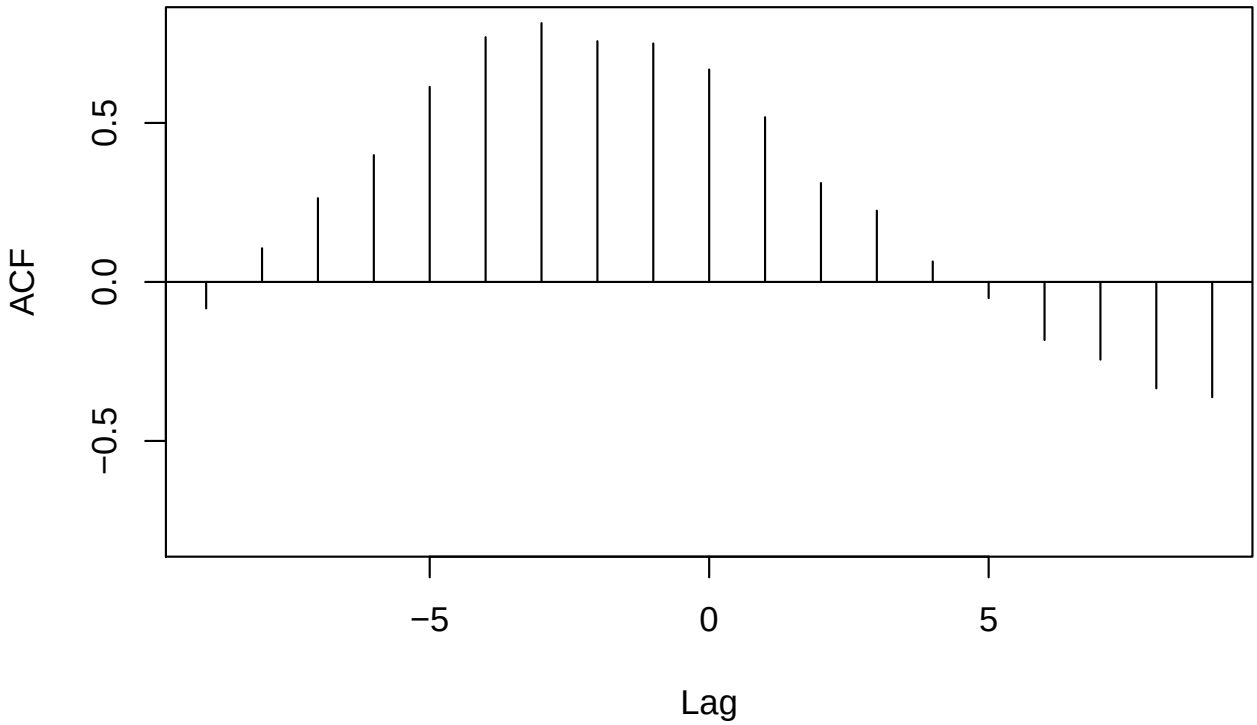
```
FALSE # A tibble: 6 x 4
FALSE # Groups:   horizon, time [4]
FALSE   horizon time          pos    prop
FALSE   <chr>   <chr>         <lgl> <dbl>
FALSE 1 2       Stable period FALSE     0
FALSE 2 2       Stable period TRUE     11
FALSE 3 2       Up period     TRUE     18
FALSE 4 4       Stable period FALSE     0
FALSE 5 4       Stable period TRUE      8
FALSE 6 4       Up period     TRUE     23
```

```
FALSE # A tibble: 4 x 3
FALSE # Groups:   horizon [2]
FALSE   horizon time          `n()`
FALSE   <chr>   <chr>         <int>
FALSE 1 2       Stable period    36
FALSE 2 2       Up period        36
FALSE 3 4       Stable period    36
FALSE 4 4       Up period        36
```

2 week ahead,
Winter 2020/21



4 week ahead,
Winter 2020/21



2 week ahead,
Summer 2021

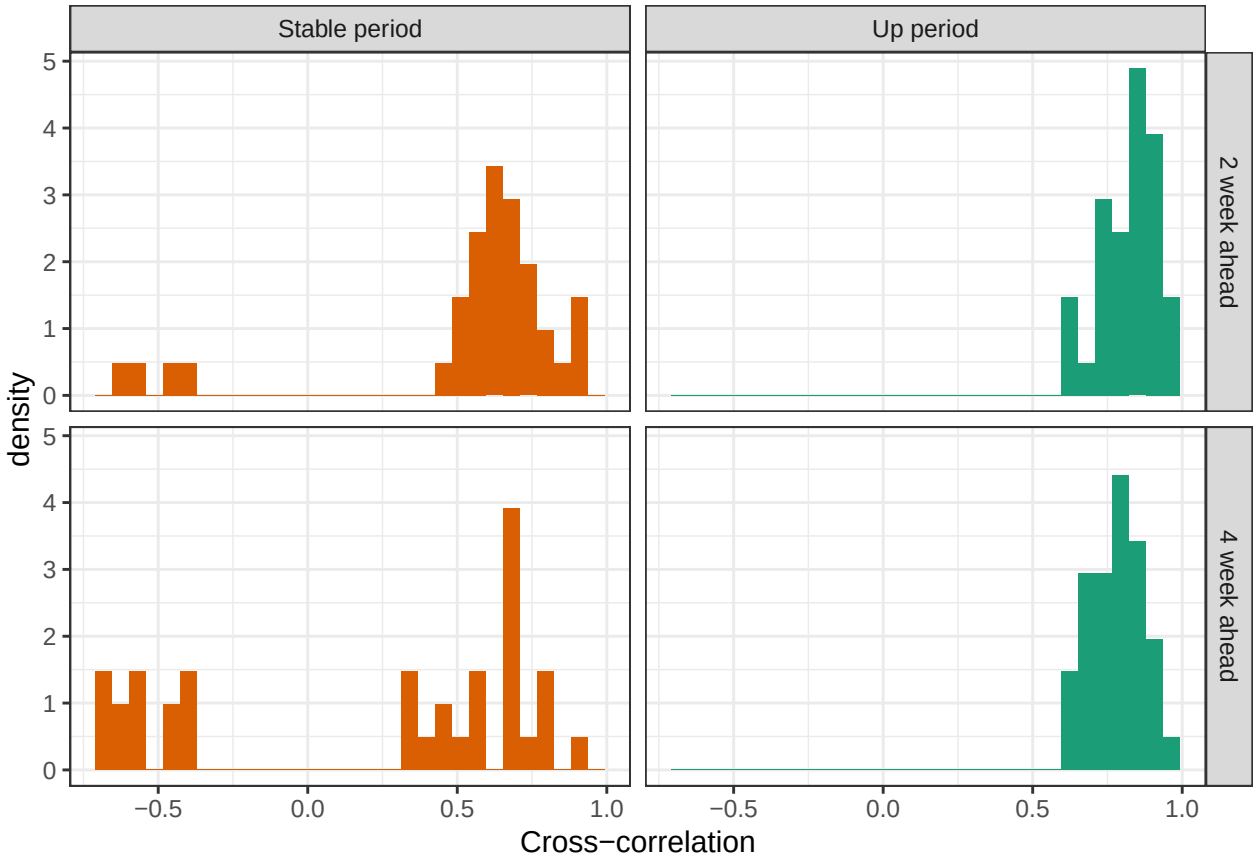


Figure 6: Distribution of max. correlations by horizon and time

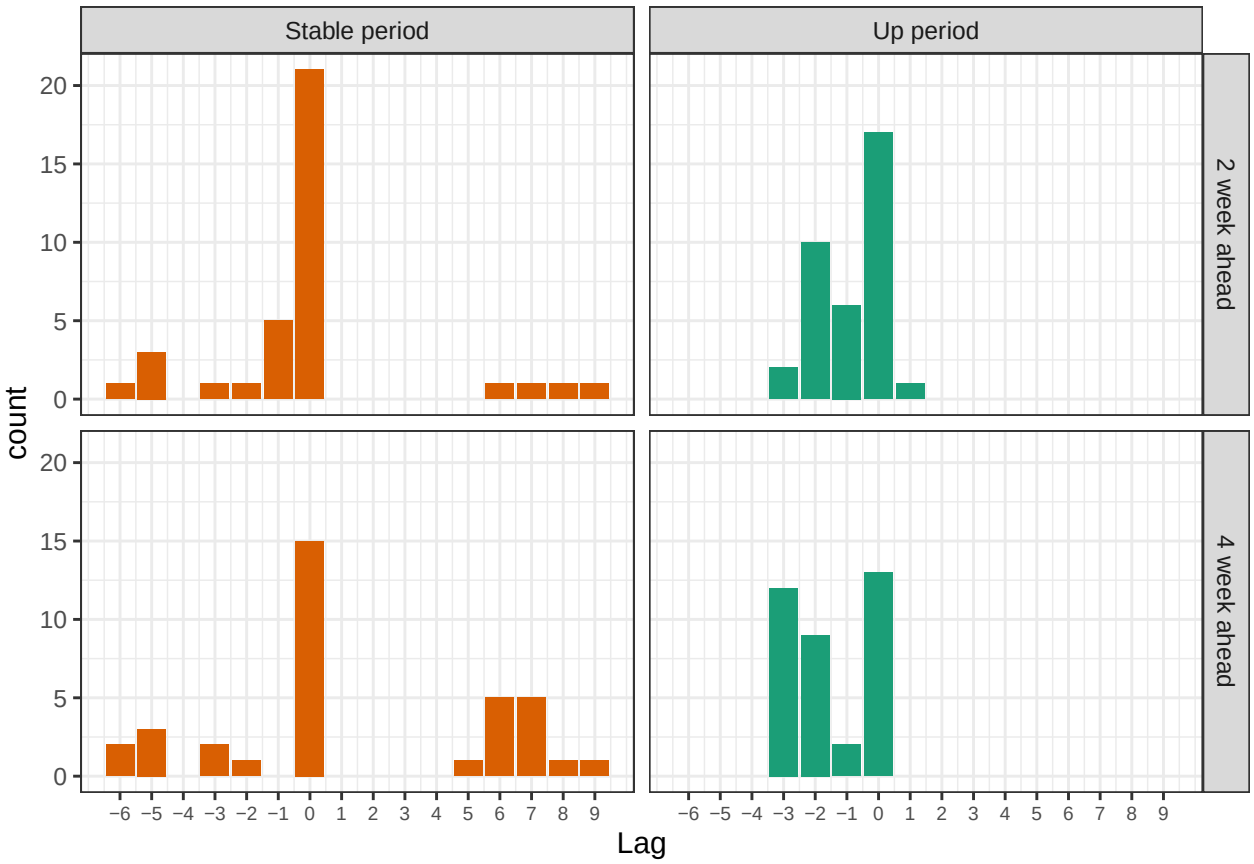


Figure 7: Count of lags of max. cross-correlations