

Forecasting Model Similarity (for inc hosp)

Johannes Bracher, Evan Ray, Nick Reich, Nutch Wattanachit, Li Shandross

06/23/2021

COVID-19 Forecasting Model Similarity Analysis for 1-4 Week Ahead Incident Hospitalization

In this extension of examining model similarity among Covid-19 models using Cramer’s distance for incident hospitalizations (inc hosp), we must make some adjustments to the inc hosp forecast data, which have a temporal resolution of “day” instead of “week,” unlike inc case and inc death. This presents a challenge because the horizons will be different for the same target end date if the forecast dates between two models differ by only a single day. (This is not an issue when the temporal resolution is in terms of weeks, which are defined by epidemiological week, not the number of days between forecast date and target end date.) Thus, we create a new variable called horizon week to solve this issue. This variable counts horizons between 1 and 7 to have a horizon week of 1, horizons between 8 and 14 to have a horizon week of 2, etc. Hence, the analyses used on inc death and inc cases should be able to be applied easily to the inc hosp data.

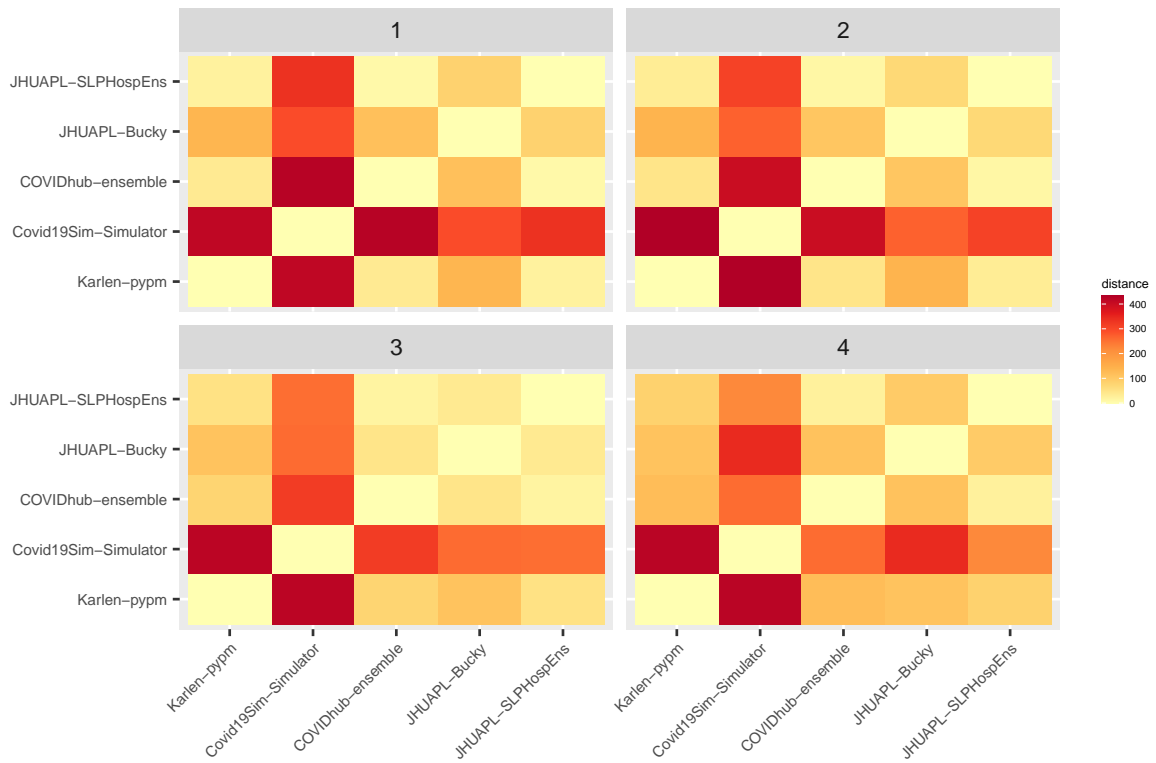
This initial analysis only examines target end dates for a single day of the week, Thursday, to account for models that include day of the week effects. However, this will be expanded later to more days, perhaps even exploring if performing the analysis on different days of the week leads to different results in terms of model similarity.

5 locations with the highest number of COVID-19 hospitalizations by the end of February 2021

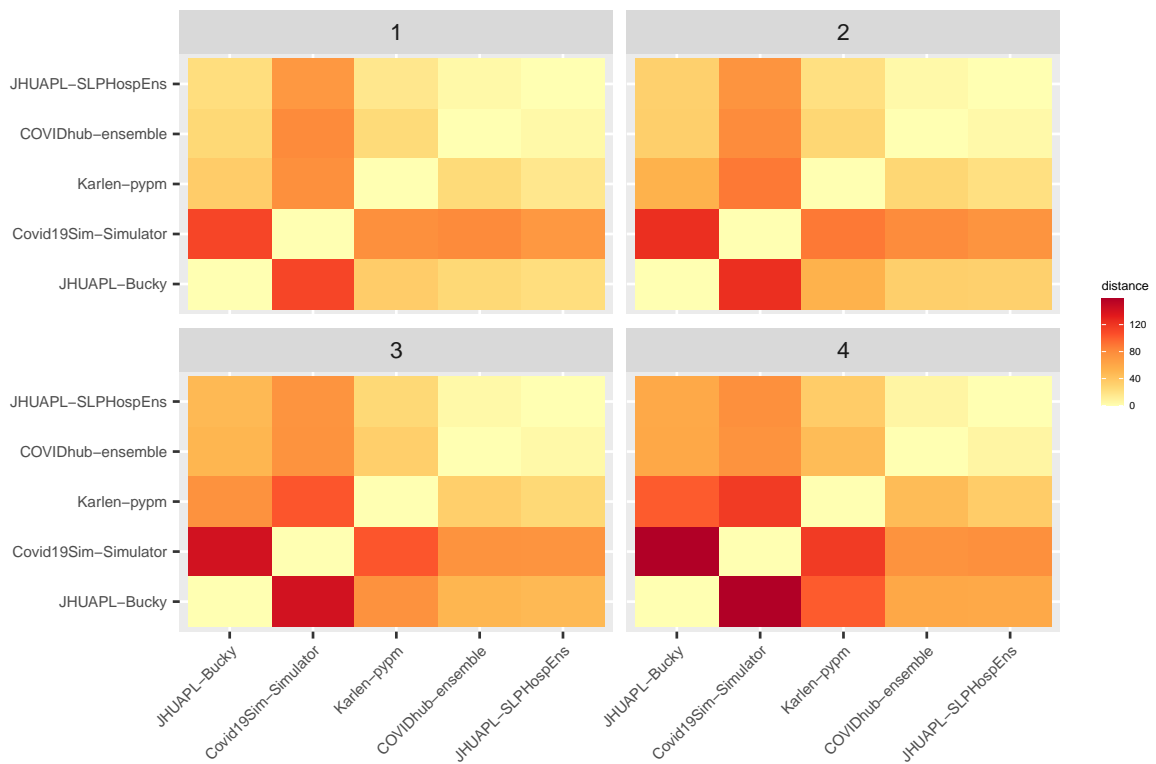
The pairwise approximated Cramer’s distances are calculated for the models that have complete submissions for all target, all 5 locations with the highest number of COVID-19 hospitalizations by end of May 2021, all probability levels, from the target end date of December, 17th 2020 to June 10th, 2021.

We can visualize the mean approximated pairwise distances across all time points in a heat map shown below. The distance from the model to itself is zero. The x -axis is arranged based in an ascending order of the model’s approximate pairwise distance from the COVIDhub-ensemble. So, the first model is the model that is most dissimilar (on average) to the ensemble in this time frame.

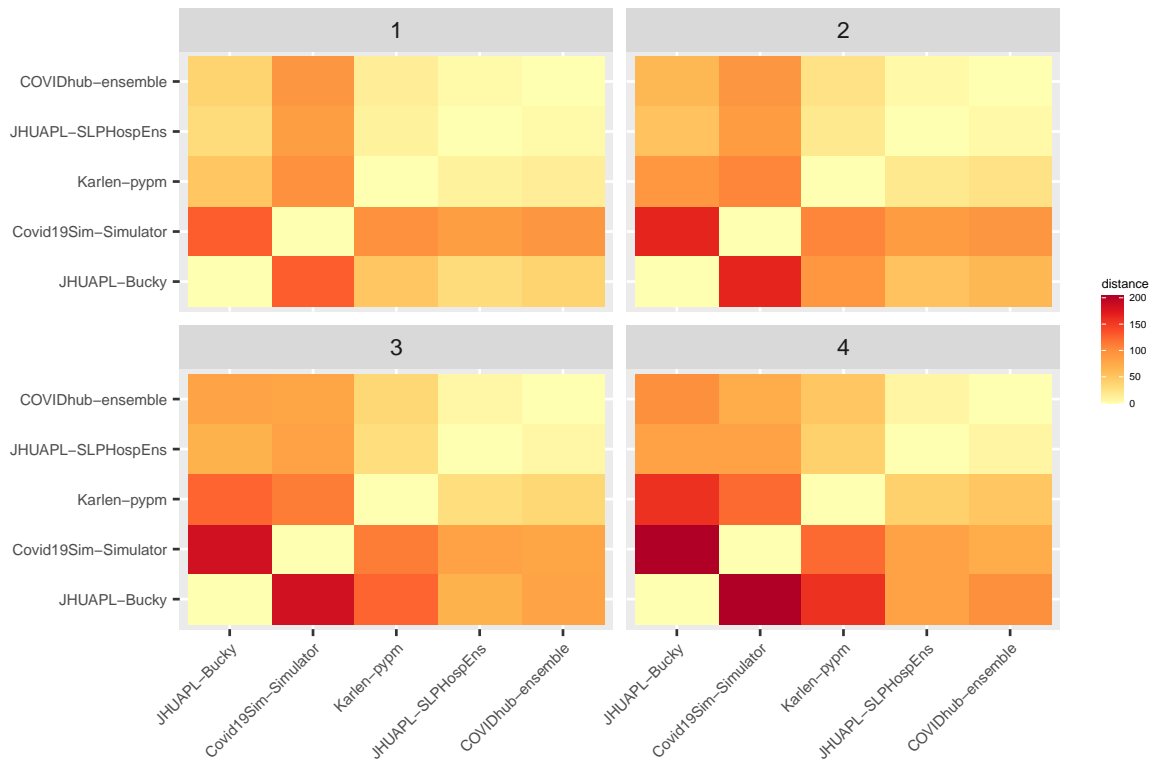
CA- Mean Approx. CD – Inc Hosp Forecasts by Horizon



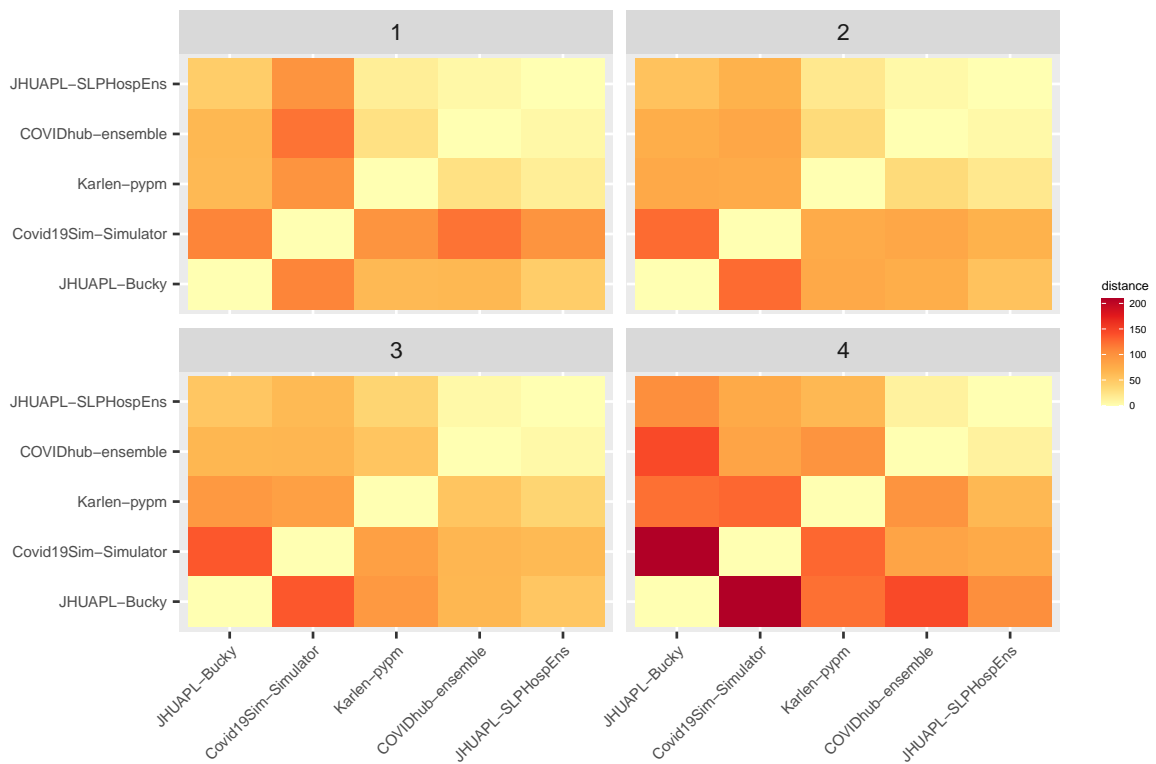
FL- Mean Approx. CD – Inc Hosp Forecasts by Horizon

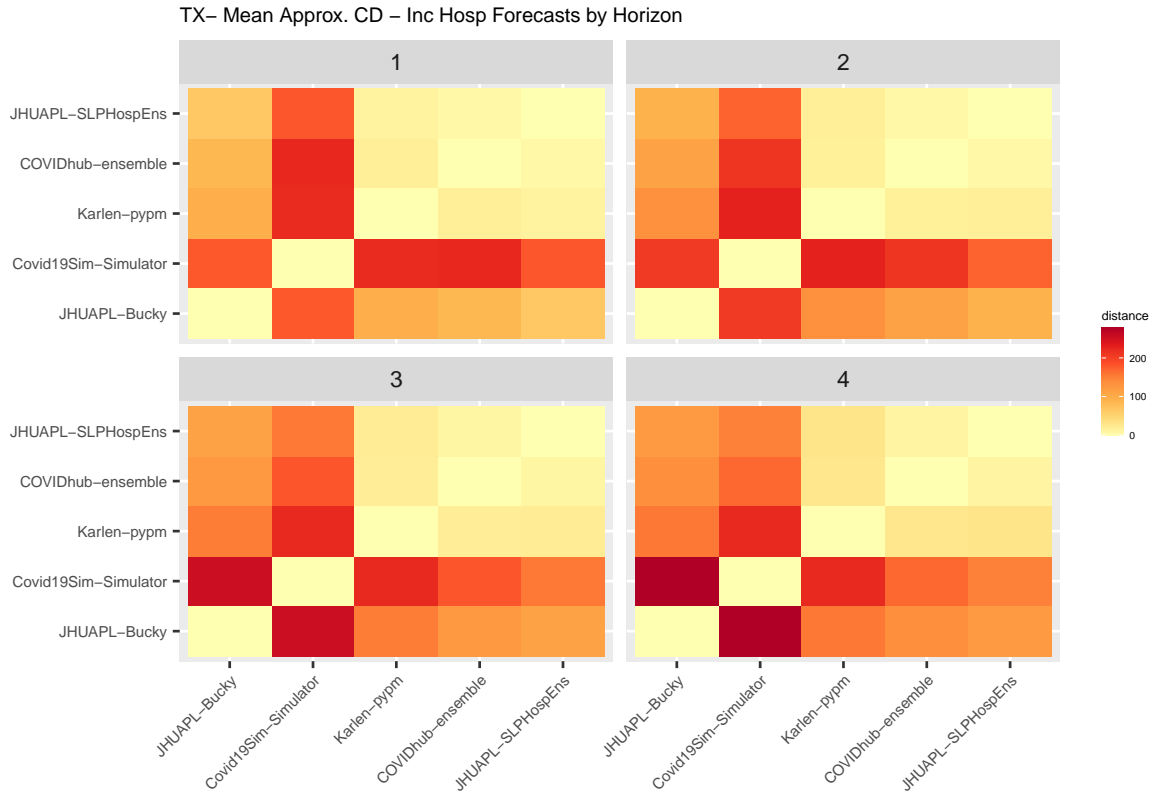


GA– Mean Approx. CD – Inc Hosp Forecasts by Horizon



NY– Mean Approx. CD – Inc Hosp Forecasts by Horizon

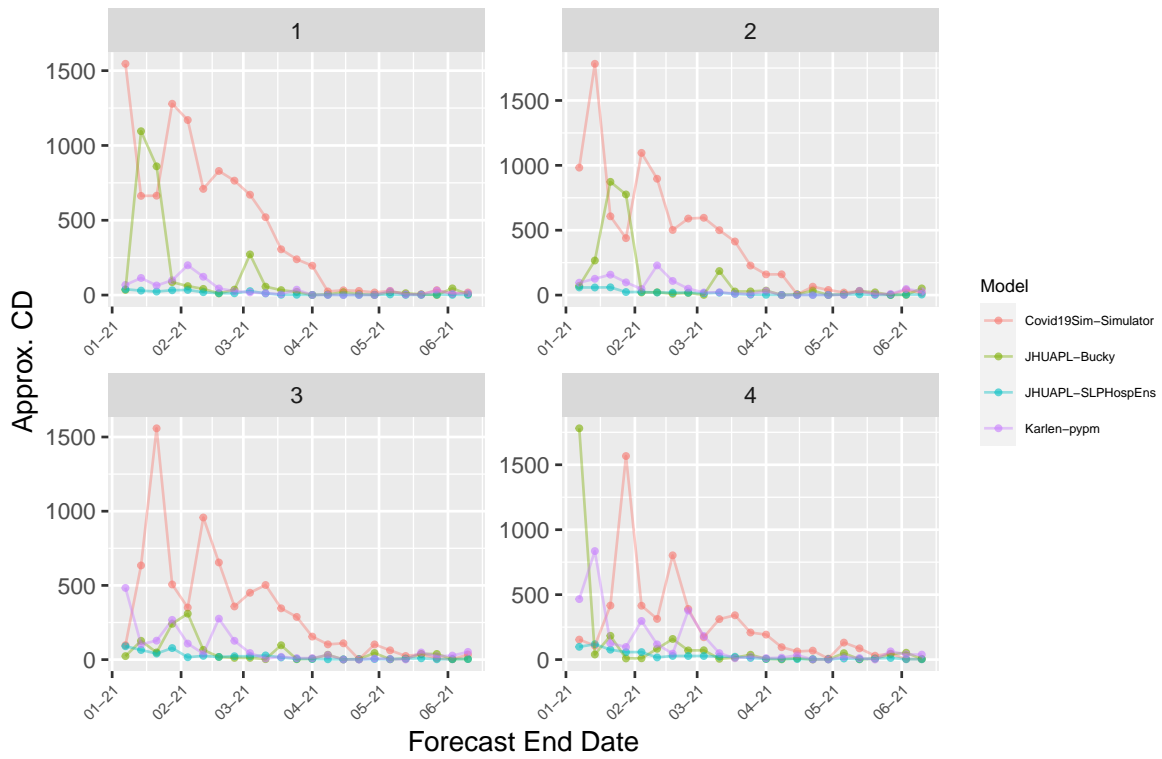




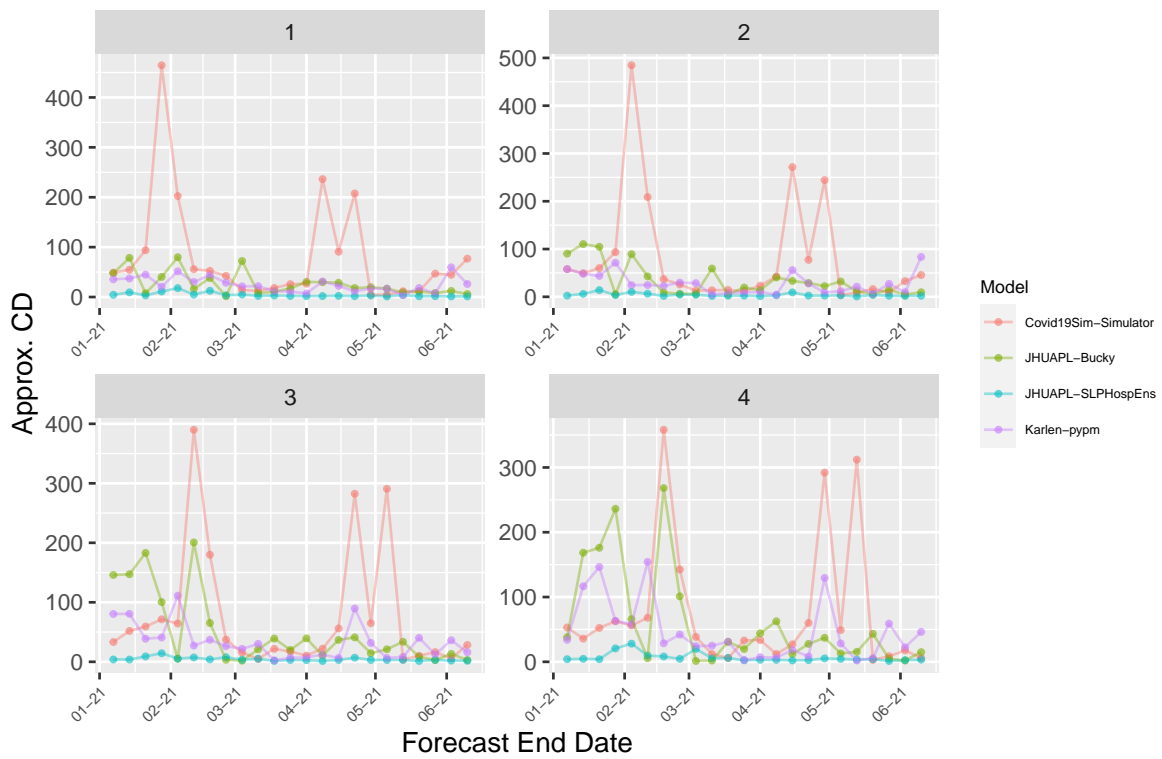
It appears that the Covid19Sim-Simulator is the least similar to the other models, across all five high count locations and horizons. It is the least similar in California and Texas and less dissimilar in the other three locations, Florida, Georgia, and New York. In Texas, JHUAPL-Bucky seems to show some difference from other models but not to the extent of the Covid19Sim-Simulator.

We can also look at the approximated pairwise distances to see how the models become more similar or dissimilar over time.

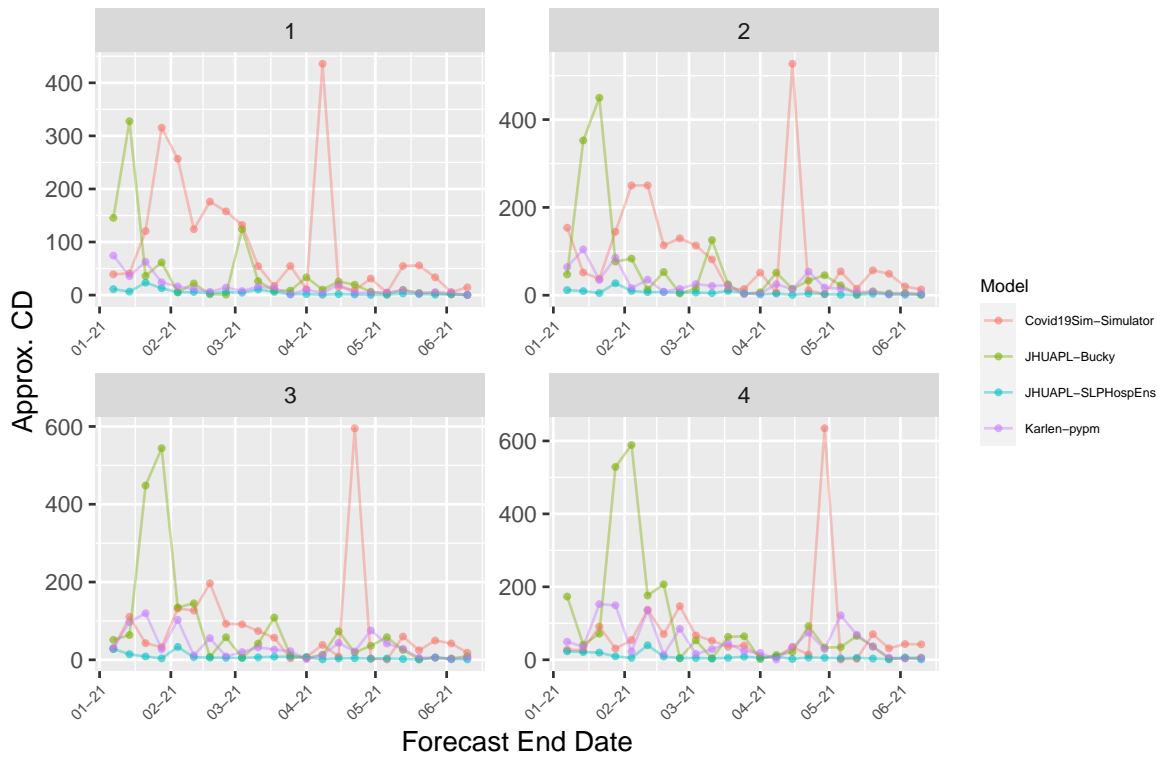
Approx. CD from COVIDhub–ensemble Over Time – CA



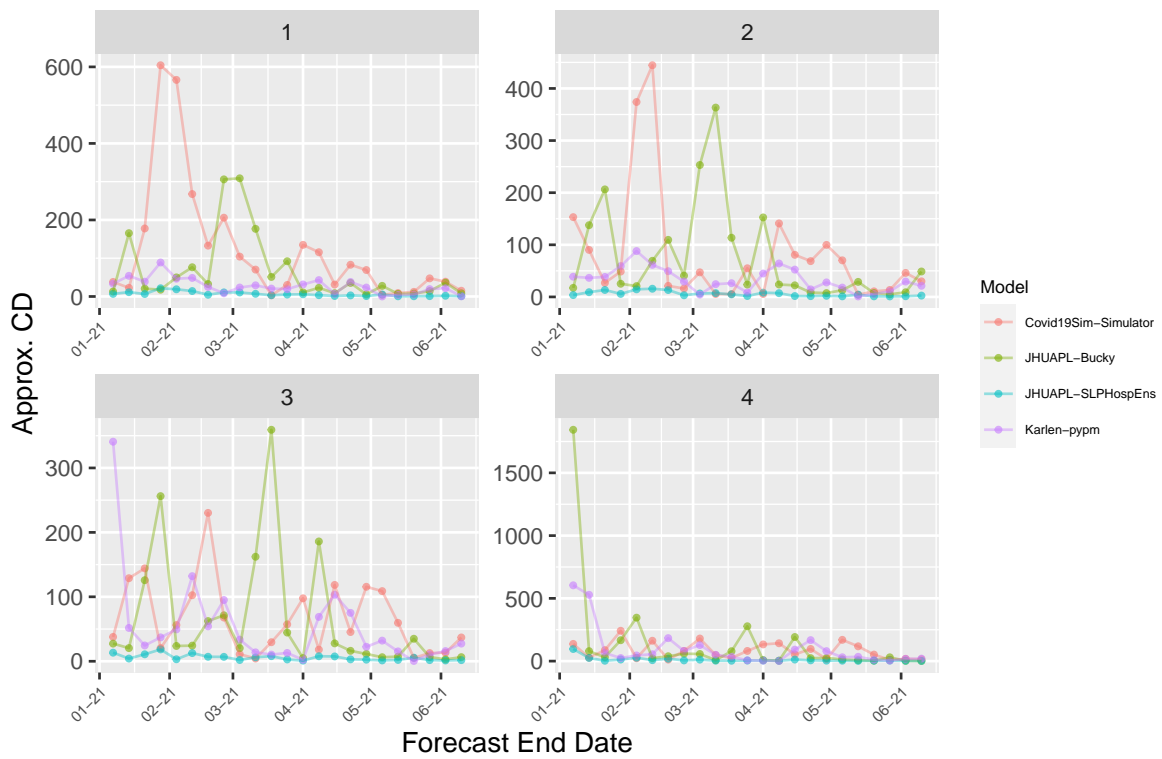
Approx. CD from COVIDhub–ensemble Over Time – FL

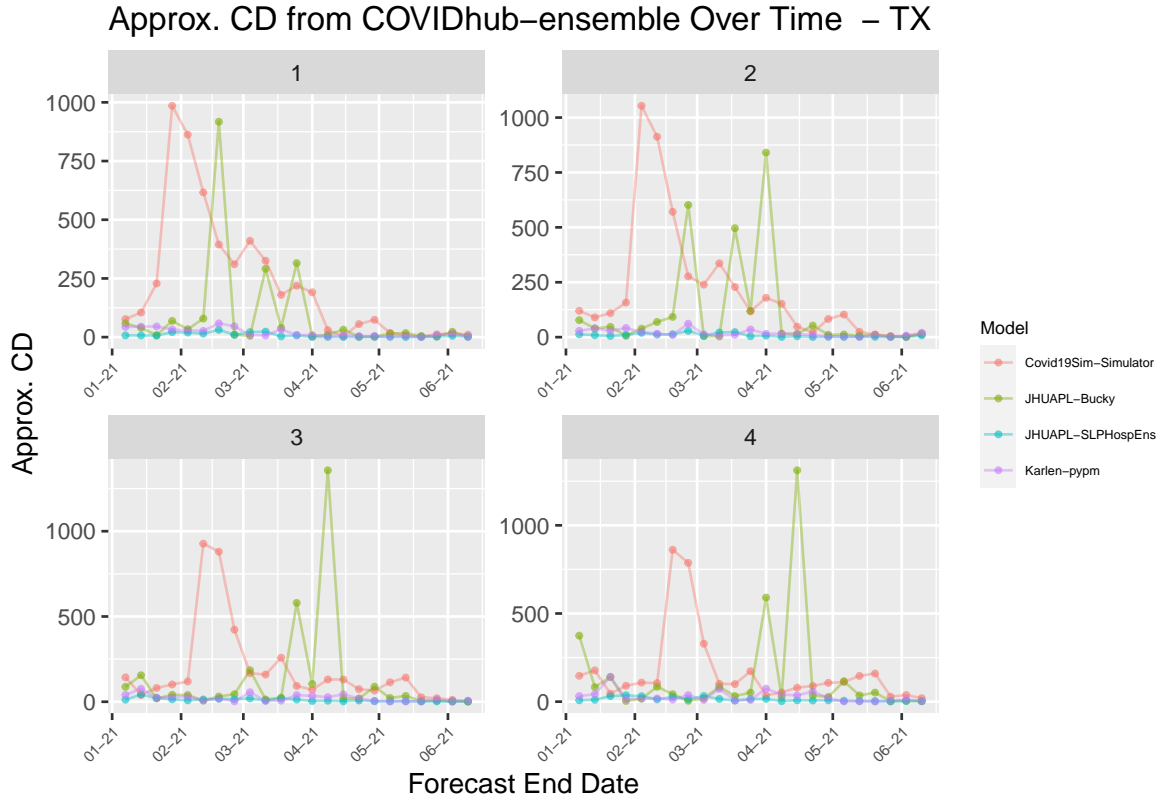


Approx. CD from COVIDhub-ensemble Over Time – GA



Approx. CD from COVIDhub-ensemble Over Time – NY

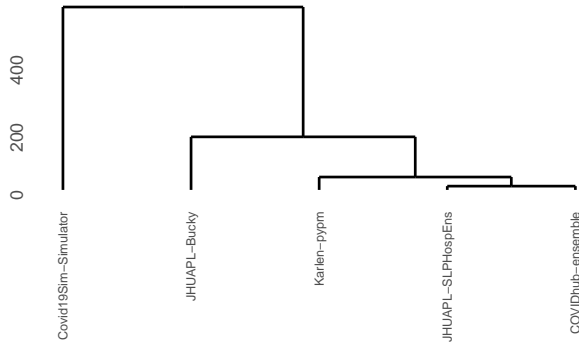




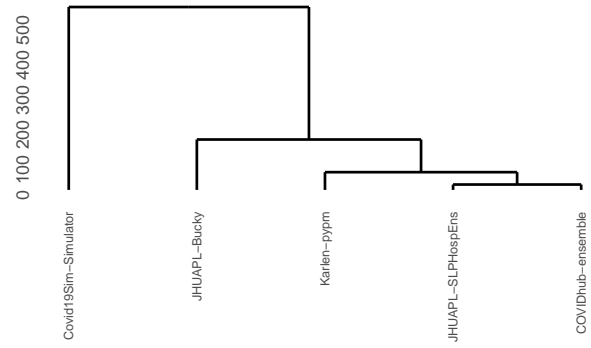
The scatterplots show that the Covid19Sim-Simulator and JHUAPL-Bucky models tend to have higher approximate Cramer's Distance from the Covidhub-ensemble model compared to the other three models. This seems to align with the results shown in the heat maps above that show that Covid19Sim-Simulator and JHUAPL-Bucky tend to have the highest mean Cramer's Distance from the other models.

We can also cluster the distances using hierarchical clustering.

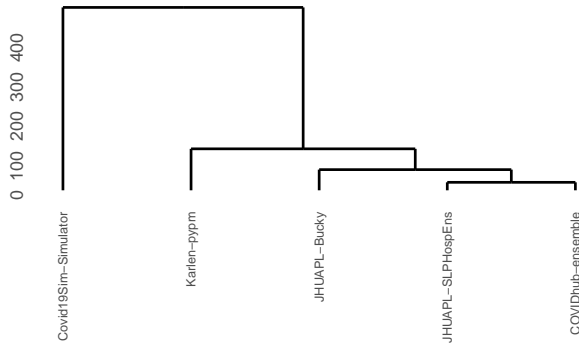
Dendrogram – 1 wk ahead inc hosp – CA



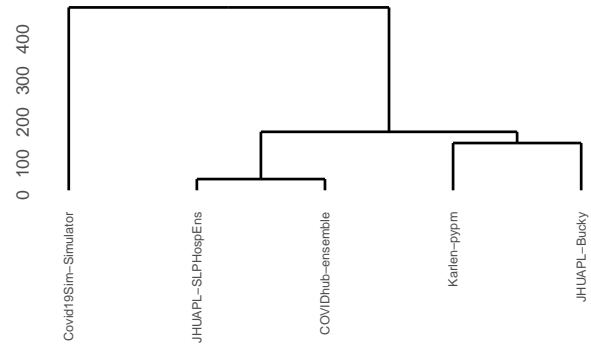
Dendrogram – 2 wk ahead inc hosp – CA



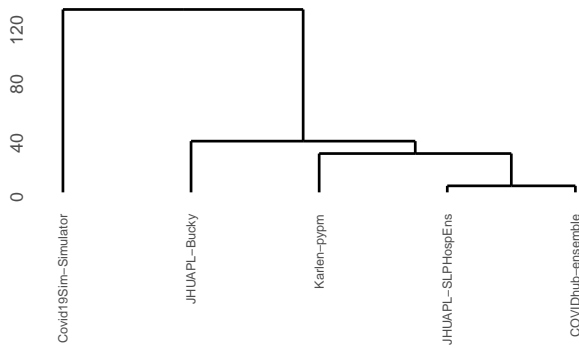
Dendrogram – 3 wk ahead inc hosp – CA



Dendrogram – 4 wk ahead inc hosp – CA



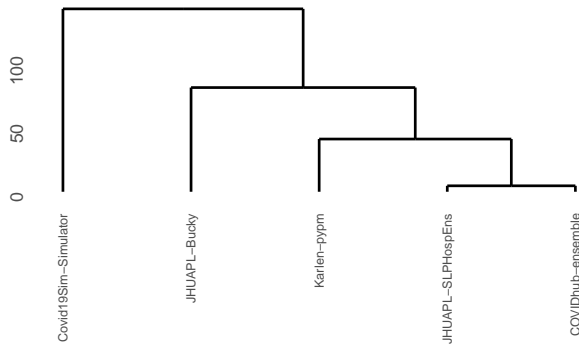
Dendrogram – 1 wk ahead inc hosp – FL



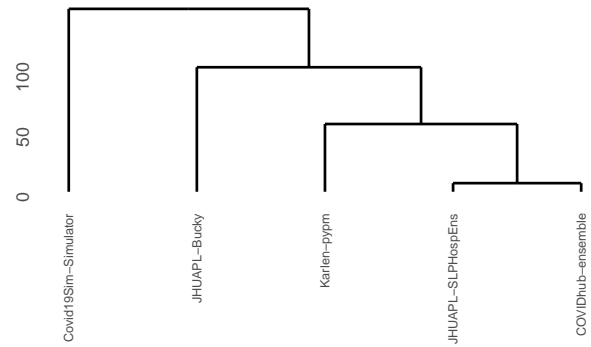
Dendrogram – 2 wk ahead inc hosp – FL



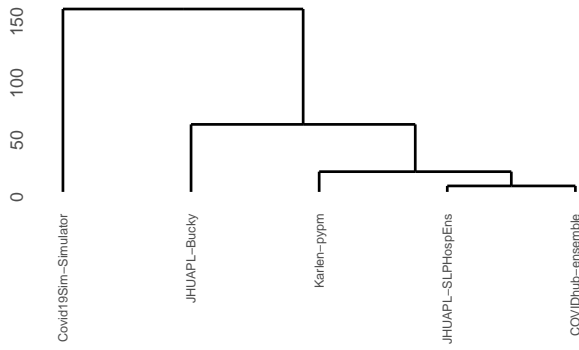
Dendrogram – 3 wk ahead inc hosp – FL



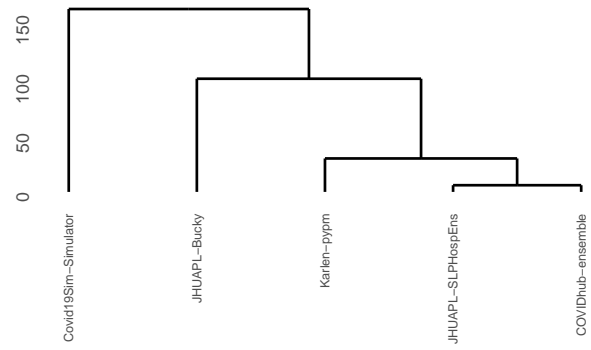
Dendrogram – 4 wk ahead inc hosp – FL



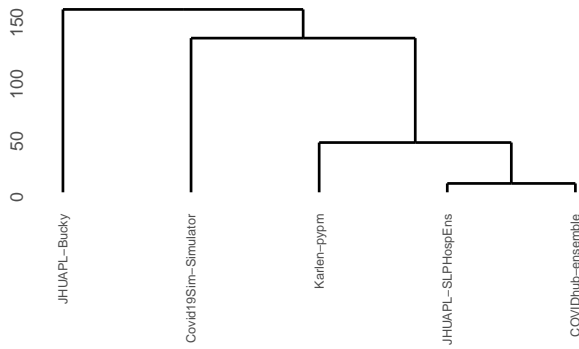
Dendrogram – 1 wk ahead inc hosp – GA



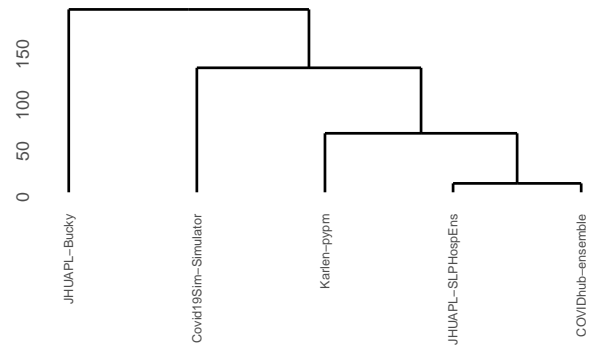
Dendrogram – 2 wk ahead inc hosp – GA



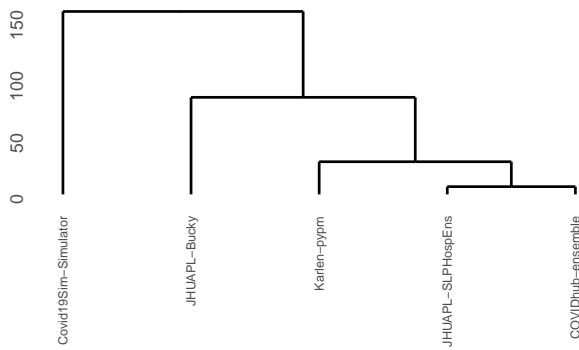
Dendrogram – 3 wk ahead inc hosp – GA



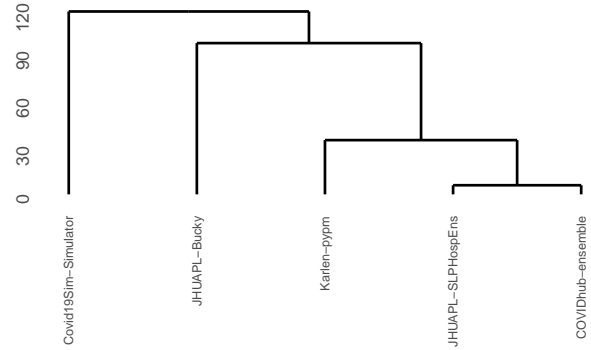
Dendrogram – 4 wk ahead inc hosp – GA



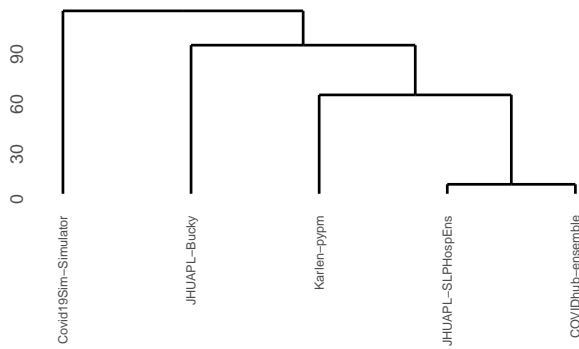
Dendrogram – 1 wk ahead inc hosp – NY



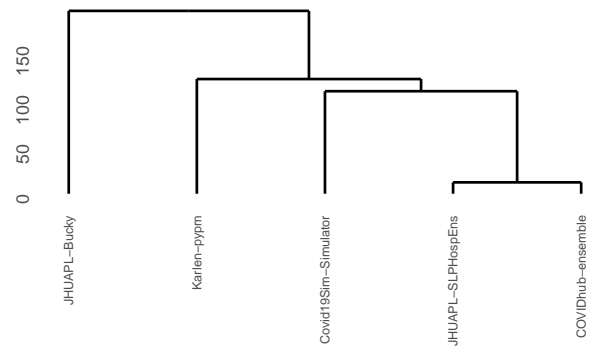
Dendrogram – 2 wk ahead inc hosp – NY

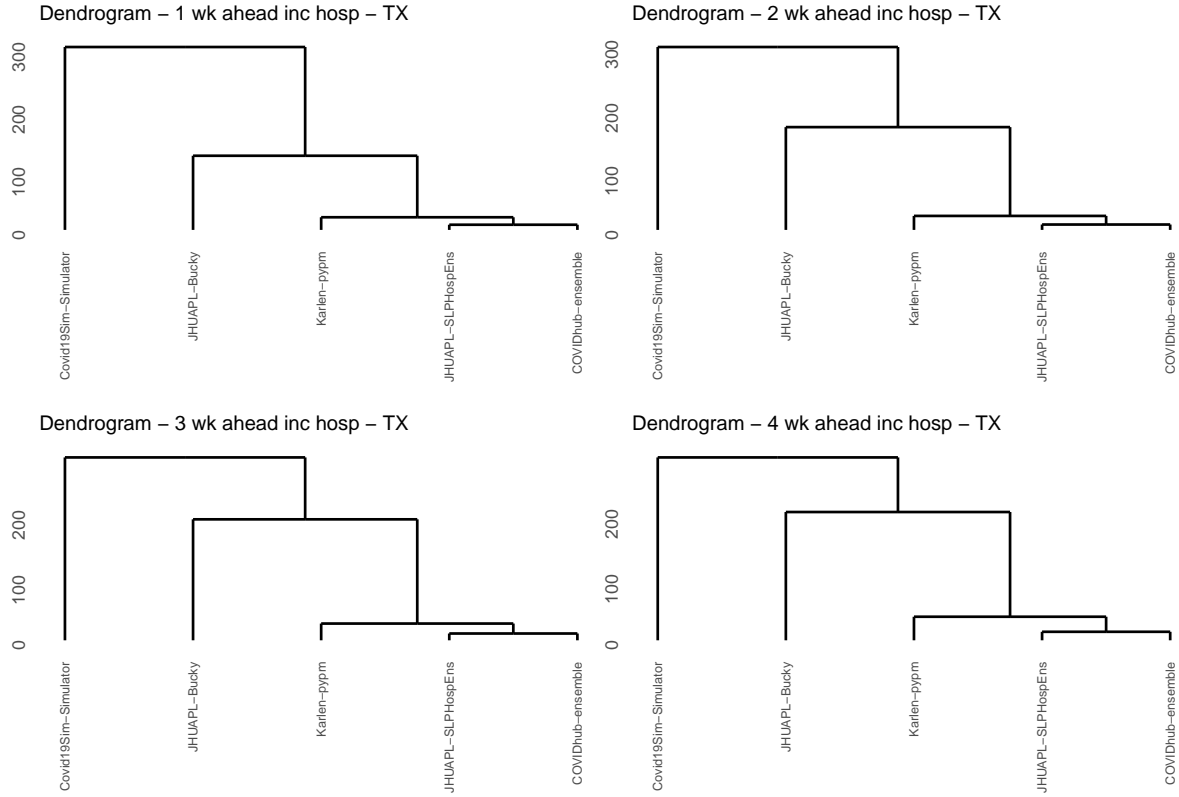


Dendrogram – 3 wk ahead inc hosp – NY



Dendrogram – 4 wk ahead inc hosp – NY





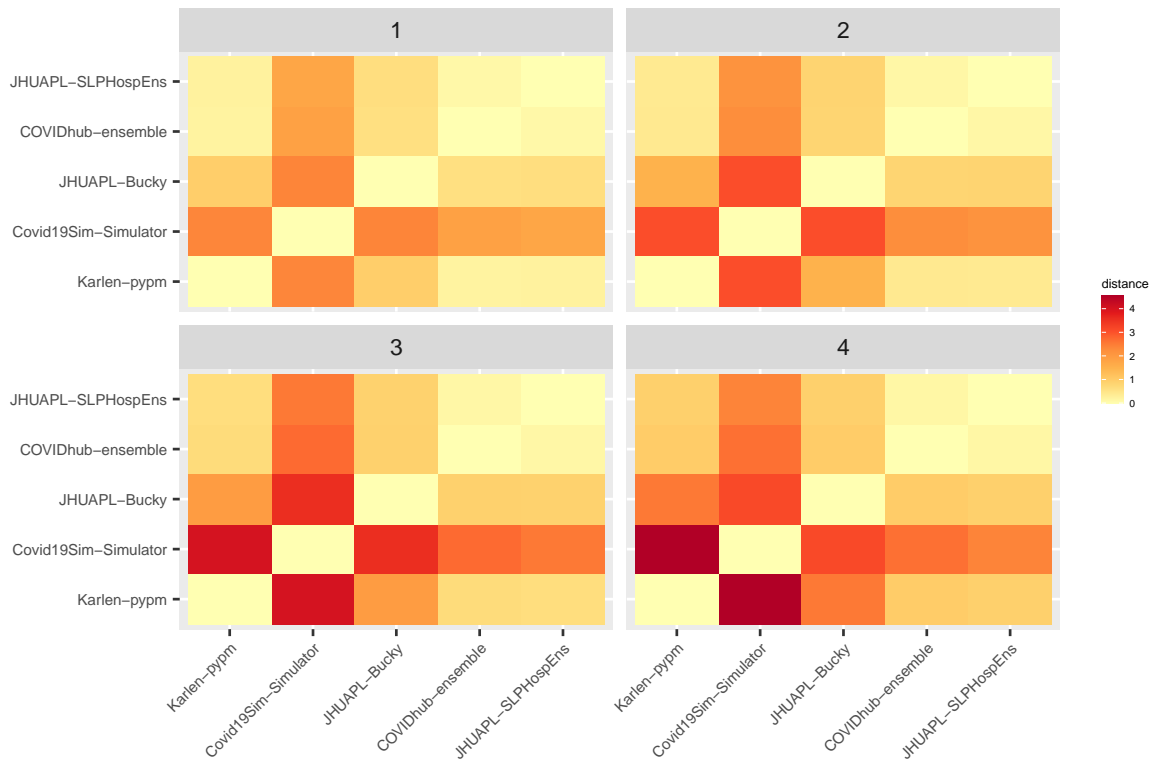
Except for the 4 week ahead inc hosp dendrogram for California, all of the dendrograms show that Covid19Sim-Simulator and JHUAPL-Bucky are the least similar to the other models, splitting away in the first or second groupings, respectively.

5 locations with the lowest number of COVID-19 deaths by the end of February 2021

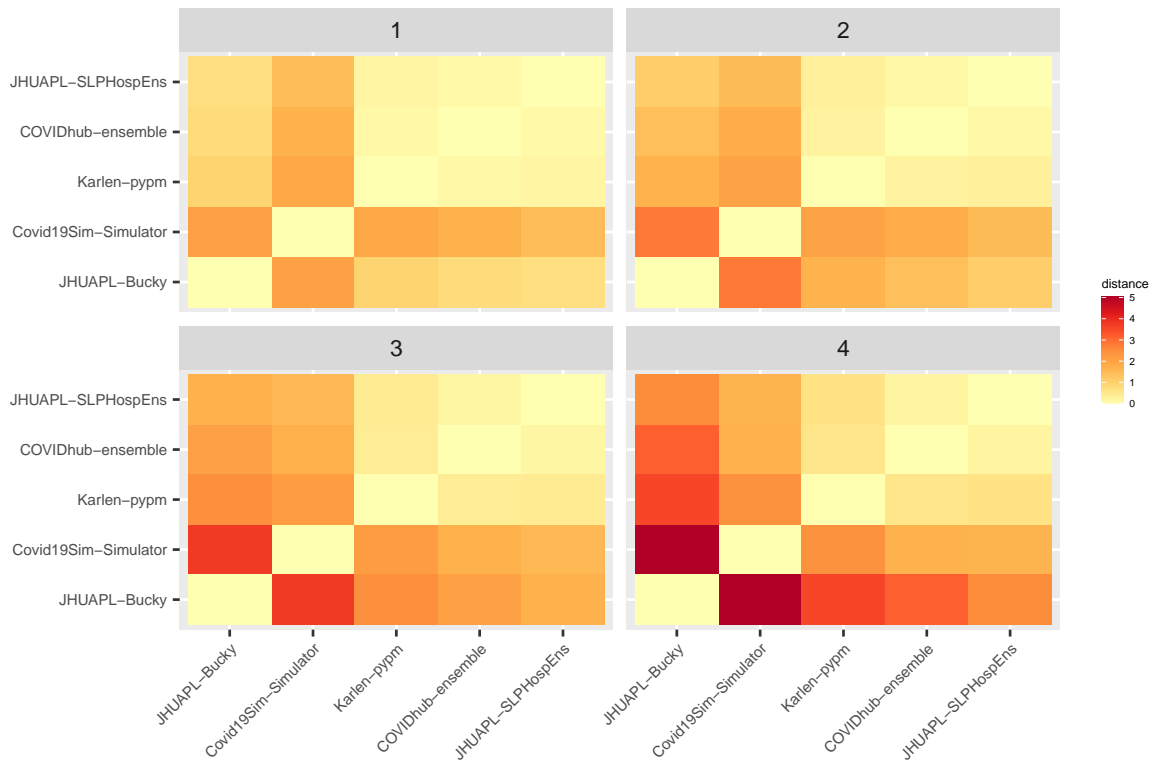
The pairwise approximated Cramer’s distances are calculated for the models that have complete submissions for all target, all 5 locations with the lowest number of COVID-19 deaths by the end of February 2021, all probability levels, from the target end date of December, 17th 2020 to June 10th, 2021 .

We can visualize the mean approximated pairwise distances across all time points in a heat map shown below. The distance from the model to itself is zero. The x -axis is arranged based in an ascending order of the model’s approximate pairwise distance from the COVIDhub-ensemble. So, the first model is the model that is most dissimilar (on average) to the ensemble in this time frame.

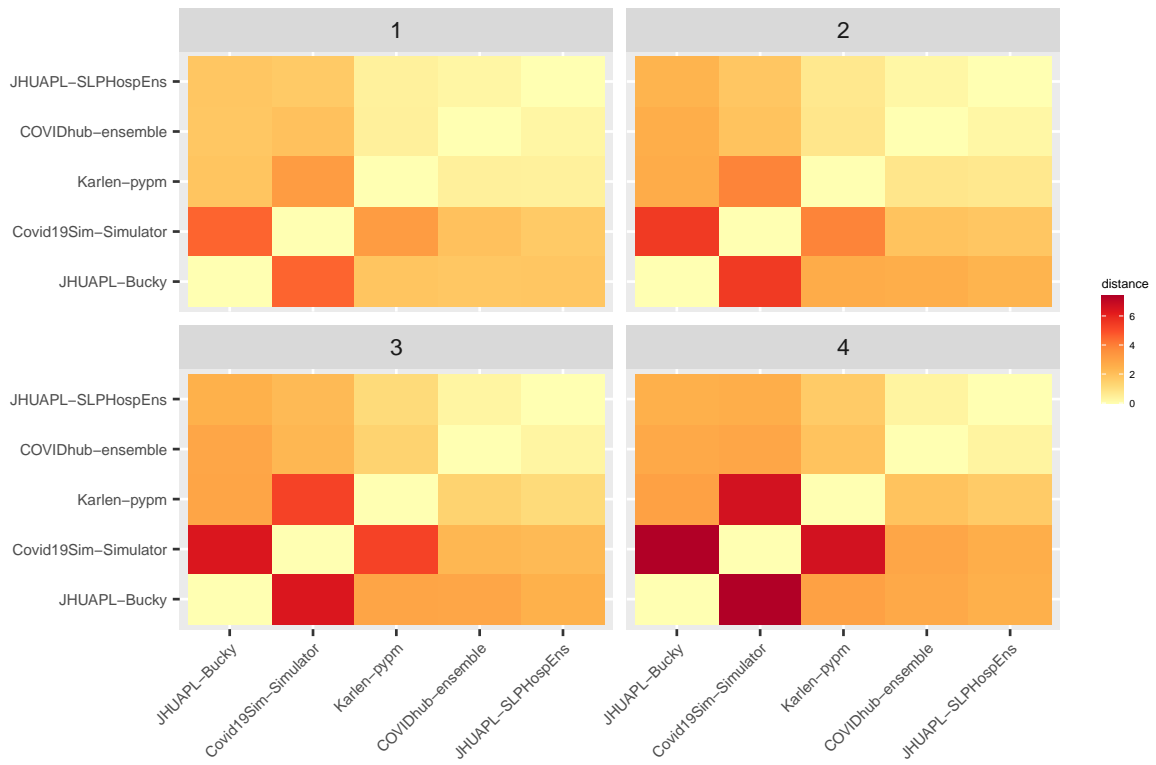
AK– Mean Approx. CD – Inc Hosp Forecasts by Horizon



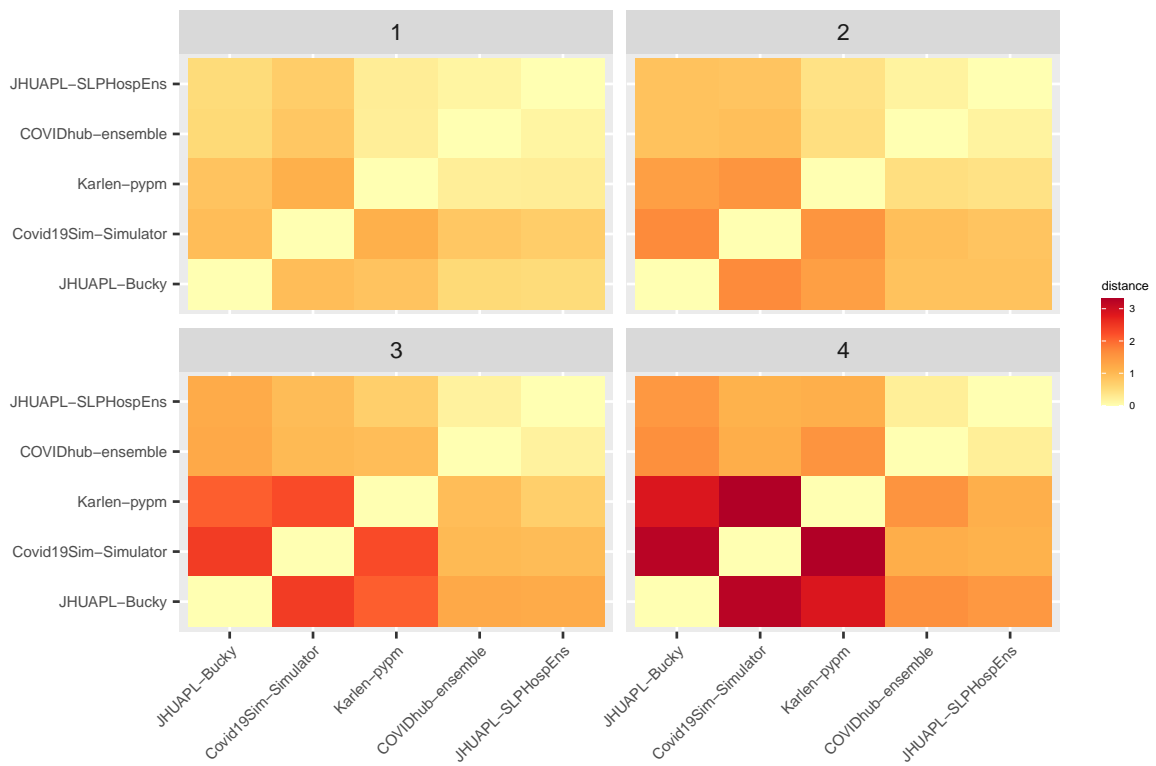
HI– Mean Approx. CD – Inc Hosp Forecasts by Horizon

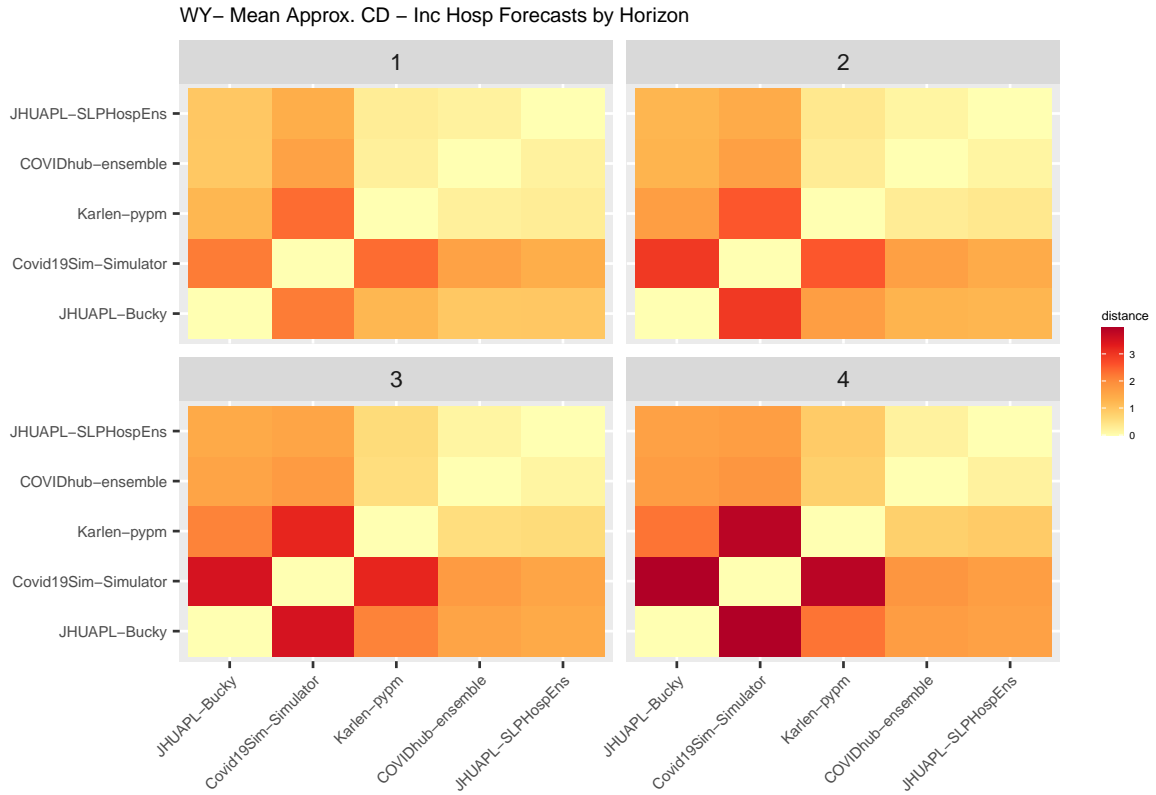


ND- Mean Approx. CD – Inc Hosp Forecasts by Horizon



VT- Mean Approx. CD – Inc Hosp Forecasts by Horizon

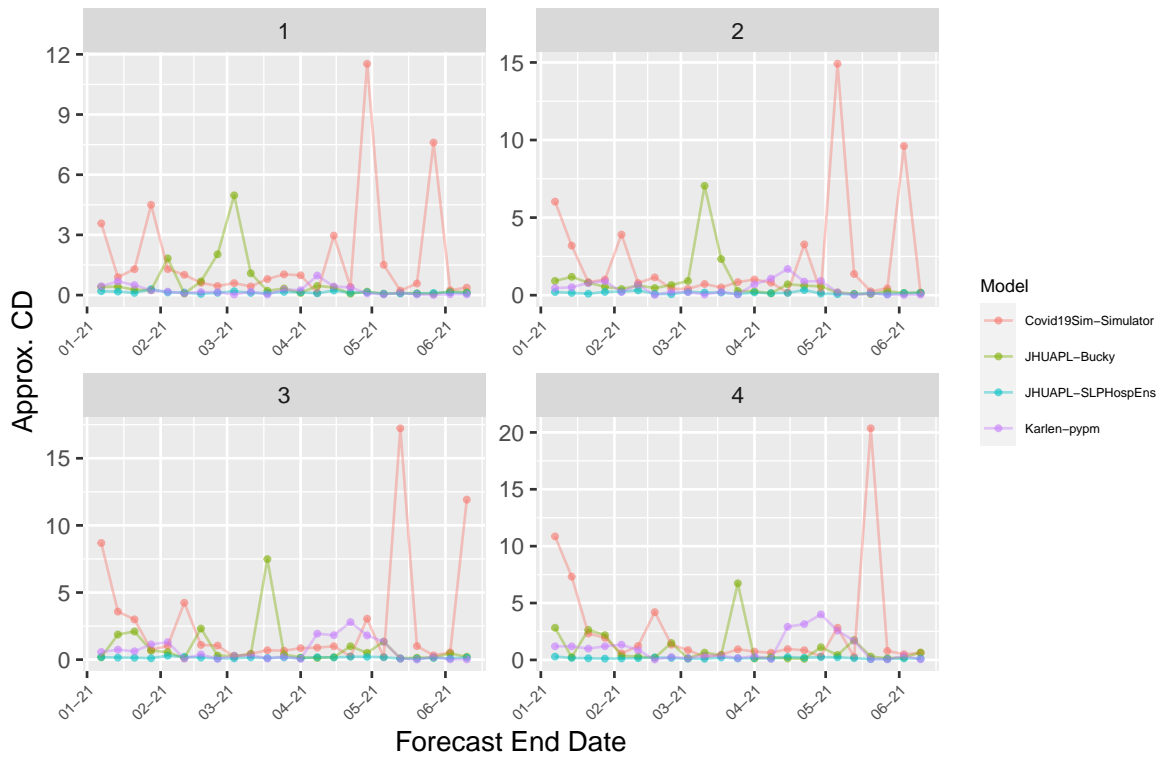




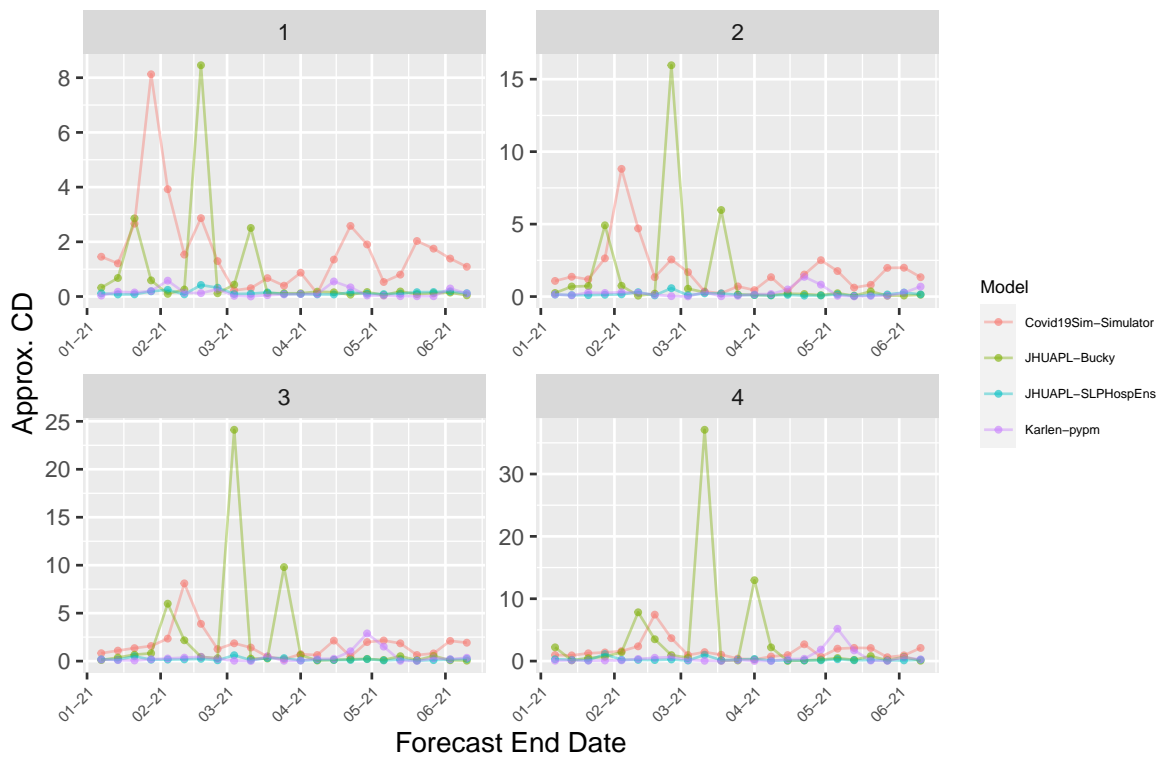
For the five low count locations, Covid19Sim-Simulator is likewise the most dissimilar from the other models with the largest calculator Cramer's distance over most locations and horizons. However, JHUAPL-Bucky is also fairly dissimilar from other models at horizons of three or four weeks, surpassing the Covid19Sim-Simulator in Cramer's Distance from other models in Hawaii and Vermont for the two larger horizons.

We can also look at the approximated pairwise distances to see how the models become more similar or dissimilar over time.

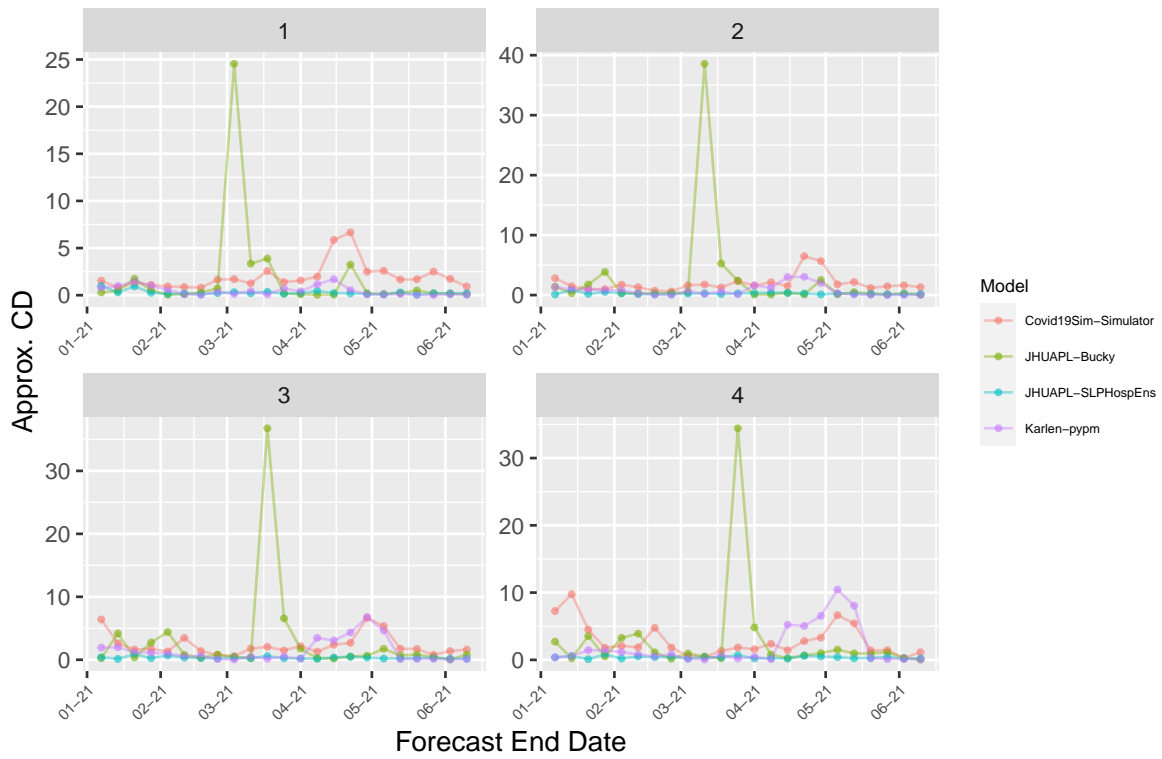
Approx. CD from COVIDhub–ensemble Over Time – AK



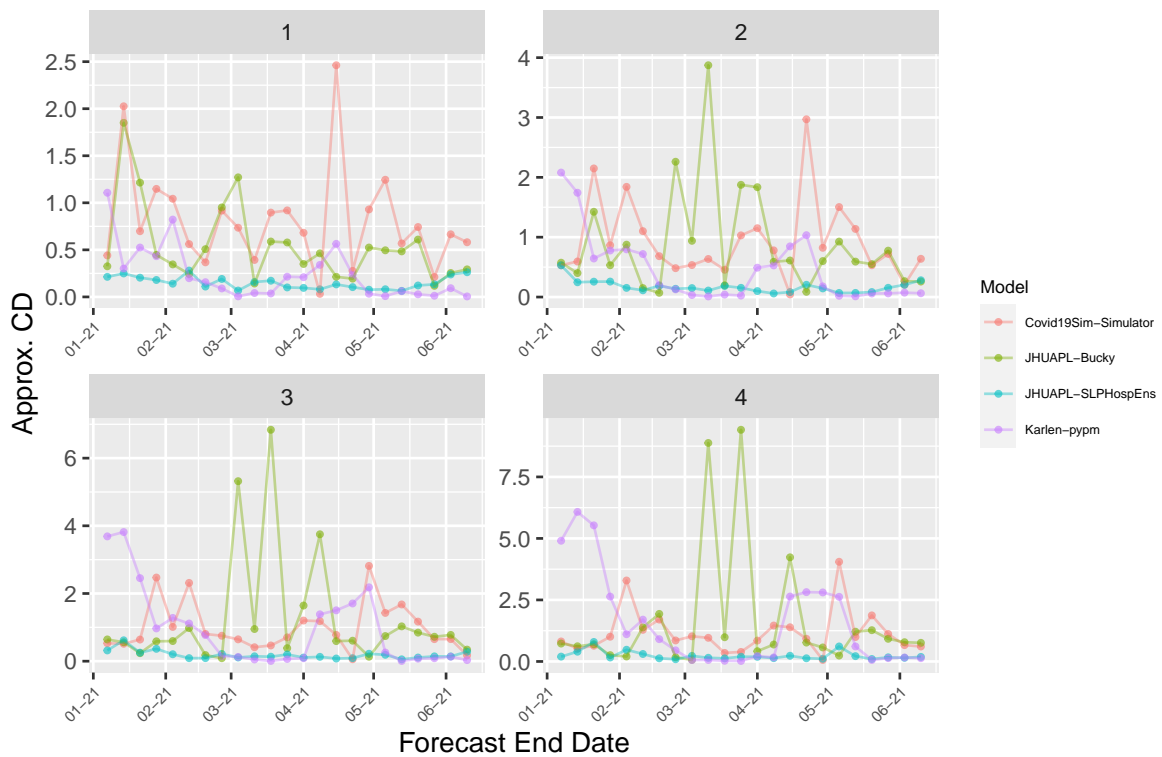
Approx. CD from COVIDhub–ensemble Over Time – HI



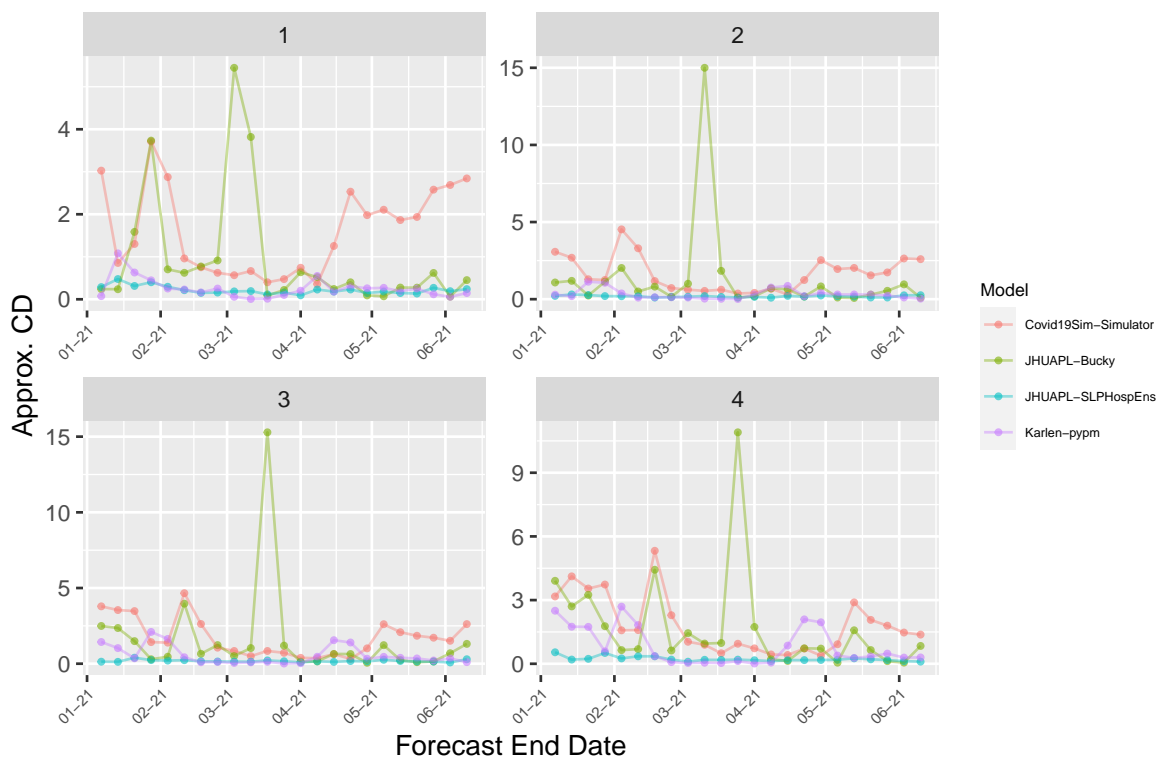
Approx. CD from COVIDhub–ensemble Over Time – ND



Approx. CD from COVIDhub–ensemble Over Time – VT



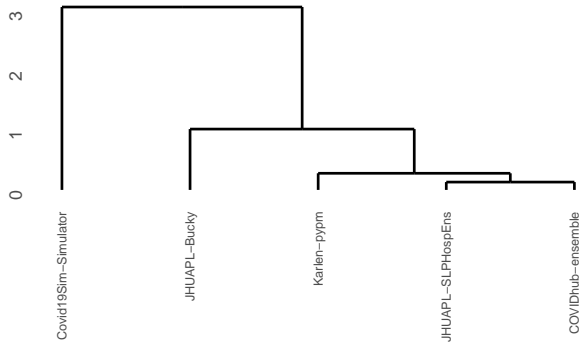
Approx. CD from COVIDhub–ensemble Over Time – WY



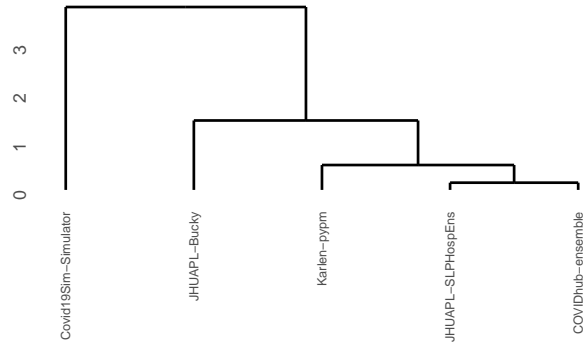
Like for the high count locations, Covid19Sim-Simulator and JHUAPL-Bucky are shown to generally have the highest Cramer's Distance from the Covidhub-ensemble model. However, we note that in Arkansas only Covid19Sim-Simulator has a noticeably larger Cramer's Distance while in North Dakota only JHUAPL-Bucky has a noticeably larger Cramer's Distance.

We can cluster the distances using hierarchical clustering. Different linkages will result in different clusters, we probably should investigate more later.

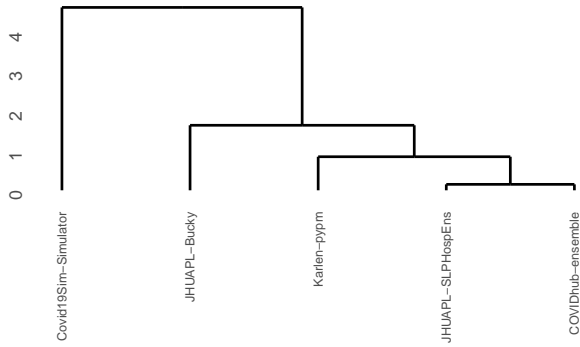
Dendrogram – 1 wk ahead inc hosp – AK



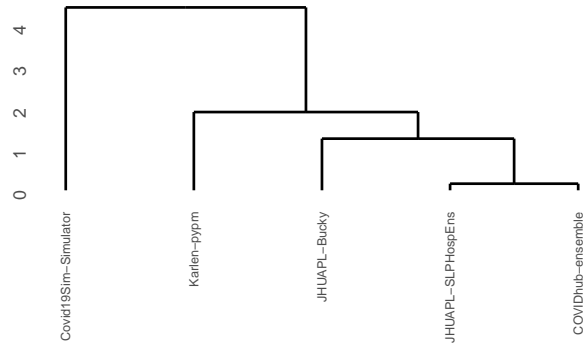
Dendrogram – 2 wk ahead inc hosp – AK



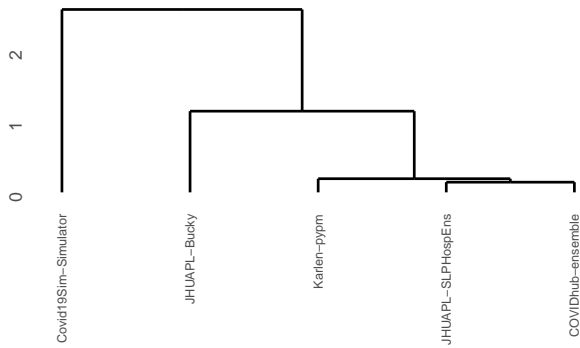
Dendrogram – 3 wk ahead inc hosp – AK



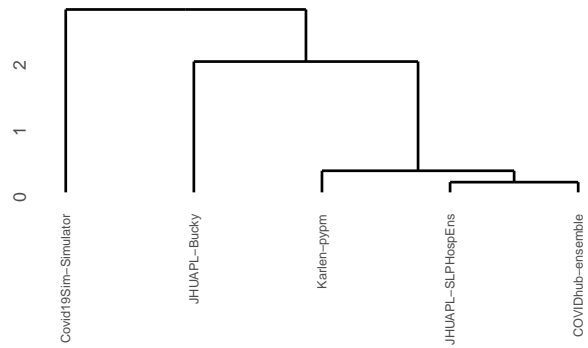
Dendrogram – 4 wk ahead inc hosp – AK



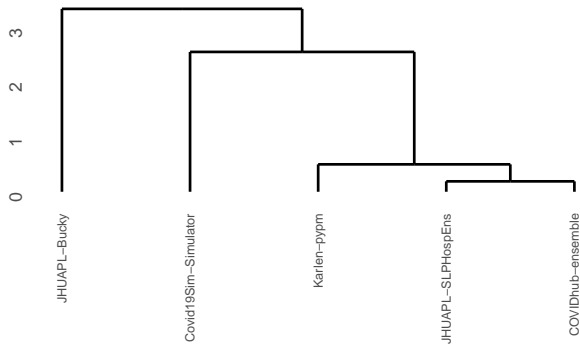
Dendrogram – 1 wk ahead inc hosp – HI



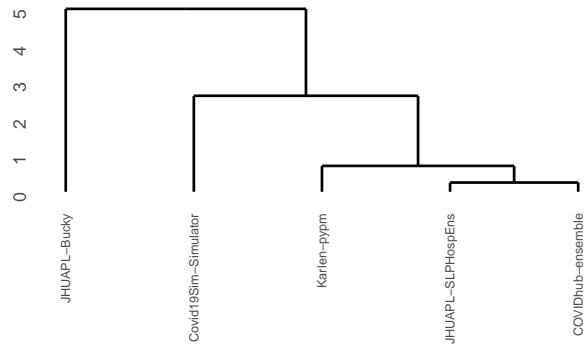
Dendrogram – 2 wk ahead inc hosp – HI

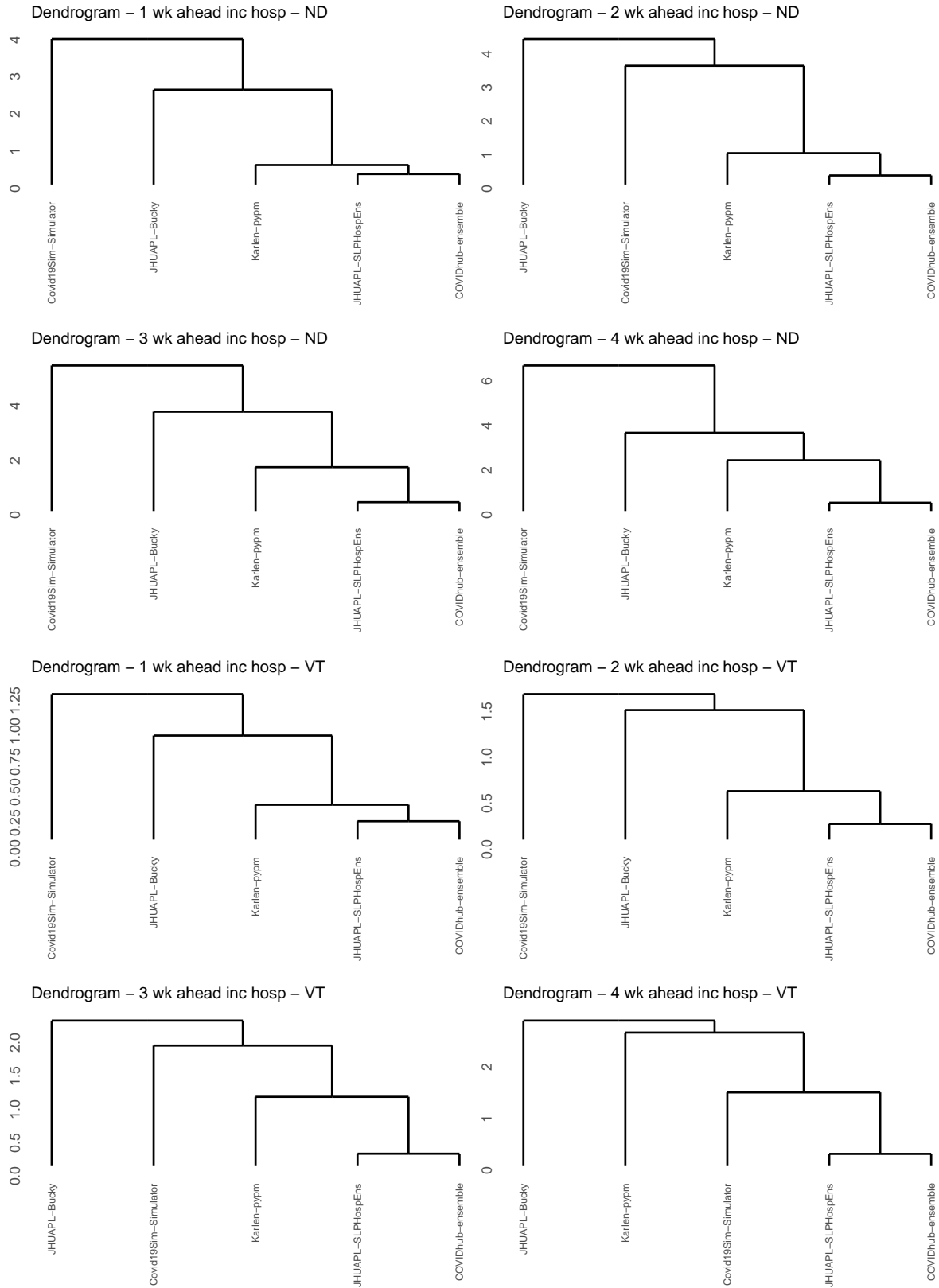


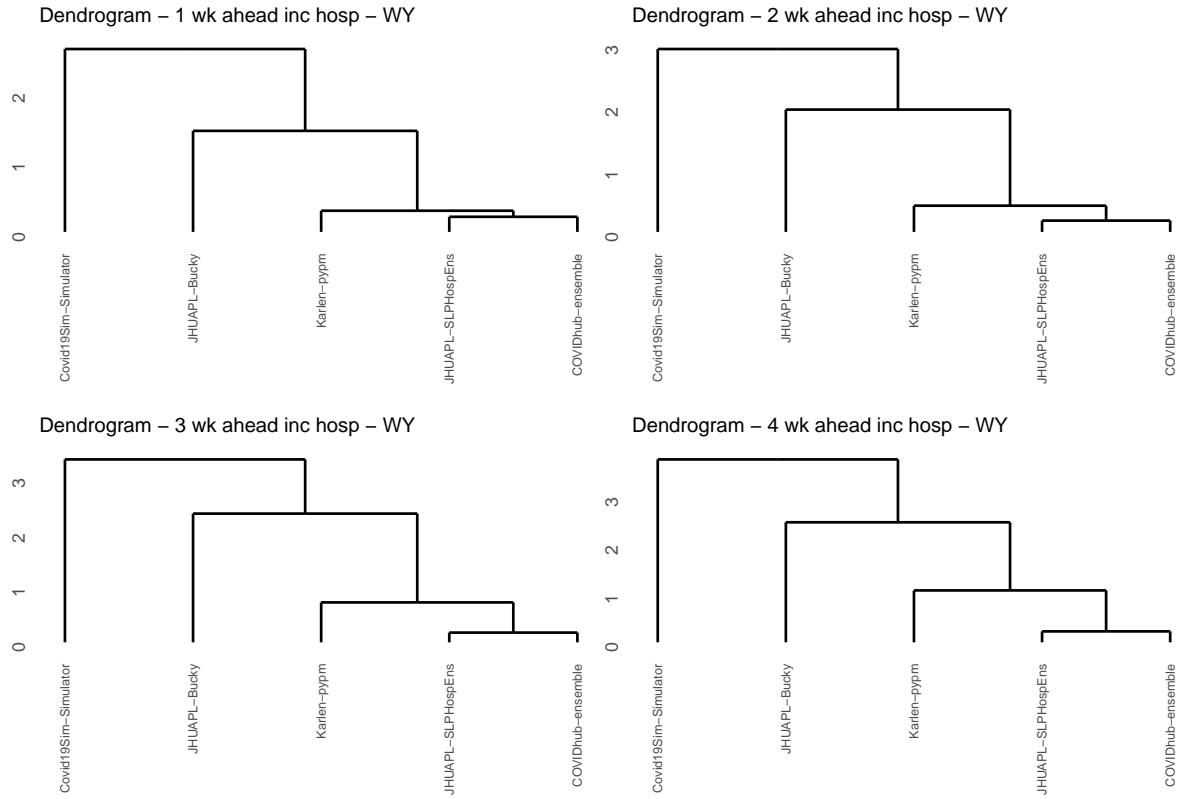
Dendrogram – 3 wk ahead inc hosp – HI



Dendrogram – 4 wk ahead inc hosp – HI







Covid19Sim-Simulator and JHUAPL-Bucky are the least similar according to the dendrograms for the five low locations, except for 4 wk ahead in Arkansas and 4 wk ahead in Vermont, which is similar to the results shown for the high count locations.