Forecasting Model Similarity (for inc hosp)

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

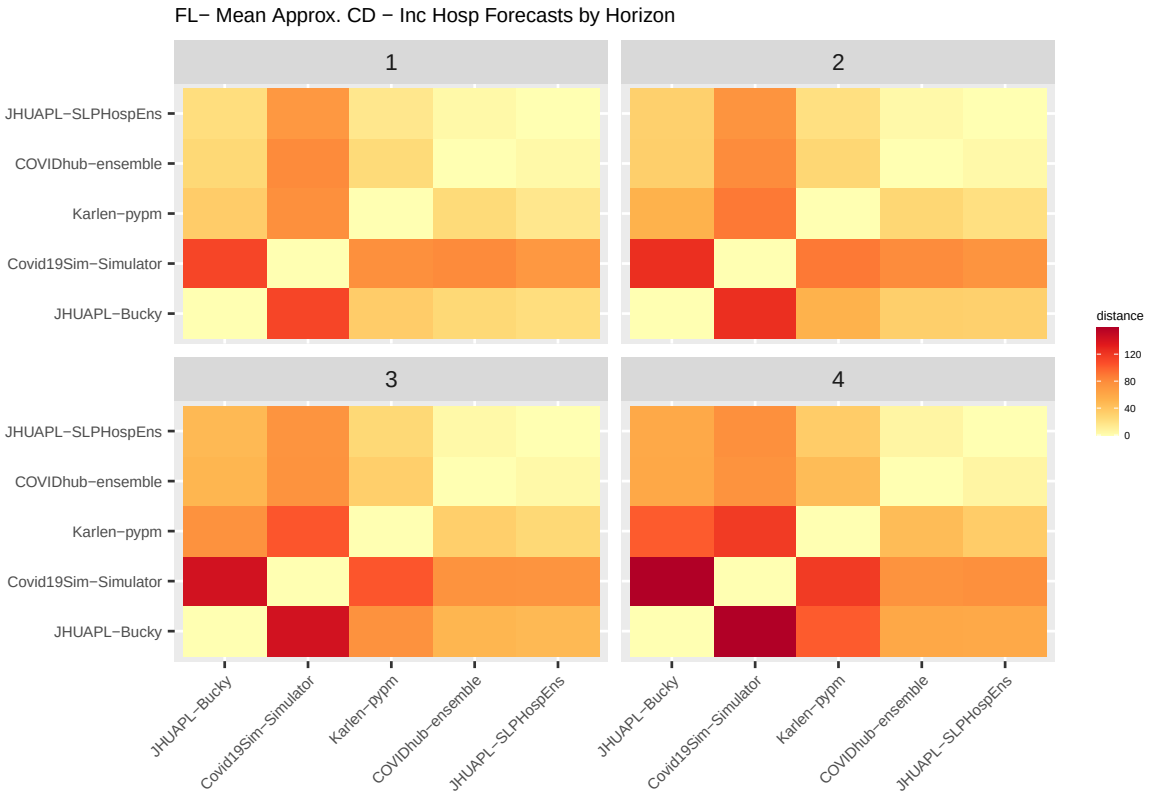
06/10/2021

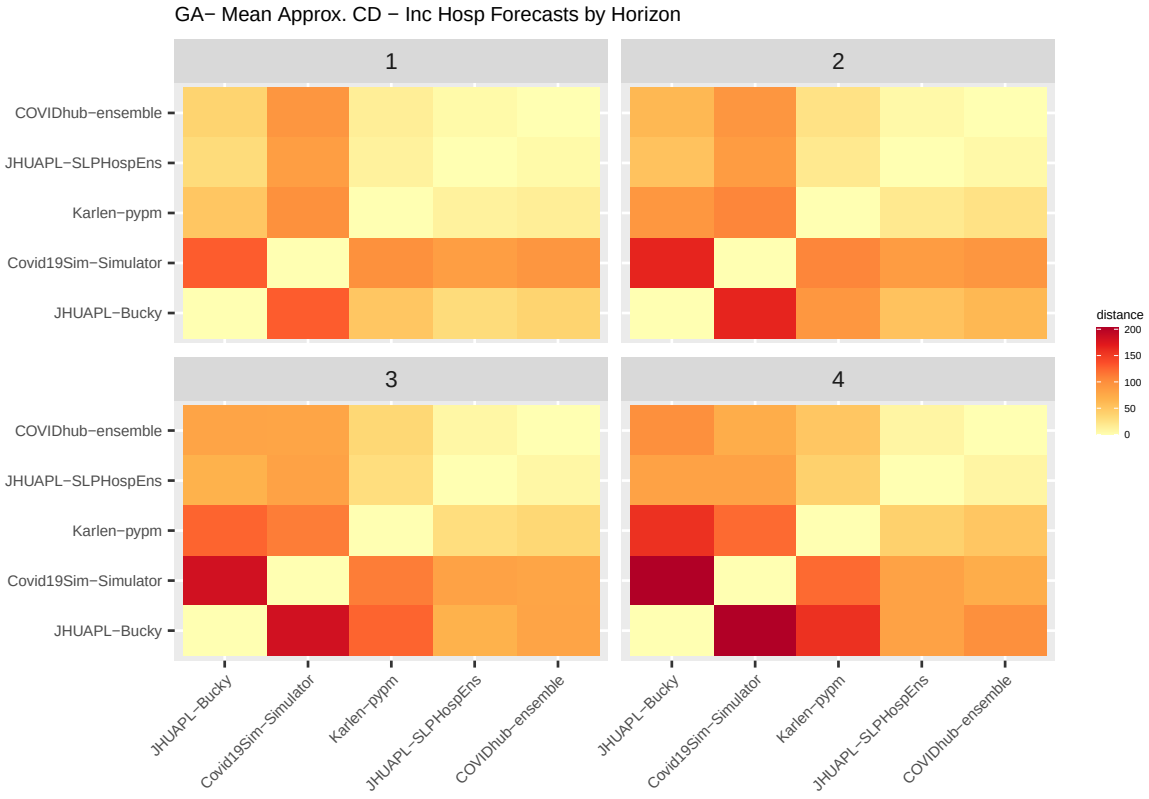
COVID-19 Forecasting Model Similarity Analysis for 1-4 Week Ahead Incident Hospitalization

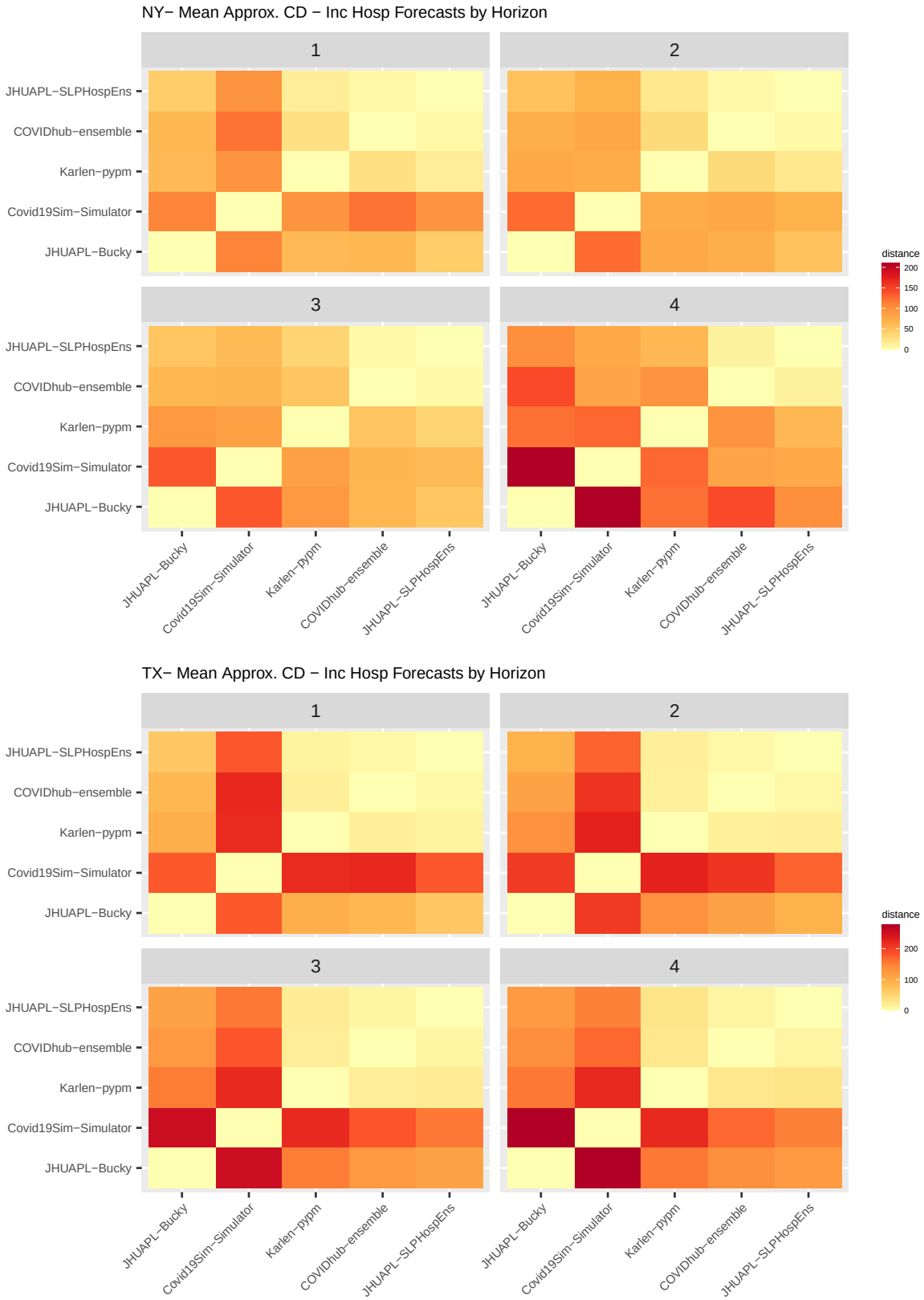
5 locations with the highest number of COVID-19 hospitalizations by the end of February 2021

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.

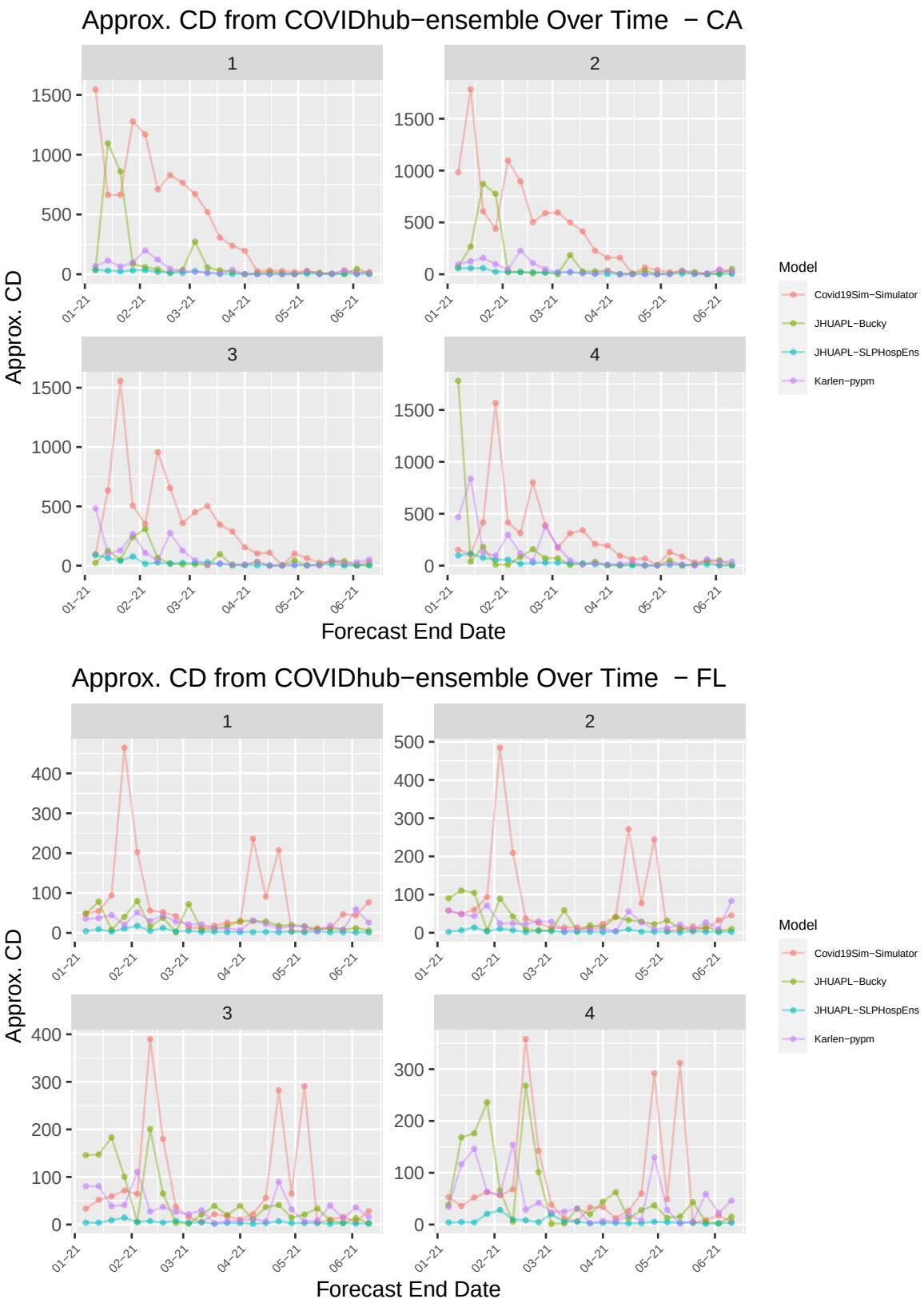

FL− Mean Approx. CD − Inc Hosp Forecasts by Horizon
1
2
3
4
JHUAPL−SLPHospEns
COVIDhub−ensemble
Karlen−pypm
Covid19Sim−Simulator
JHUAPL−Bucky
distance
120
80
40
0

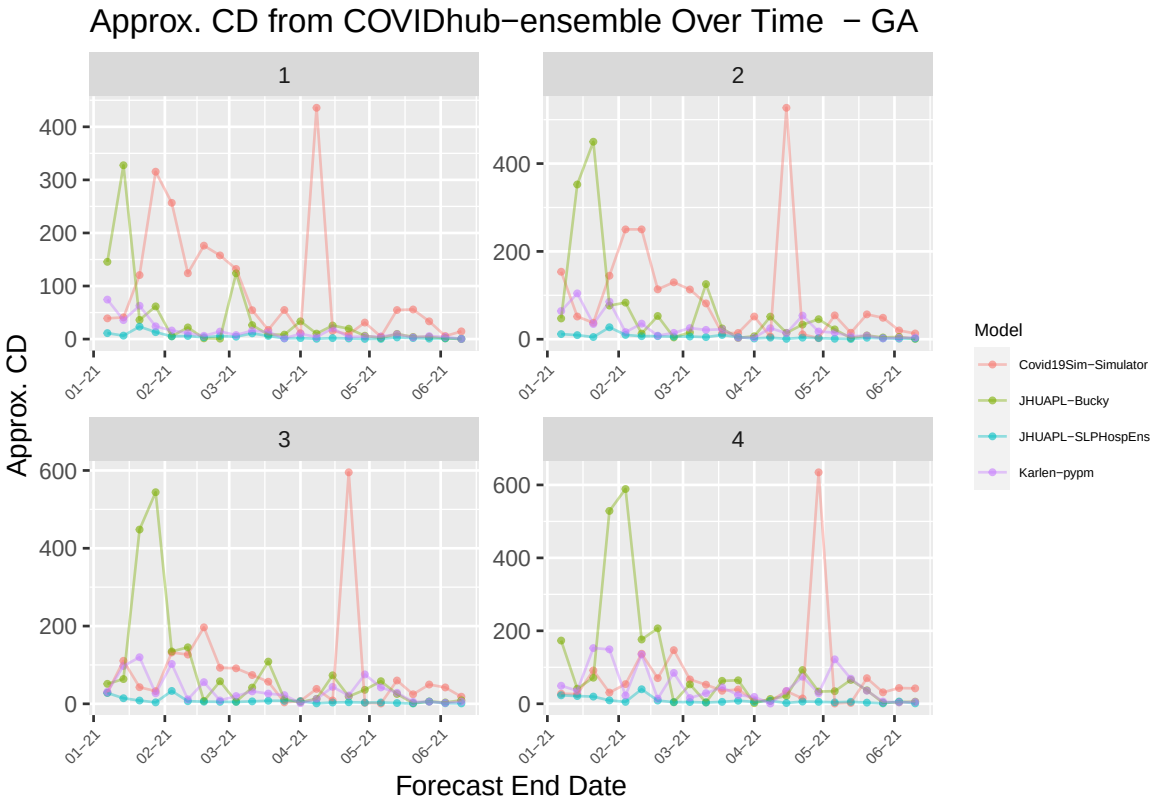

GA− Mean Approx. CD − Inc Hosp Forecasts by Horizon
1
2
3
4
COVIDhub−ensemble
JHUAPL−SLPHospEns
Karlen−pypm
Covid19Sim−Simulator
JHUAPL−Bucky
distance
200
150
100
50
0

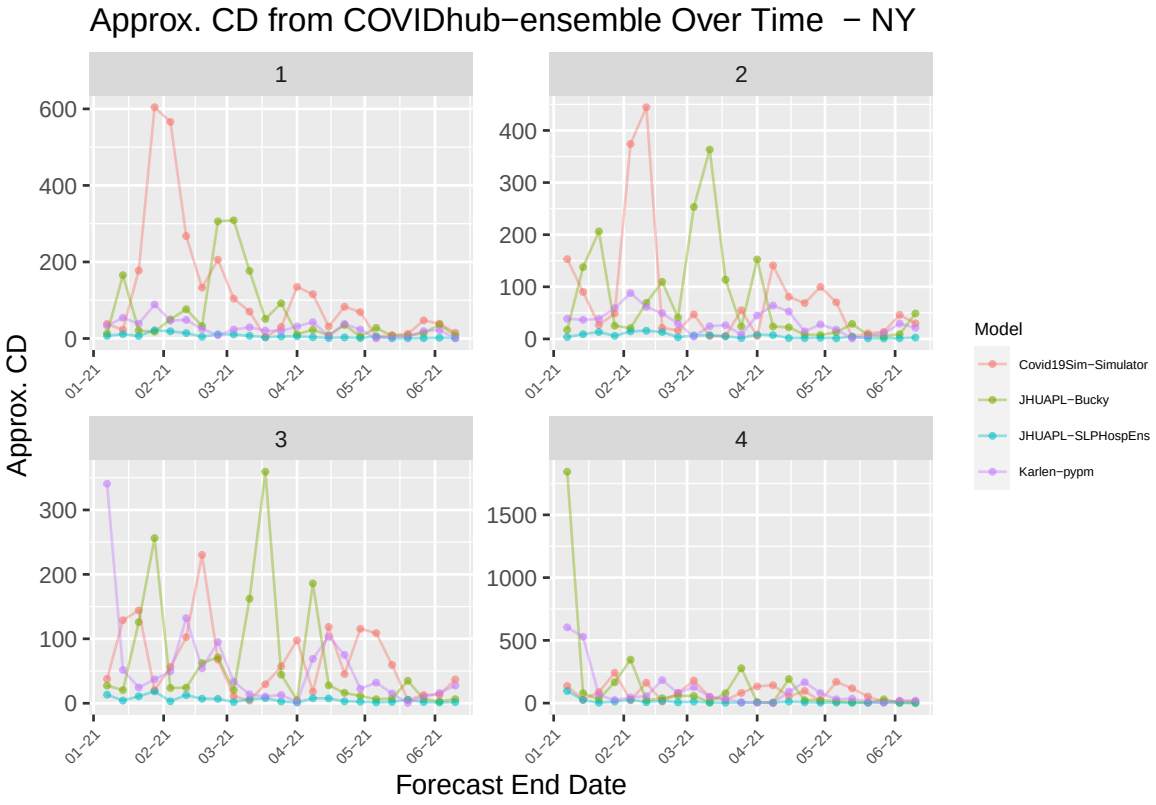


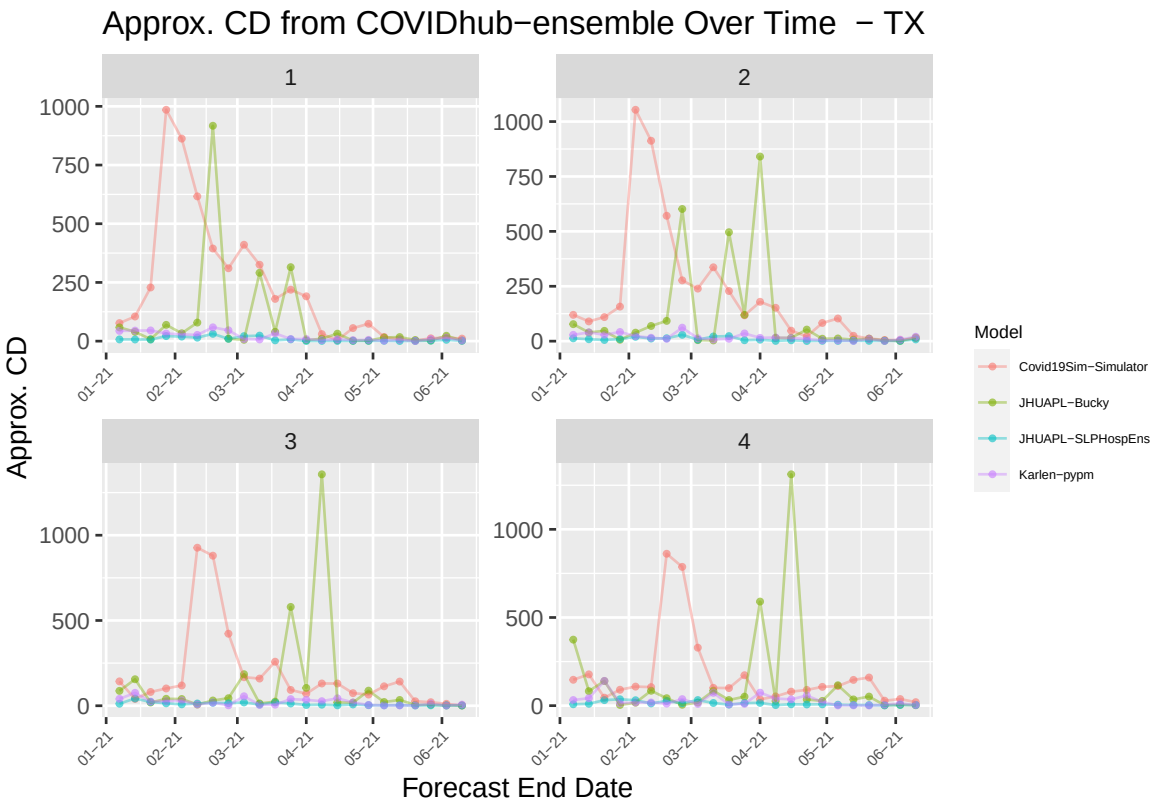
We can also look at the approximated pairwise distances to see how the models become more similar or

dissimilar over time.

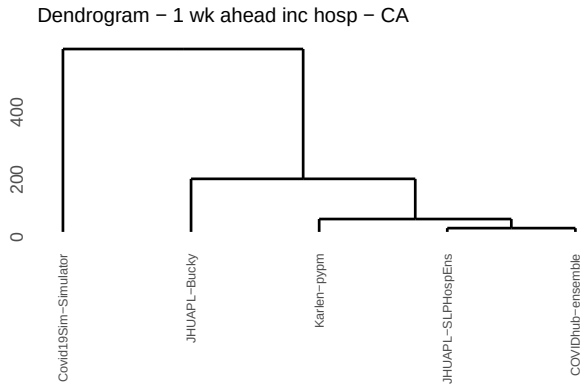


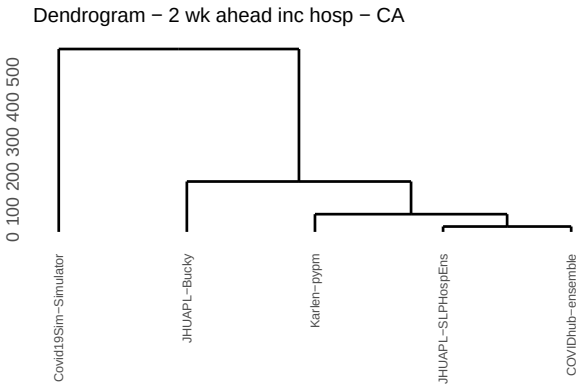

Approx. CD from COVIDhub−ensemble Over Time − GA
1
2
3
4
Approx. CD
Forecast End Date
01−21
02−21
03−21
04−21
05−21
06−21
Model
Covid19Sim−Simulator
JHUAPL−Bucky
JHUAPL−SLPHospEns
Karlen−pypm

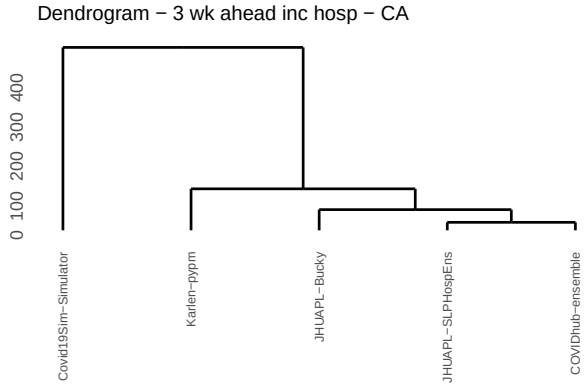

Approx. CD from COVIDhub−ensemble Over Time − NY
1
2
3
4
Approx. CD
Forecast End Date
01−21
02−21
03−21
04−21
05−21
06−21
Model
Covid19Sim−Simulator
JHUAPL−Bucky
JHUAPL−SLPHospEns
Karlen−pypm

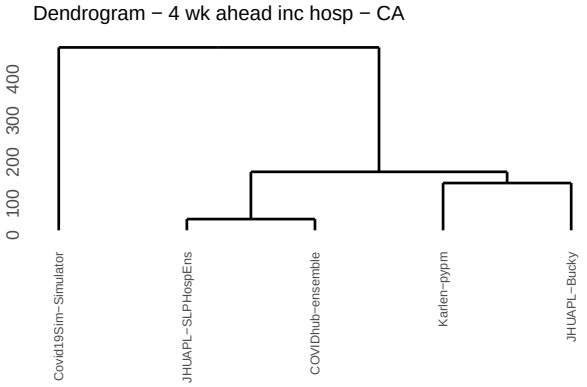


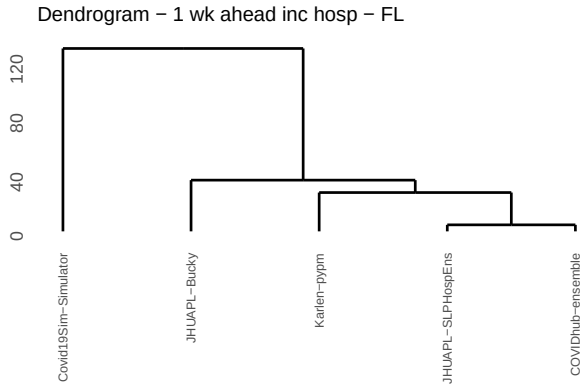
We can cluster the distances using hierarchical clustering. Different linkages will result in different clusters, we probably should investigate more later.

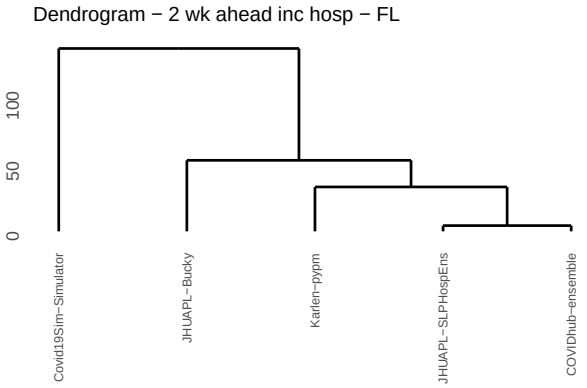

Dendrogram − 1 wk ahead inc hosp − CA
0
200
400
Covid19Sim−Simulator
JHUAPL−Bucky
Karlen−pypm
JHUAPL−SLPHospEns
COVIDhub−ensemble

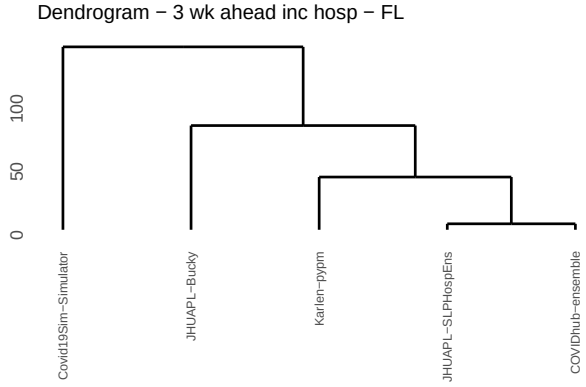

Dendrogram − 2 wk ahead inc hosp − CA
0
100
200
300
400
500
Covid19Sim−Simulator
JHUAPL−Bucky
Karlen−pypm
JHUAPL−SLPHospEns
COVIDhub−ensemble

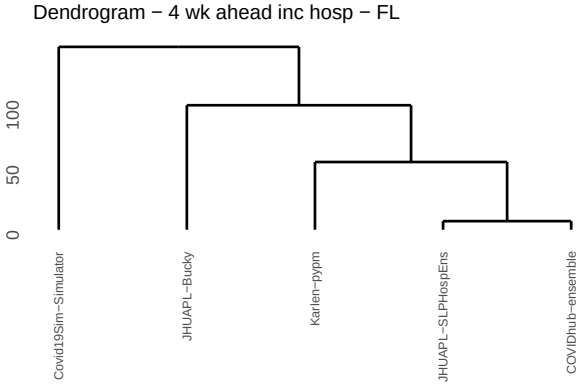

Dendrogram − 3 wk ahead inc hosp − CA
0
100
200
300
400
Covid19Sim−Simulator
Karlen−pypm
JHUAPL−Bucky
JHUAPL−SLPHospEns
COVIDhub−ensemble

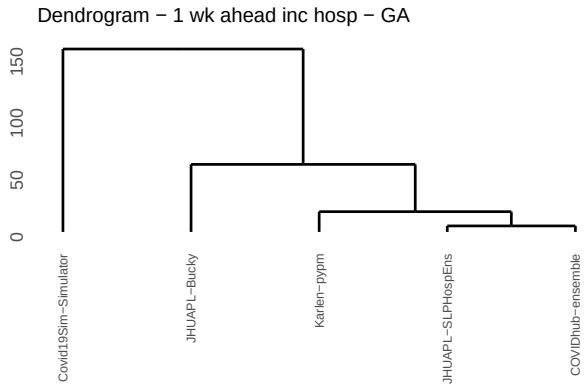

Dendrogram − 4 wk ahead inc hosp − CA
0
100
200
300
400
Covid19Sim−Simulator
JHUAPL−SLPHospEns
COVIDhub−ensemble
Karlen−pypm
JHUAPL−Bucky

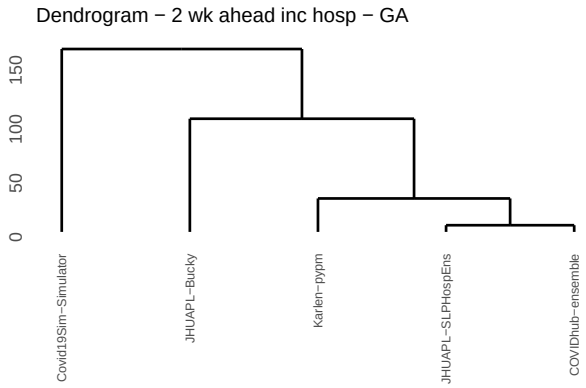

Dendrogram − 1 wk ahead inc hosp − FL
0
40
80
120
Covid19Sim−Simulator
JHUAPL−Bucky
Karlen−pypm
JHUAPL−SLPHospEns
COVIDhub−ensemble

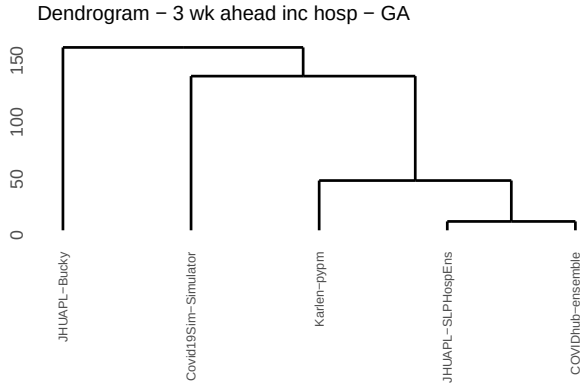

Dendrogram − 2 wk ahead inc hosp − FL
0
50
100
Covid19Sim−Simulator
JHUAPL−Bucky
Karlen−pypm
JHUAPL−SLPHospEns
COVIDhub−ensemble

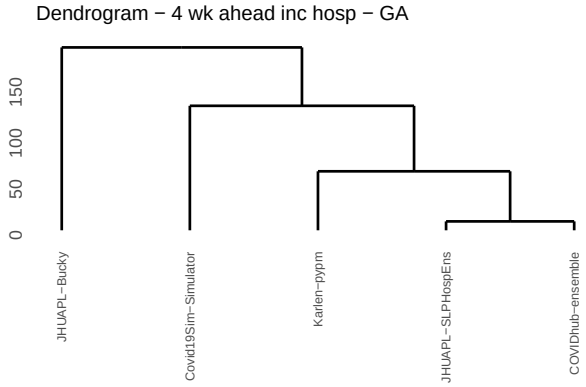

Dendrogram − 3 wk ahead inc hosp − FL
0
50
100
Covid19Sim−Simulator
JHUAPL−Bucky
Karlen−pypm
JHUAPL−SLPHospEns
COVIDhub−ensemble

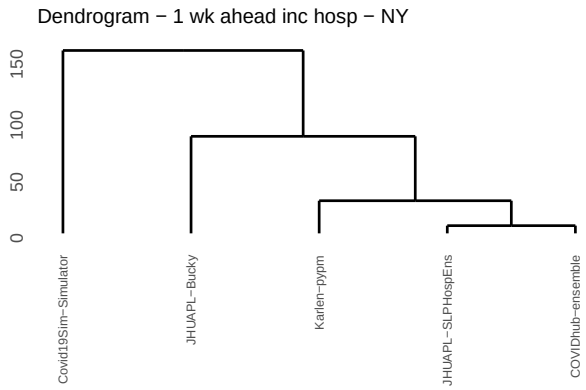

Dendrogram − 4 wk ahead inc hosp − FL
0
50
100
Covid19Sim−Simulator
JHUAPL−Bucky
Karlen−pypm
JHUAPL−SLPHospEns
COVIDhub−ensemble

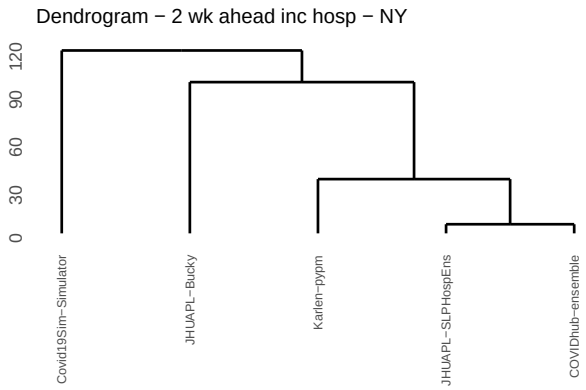

Dendrogram − 1 wk ahead inc hosp − GA
0
50
100
150
Covid19Sim−Simulator
JHUAPL−Bucky
Karlen−pypm
JHUAPL−SLPHospEns
COVIDhub−ensemble

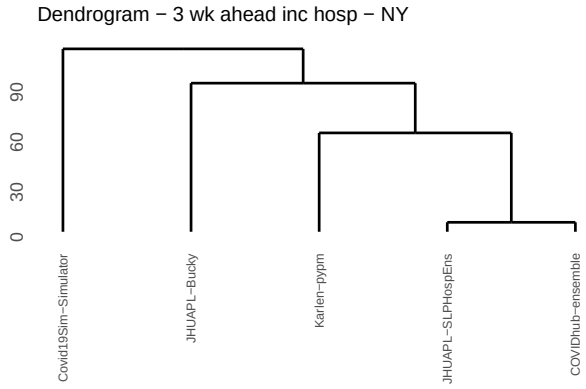

Dendrogram − 2 wk ahead inc hosp − GA
0
50
100
150
Covid19Sim−Simulator
JHUAPL−Bucky
Karlen−pypm
JHUAPL−SLPHospEns
COVIDhub−ensemble

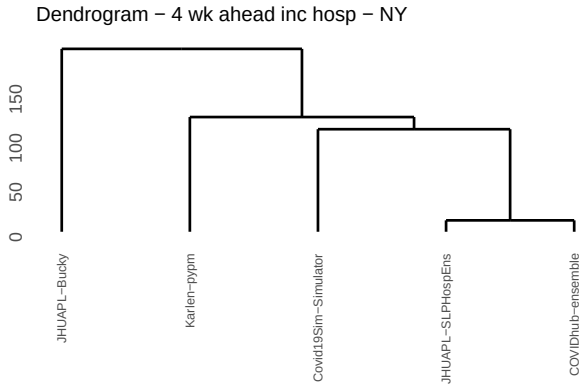

Dendrogram − 3 wk ahead inc hosp − GA
0
50
100
150
JHUAPL−Bucky
Covid19Sim−Simulator
Karlen−pypm
JHUAPL−SLPHospEns
COVIDhub−ensemble

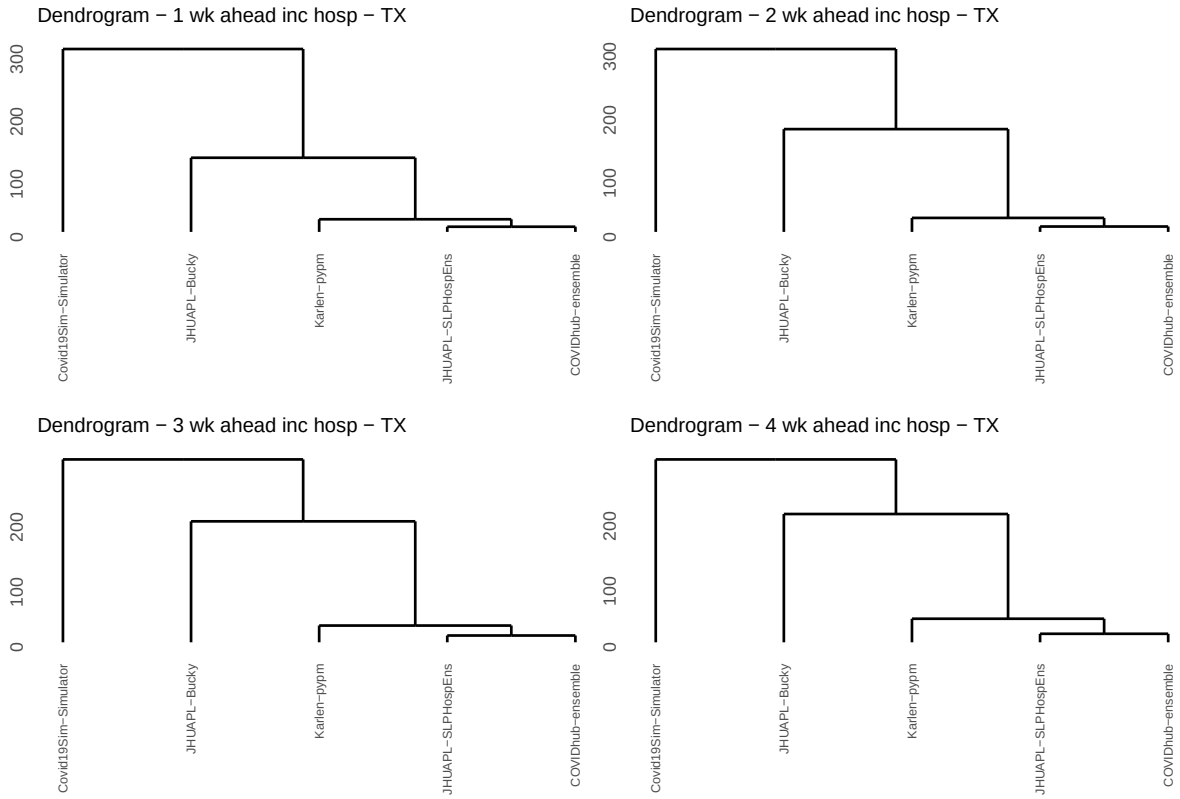

Dendrogram − 4 wk ahead inc hosp − GA
0
50
100
150
JHUAPL−Bucky
Covid19Sim−Simulator
Karlen−pypm
JHUAPL−SLPHospEns
COVIDhub−ensemble


Dendrogram − 1 wk ahead inc hosp − NY
0
50
100
150
Covid19Sim−Simulator
JHUAPL−Bucky
Karlen−pypm
JHUAPL−SLPHospEns
COVIDhub−ensemble


Dendrogram − 2 wk ahead inc hosp − NY
0
30
60
90
120
Covid19Sim−Simulator
JHUAPL−Bucky
Karlen−pypm
JHUAPL−SLPHospEns
COVIDhub−ensemble


Dendrogram − 3 wk ahead inc hosp − NY
0
30
60
90
Covid19Sim−Simulator
JHUAPL−Bucky
Karlen−pypm
JHUAPL−SLPHospEns
COVIDhub−ensemble


Dendrogram − 4 wk ahead inc hosp − NY
0
50
100
150
JHUAPL−Bucky
Karlen−pypm
Covid19Sim−Simulator
JHUAPL−SLPHospEns
COVIDhub−ensemble

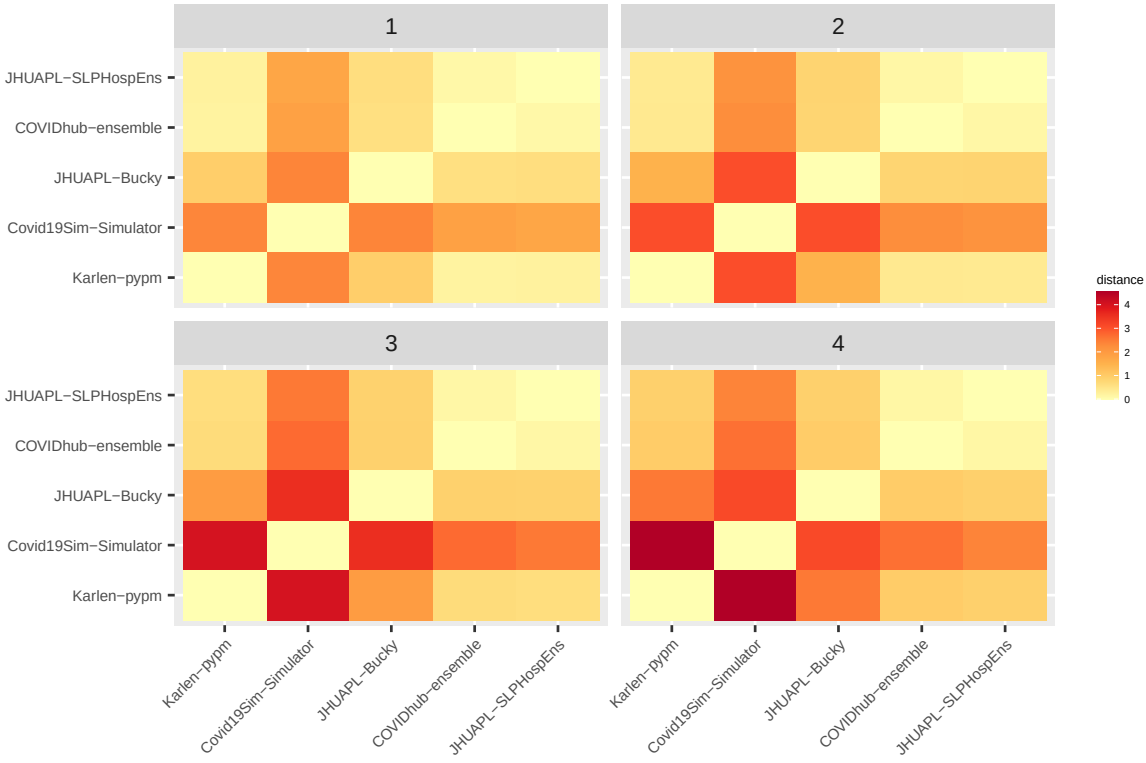


5 locations with the lowest number of COVID-19 deaths by the end of February 2021

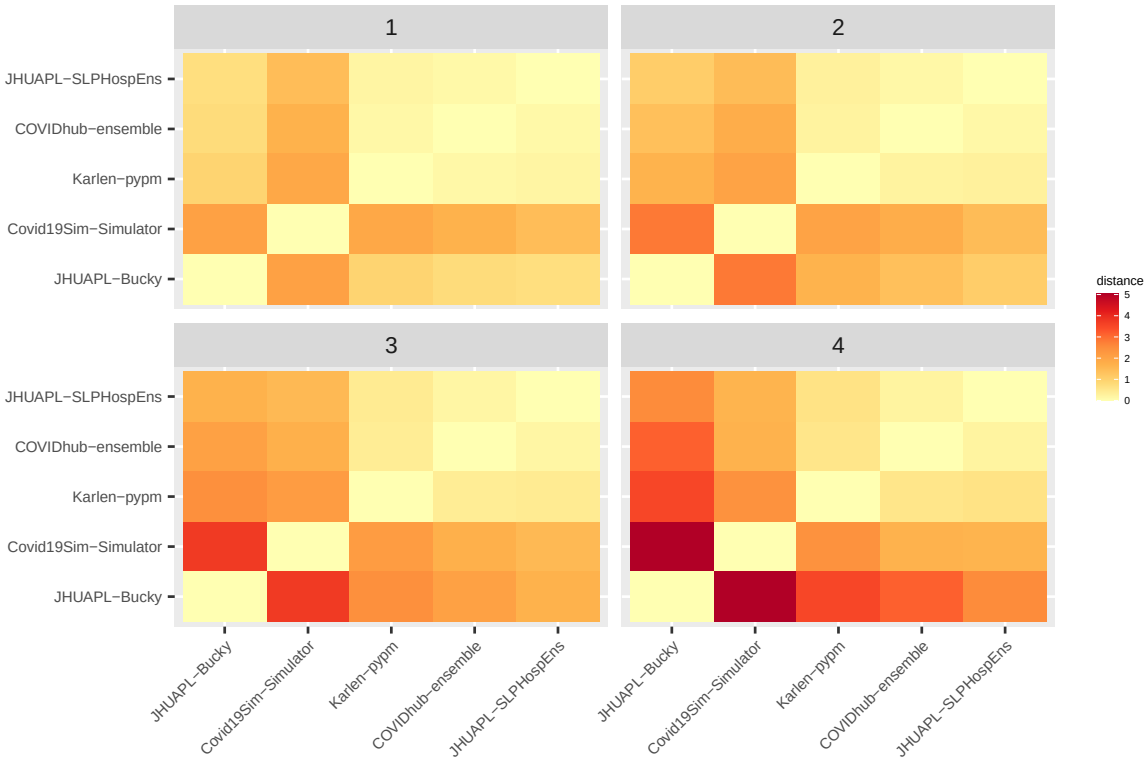
The pairwise approximated Cramer's distances are calculated for the models that have complete submissions for all target, all 5 locations with the lowest number of COVID-19 deaths by the end of February 2021, all probability levels, from the target end date of December, 17th 2020 to June 10th, 2021 .

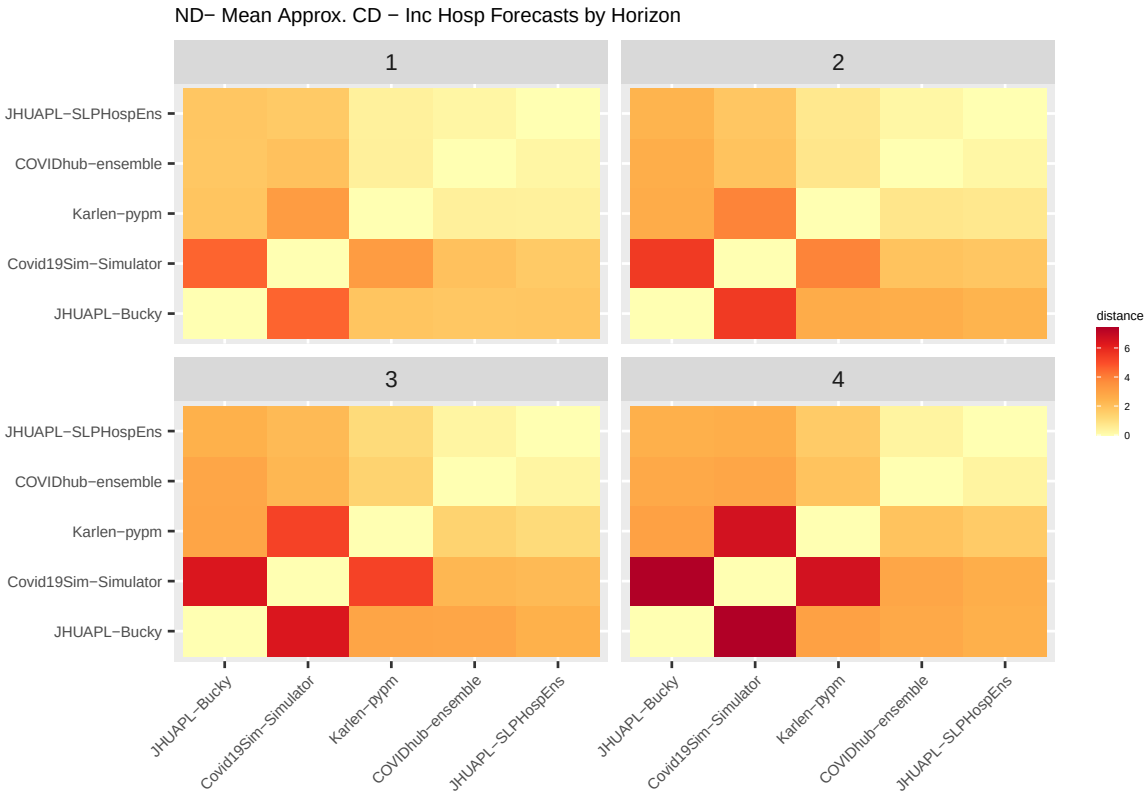
We can visualize the mean approximated pairwise distances across all time points in a heat map shown below. The distance from the model to itself is zero. The $x-$axis is arranged based in an ascending order of the model's approximate pairwise distance from the COVIDhub-ensemble. So, the first model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

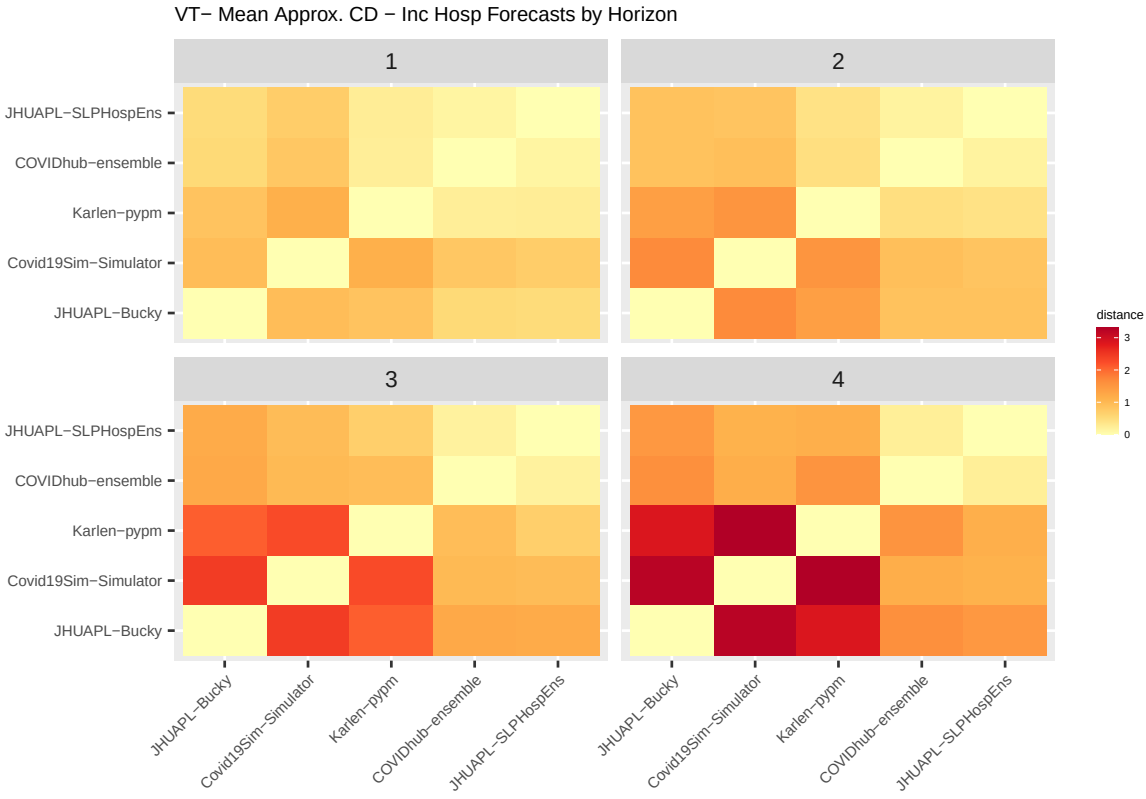
AK− Mean Approx. CD − Inc Hosp Forecasts by Horizon

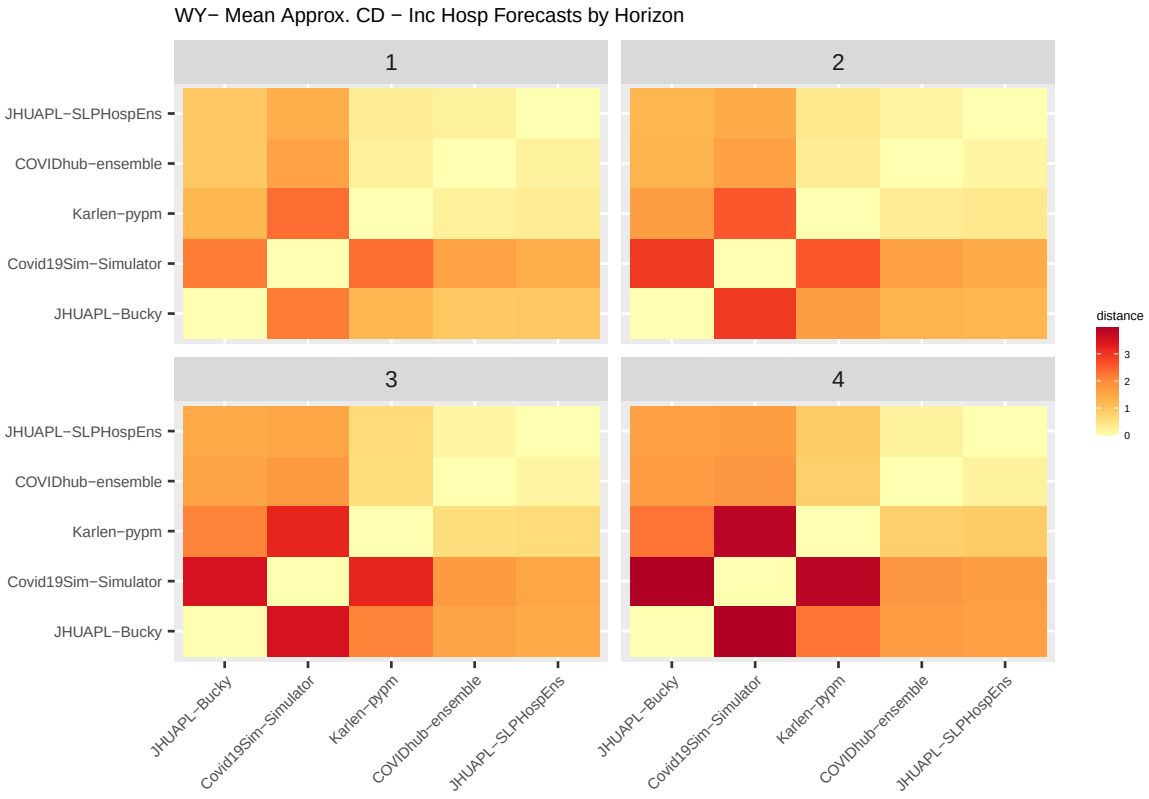


HI− Mean Approx. CD − Inc Hosp Forecasts by Horizon

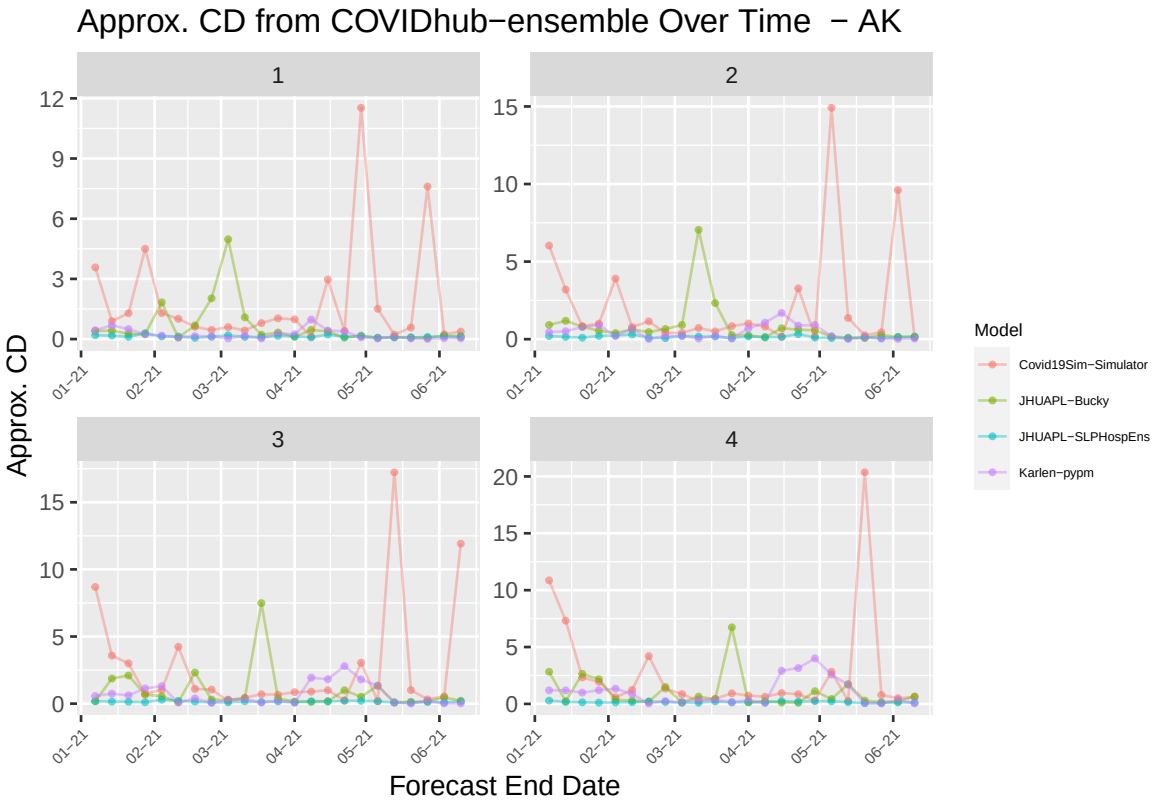


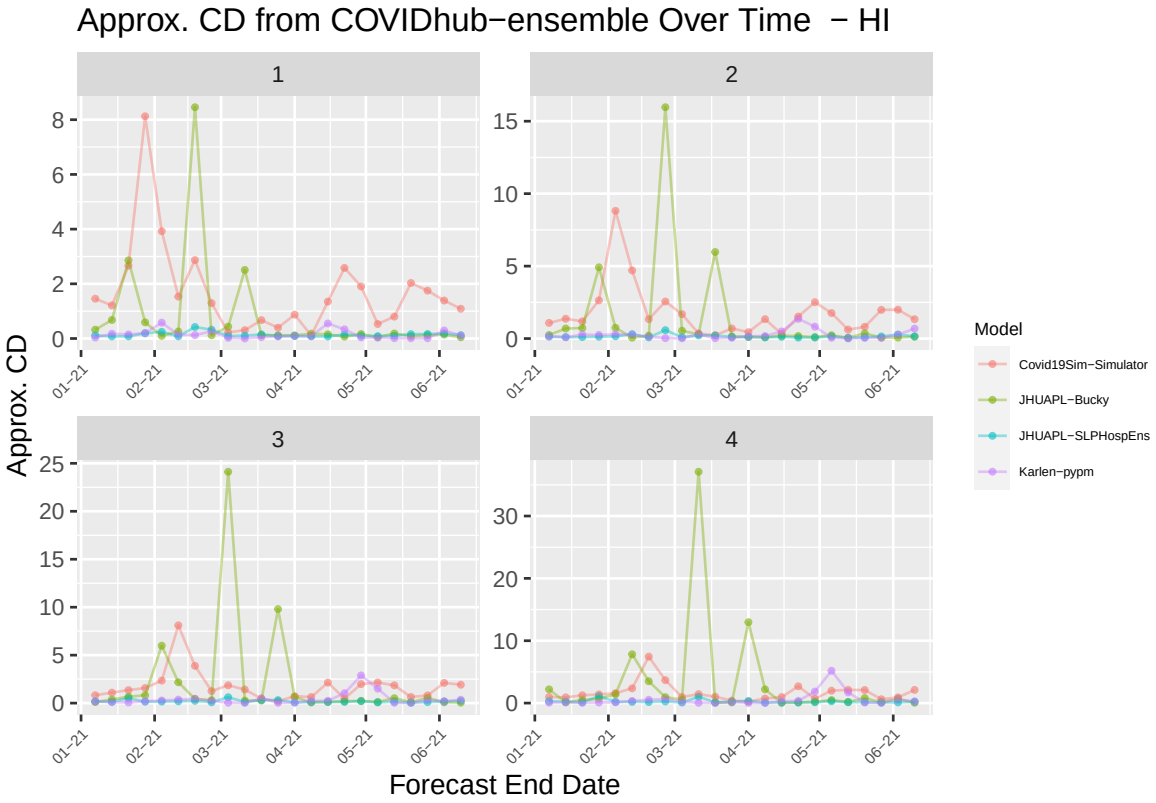

ND− Mean Approx. CD − Inc Hosp Forecasts by Horizon
1
2
3
4
JHUAPL−SLPHospEns
COVIDhub−ensemble
Karlen−pypm
Covid19Sim−Simulator
JHUAPL−Bucky
distance
6
4
2
0

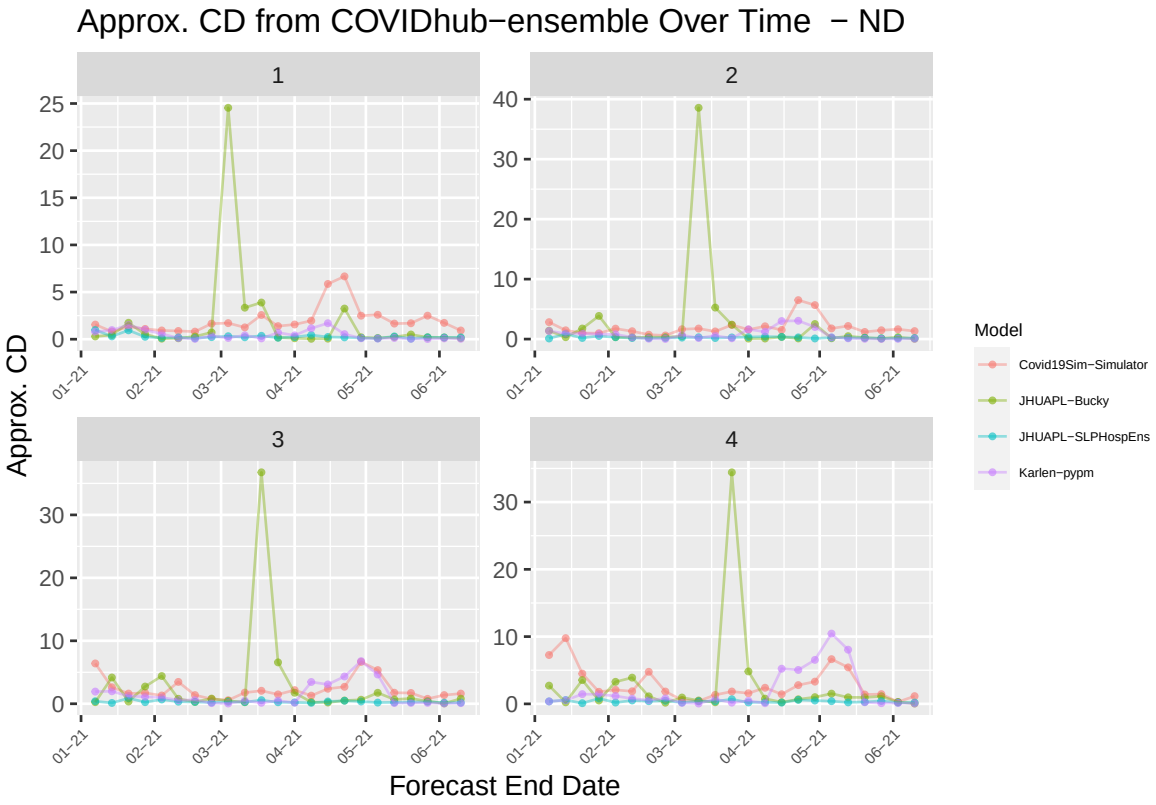

VT− Mean Approx. CD − Inc Hosp Forecasts by Horizon
1
2
3
4
JHUAPL−SLPHospEns
COVIDhub−ensemble
Karlen−pypm
Covid19Sim−Simulator
JHUAPL−Bucky
distance
3
2
1
0

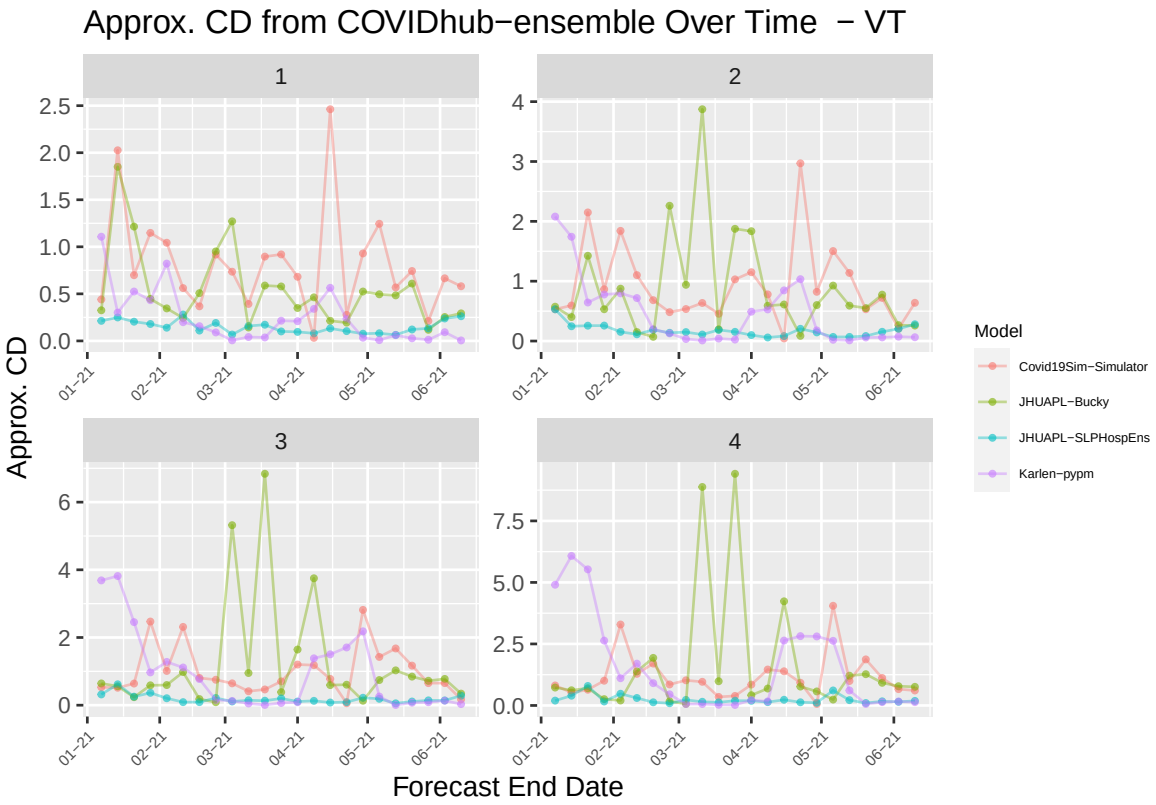


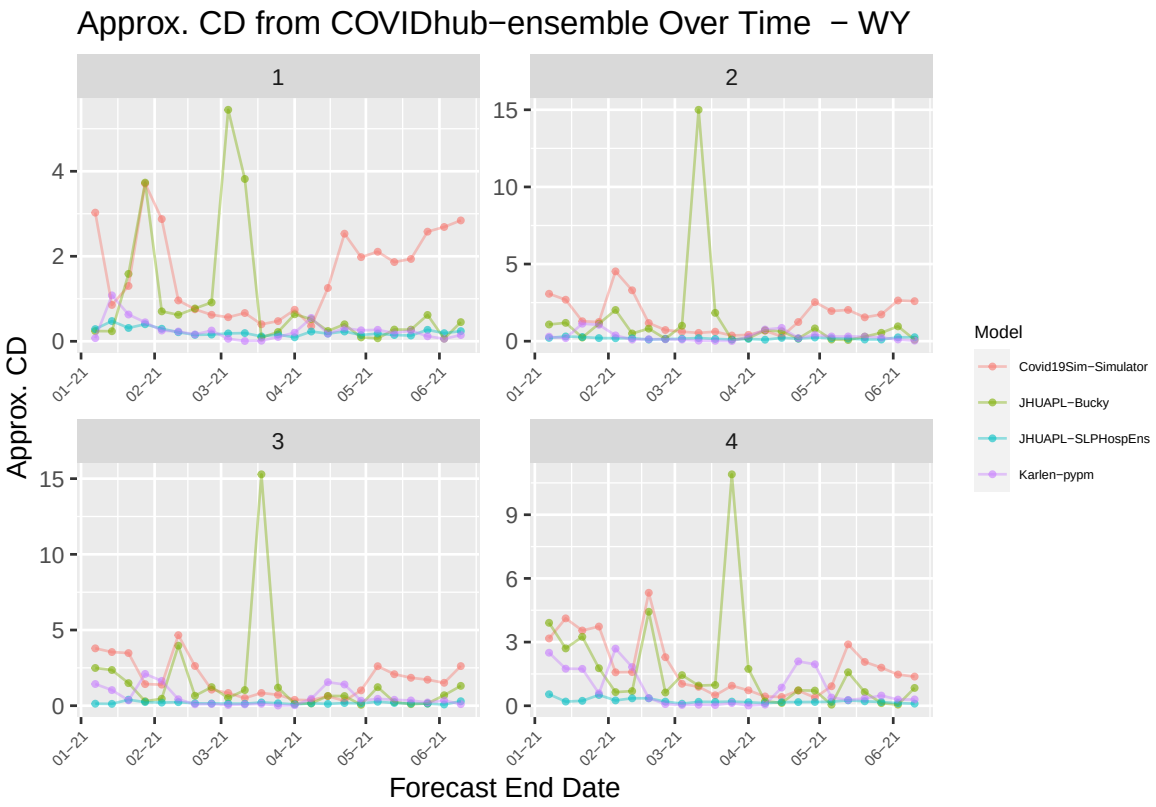
We can also look at the approximated pairwise distances to see how the models become more similar or dissimilar over time.


Approx. CD from COVIDhub−ensemble Over Time − AK
1
2
3
4
Approx. CD
Forecast End Date
01−21
02−21
03−21
04−21
05−21
06−21
Model
Covid19Sim−Simulator
JHUAPL−Bucky
JHUAPL−SLPHospEns
Karlen−pypm

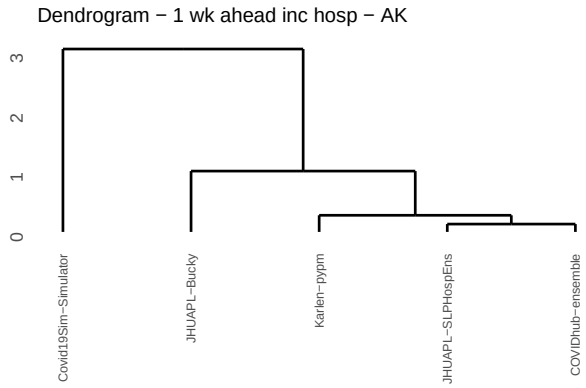

Approx. CD from COVIDhub−ensemble Over Time − HI
1
2
3
4
Approx. CD
Forecast End Date
01−21
02−21
03−21
04−21
05−21
06−21
Model
Covid19Sim−Simulator
JHUAPL−Bucky
JHUAPL−SLPHospEns
Karlen−pypm

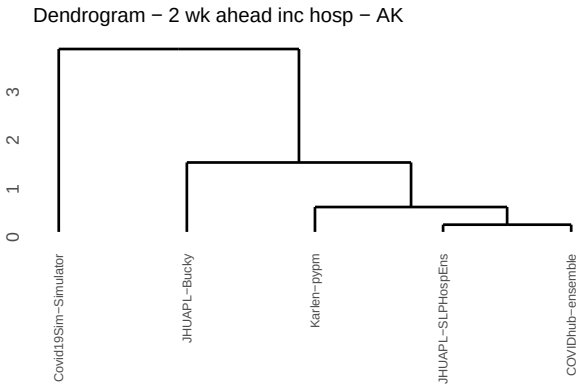

Approx. CD from COVIDhub−ensemble Over Time − ND
1
2
3
4
Approx. CD
Forecast End Date
01−21
02−21
03−21
04−21
05−21
06−21
Model
Covid19Sim−Simulator
JHUAPL−Bucky
JHUAPL−SLPHospEns
Karlen−pypm

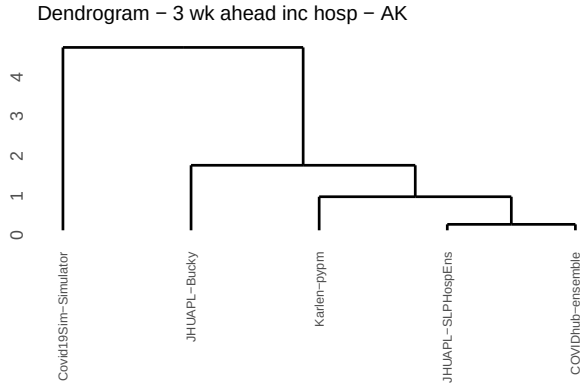

Approx. CD from COVIDhub−ensemble Over Time − VT
1
2
3
4
Approx. CD
Forecast End Date
01−21
02−21
03−21
04−21
05−21
06−21
Model
Covid19Sim−Simulator
JHUAPL−Bucky
JHUAPL−SLPHospEns
Karlen−pypm

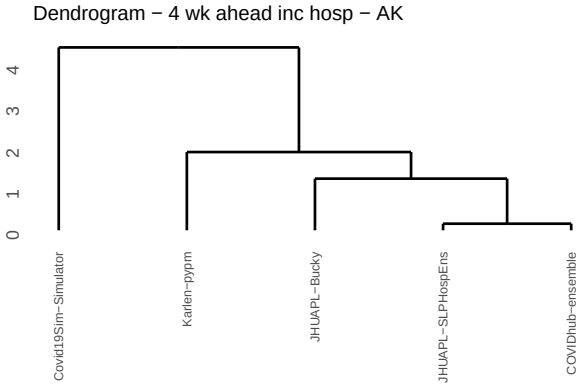


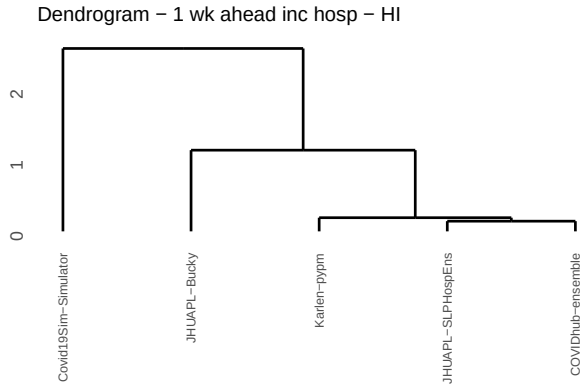
We can cluster the distances using hierarchical clustering. Different linkages will result in different clusters, we probably should investigate more later.

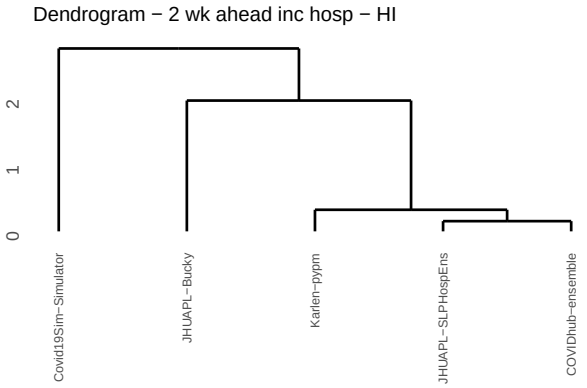

Dendrogram − 1 wk ahead inc hosp − AK
0
1
2
3
Covid19Sim−Simulator
JHUAPL−Bucky
Karlen−pypm
JHUAPL−SLPHospEns
COVIDhub−ensemble

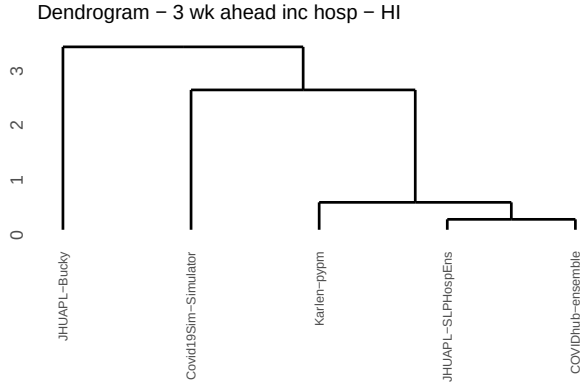

Dendrogram − 2 wk ahead inc hosp − AK
0
1
2
3
Covid19Sim−Simulator
JHUAPL−Bucky
Karlen−pypm
JHUAPL−SLPHospEns
COVIDhub−ensemble

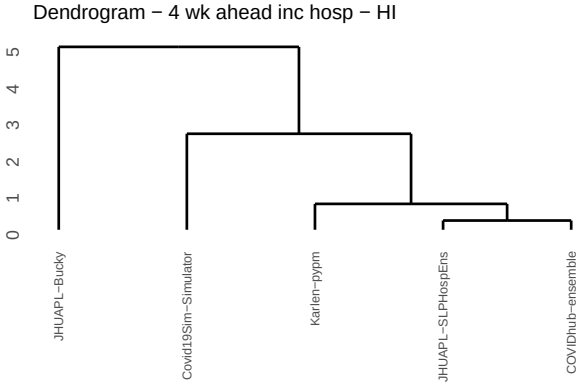

Dendrogram − 3 wk ahead inc hosp − AK
0
1
2
3
4
Covid19Sim−Simulator
JHUAPL−Bucky
Karlen−pypm
JHUAPL−SLPHospEns
COVIDhub−ensemble

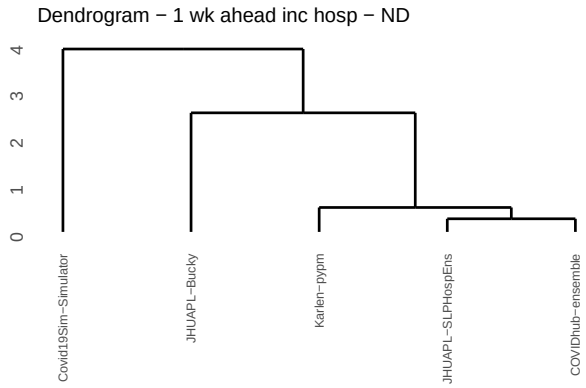

Dendrogram − 4 wk ahead inc hosp − AK
0
1
2
3
4
Covid19Sim−Simulator
Karlen−pypm
JHUAPL−Bucky
JHUAPL−SLPHospEns
COVIDhub−ensemble

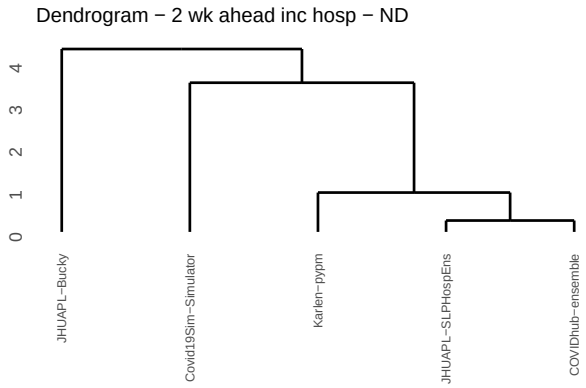

Dendrogram − 1 wk ahead inc hosp − HI
0
1
2
Covid19Sim−Simulator
JHUAPL−Bucky
Karlen−pypm
JHUAPL−SLPHospEns
COVIDhub−ensemble

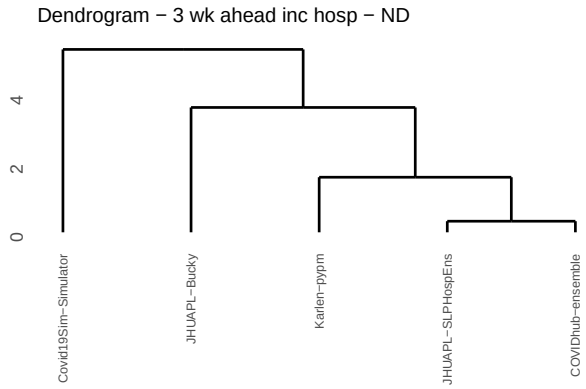

Dendrogram − 2 wk ahead inc hosp − HI
0
1
2
Covid19Sim−Simulator
JHUAPL−Bucky
Karlen−pypm
JHUAPL−SLPHospEns
COVIDhub−ensemble

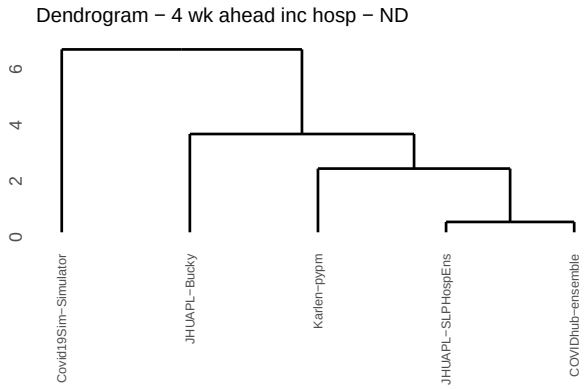

Dendrogram − 3 wk ahead inc hosp − HI
0
1
2
3
JHUAPL−Bucky
Covid19Sim−Simulator
Karlen−pypm
JHUAPL−SLPHospEns
COVIDhub−ensemble

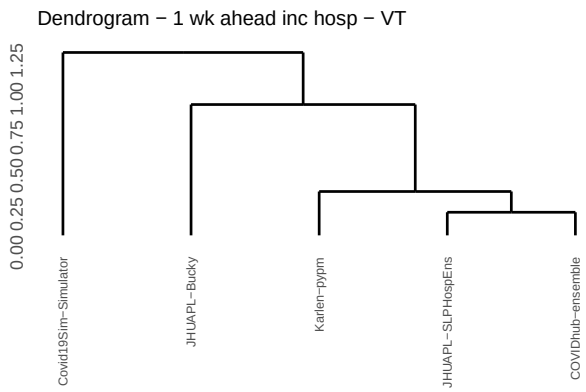

Dendrogram − 4 wk ahead inc hosp − HI
0
1
2
3
4
5
JHUAPL−Bucky
Covid19Sim−Simulator
Karlen−pypm
JHUAPL−SLPHospEns
COVIDhub−ensemble

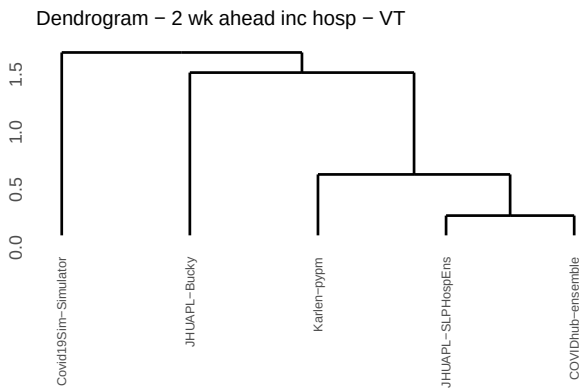

Dendrogram − 1 wk ahead inc hosp − ND
0
1
2
3
4
Covid19Sim−Simulator
JHUAPL−Bucky
Karlen−pypm
JHUAPL−SLPHospEns
COVIDhub−ensemble

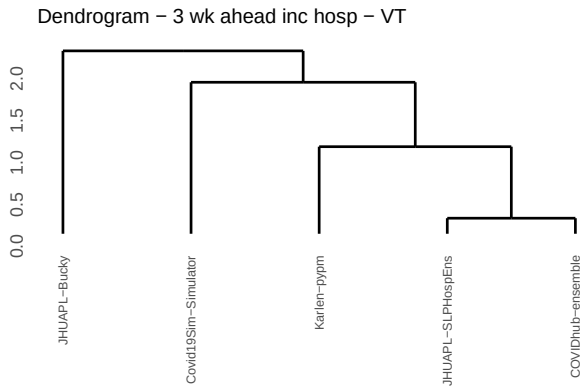

Dendrogram − 2 wk ahead inc hosp − ND
0
1
2
3
4
JHUAPL−Bucky
Covid19Sim−Simulator
Karlen−pypm
JHUAPL−SLPHospEns
COVIDhub−ensemble

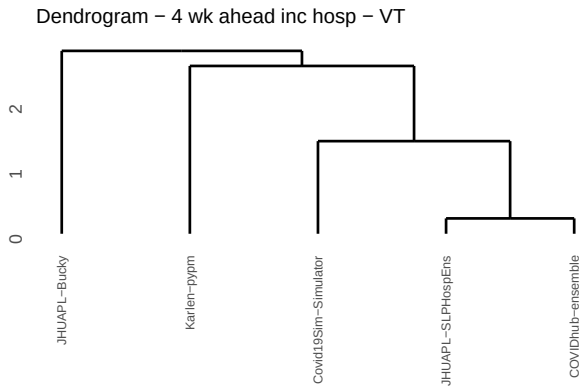

Dendrogram − 3 wk ahead inc hosp − ND
0
2
4
Covid19Sim−Simulator
JHUAPL−Bucky
Karlen−pypm
JHUAPL−SLPHospEns
COVIDhub−ensemble

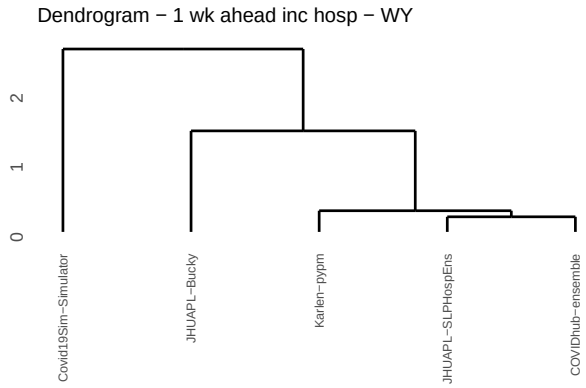

Dendrogram − 4 wk ahead inc hosp − ND
0
2
4
6
Covid19Sim−Simulator
JHUAPL−Bucky
Karlen−pypm
JHUAPL−SLPHospEns
COVIDhub−ensemble

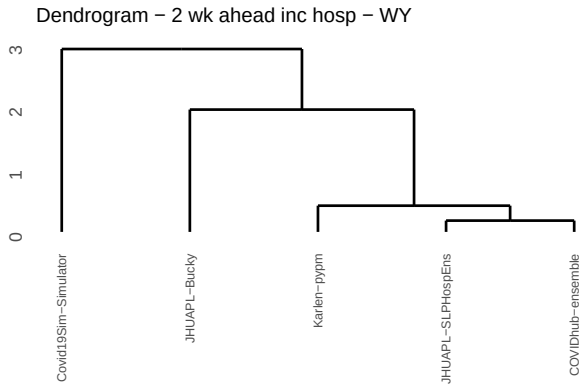

Dendrogram − 1 wk ahead inc hosp − VT
0.00
0.25
0.50
0.75
1.00
1.25
Covid19Sim−Simulator
JHUAPL−Bucky
Karlen−pypm
JHUAPL−SLPHospEns
COVIDhub−ensemble

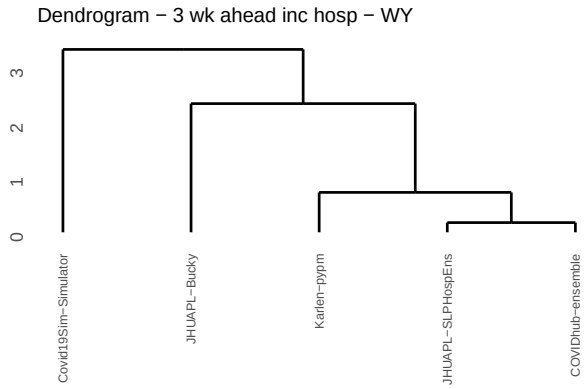

Dendrogram − 2 wk ahead inc hosp − VT
0.0
0.5
1.0
1.5
Covid19Sim−Simulator
JHUAPL−Bucky
Karlen−pypm
JHUAPL−SLPHospEns
COVIDhub−ensemble

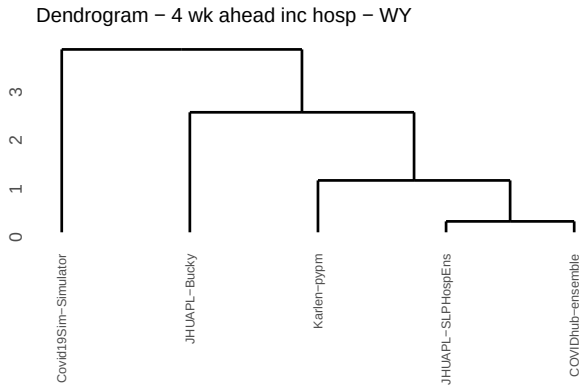

Dendrogram − 3 wk ahead inc hosp − VT
0.0
0.5
1.0
1.5
2.0
JHUAPL−Bucky
Covid19Sim−Simulator
Karlen−pypm
JHUAPL−SLPHospEns
COVIDhub−ensemble


Dendrogram − 4 wk ahead inc hosp − VT
0
1
2
JHUAPL−Bucky
Karlen−pypm
Covid19Sim−Simulator
JHUAPL−SLPHospEns
COVIDhub−ensemble


Dendrogram − 1 wk ahead inc hosp − WY
0
1
2
Covid19Sim−Simulator
JHUAPL−Bucky
Karlen−pypm
JHUAPL−SLPHospEns
COVIDhub−ensemble


Dendrogram − 2 wk ahead inc hosp − WY
0
1
2
3
Covid19Sim−Simulator
JHUAPL−Bucky
Karlen−pypm
JHUAPL−SLPHospEns
COVIDhub−ensemble


Dendrogram − 3 wk ahead inc hosp − WY
0
1
2
3
Covid19Sim−Simulator
JHUAPL−Bucky
Karlen−pypm
JHUAPL−SLPHospEns
COVIDhub−ensemble


Dendrogram − 4 wk ahead inc hosp − WY
0
1
2
3
Covid19Sim−Simulator
JHUAPL−Bucky
Karlen−pypm
JHUAPL−SLPHospEns
COVIDhub−ensemble