Multi-Model Ensemble Clustering at the Weather Prediction Center

Bill Lamberson[1,2] and Jim Nelson[2]

[1]Cooperative Institute for Research in Environmental Sciences (CIRES)

[2]Weather Prediction Center

UFS Users Workshop

28 July 2020

The Meteorological Problem

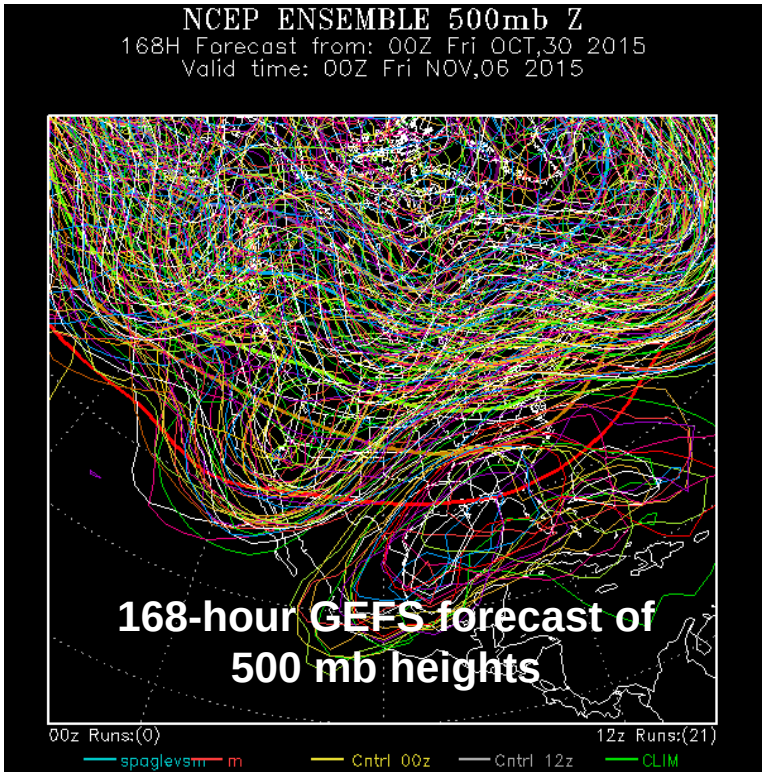
- **Problem:** The atmosphere is a chaotic system, so a way to represent the inherent uncertainty in forecasts, especially at longer lead times, is needed.

- **Solution:** Ensembles!

But Wait! We Still Have a Problem!

- **New Problem**: We exponentially increased the amount of data that we want people to look at when making a forecast but didn't necessarily give them more time to make the forecast.

- Consequently, ensembles are often only used in their rudimentary form (ensemble mean and spaghetti diagrams).



- **Potential solution: Ensemble clustering**

- Can provide forecasters with more ensemble information without data overload.

What is Ensemble Clustering

- A way to distill the ensemble down to the few forecast scenarios that are most prevalent among the ensemble's membership.

- Viewing ensemble clusters allows forecasters to quickly assess areas of uncertainty in the forecast and the different ways the forecast could play out without data overload.

- This is essential for forecaster situational awareness and the communication of the forecast.

- Ensemble clustering isn't new and there are many ways to do it.

- So far it has mostly been an academic exercise with little uptake by operational forecasters in the NWS.

Multi-Model Ensemble Clustering at WPC

- WPC began experimenting with ensemble clustering as a better way to utilize ensemble data during some of our Hydrometeorology Testbed Experiments.

- Over the past couple of years we have tinkered around but have settled on using using a variation of Fuzzy Clustering to create create clusters of ensemble members from the GEFS, GEPS, and EPS (a total of 90 members) for each day in days 3-7.

- These clusters are available to view on a website and have started to be used outside of WPC, particularly by Western and Central Region WFOs.

Zheng, M., E. K. Chang, B. A. Colle, Y. Luo, and Y. Zhu, 2017: Applying fuzzy clustering to a multimodel ensemble for U.S. East Coast winter storms: Scenario identification and forecast verification. *Wea. Forecasting*, **32**, 881–903, https://doi.org/10.1175/WAF-D-16-0112.1.

Fuzzy Clustering Recipe

1. Calculate and interpret the first two Empirical Orthogonal Functions (EOFs) of the 500-hPa height forecast over a region of the United States from the global ensembles (EPS, GEFS, and GEPS).

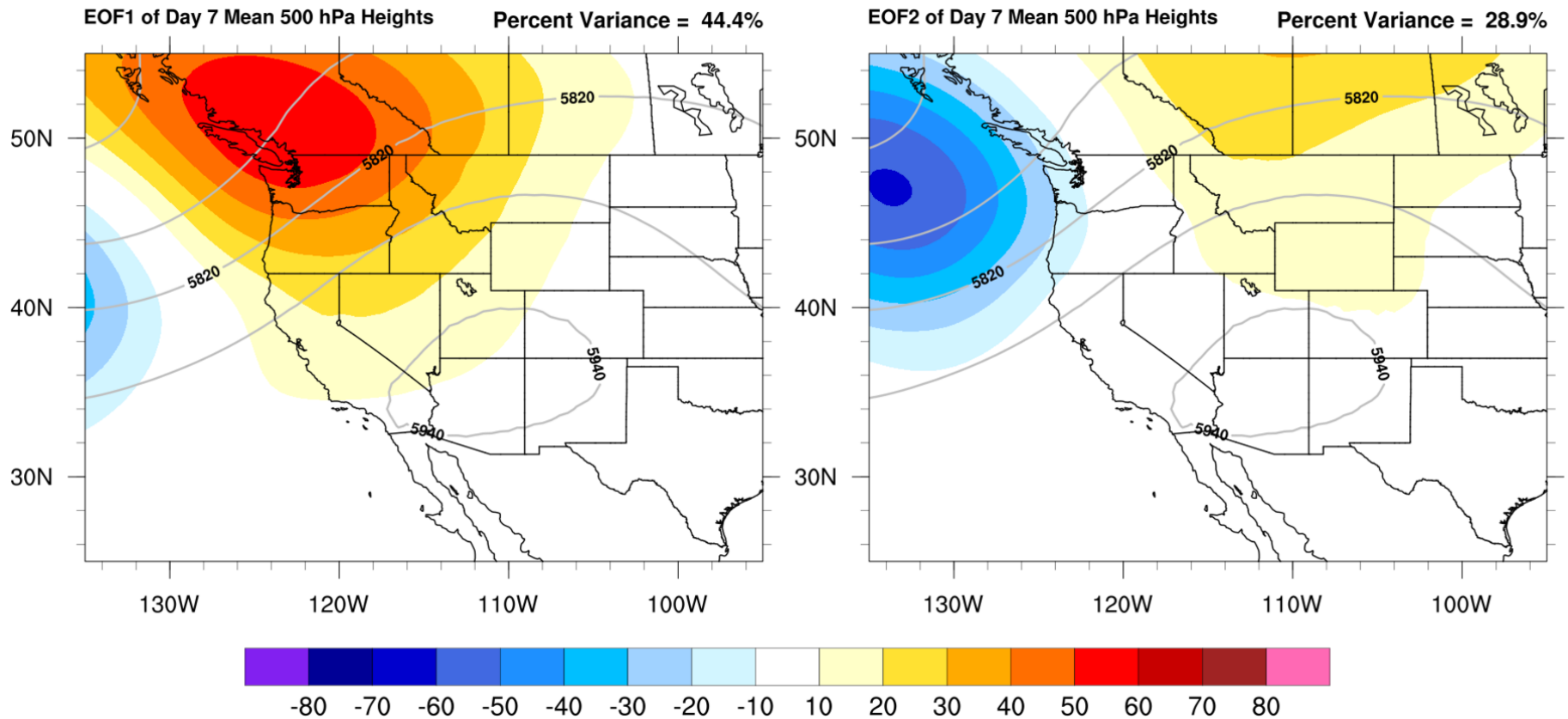
2. Use k-means clustering to create clusters of ensemble members based on their Principal Component (PCs) for the EOFs.

3. View the clusters.

Example: Let's apply fuzzy clustering to the 7-day forecasts of the weather for this Thursday over the Western United States

1. Calculate and Interpret First Two EOFs

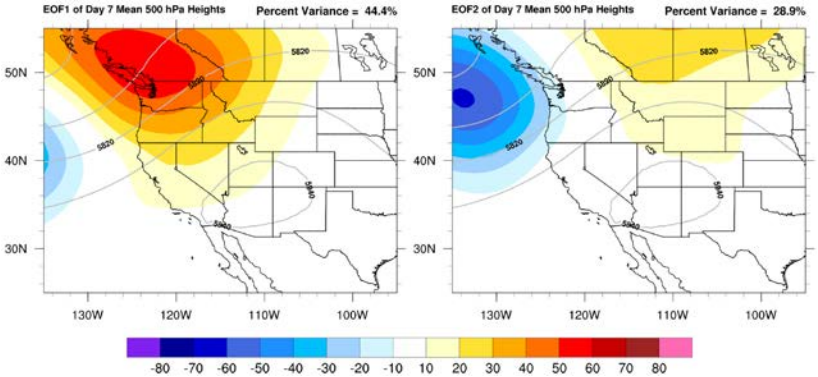
Init: 0000 UTC 23 Jul 2020
Valid: 0000 UTC 30 Jul – 0000 UTC 31 Jul

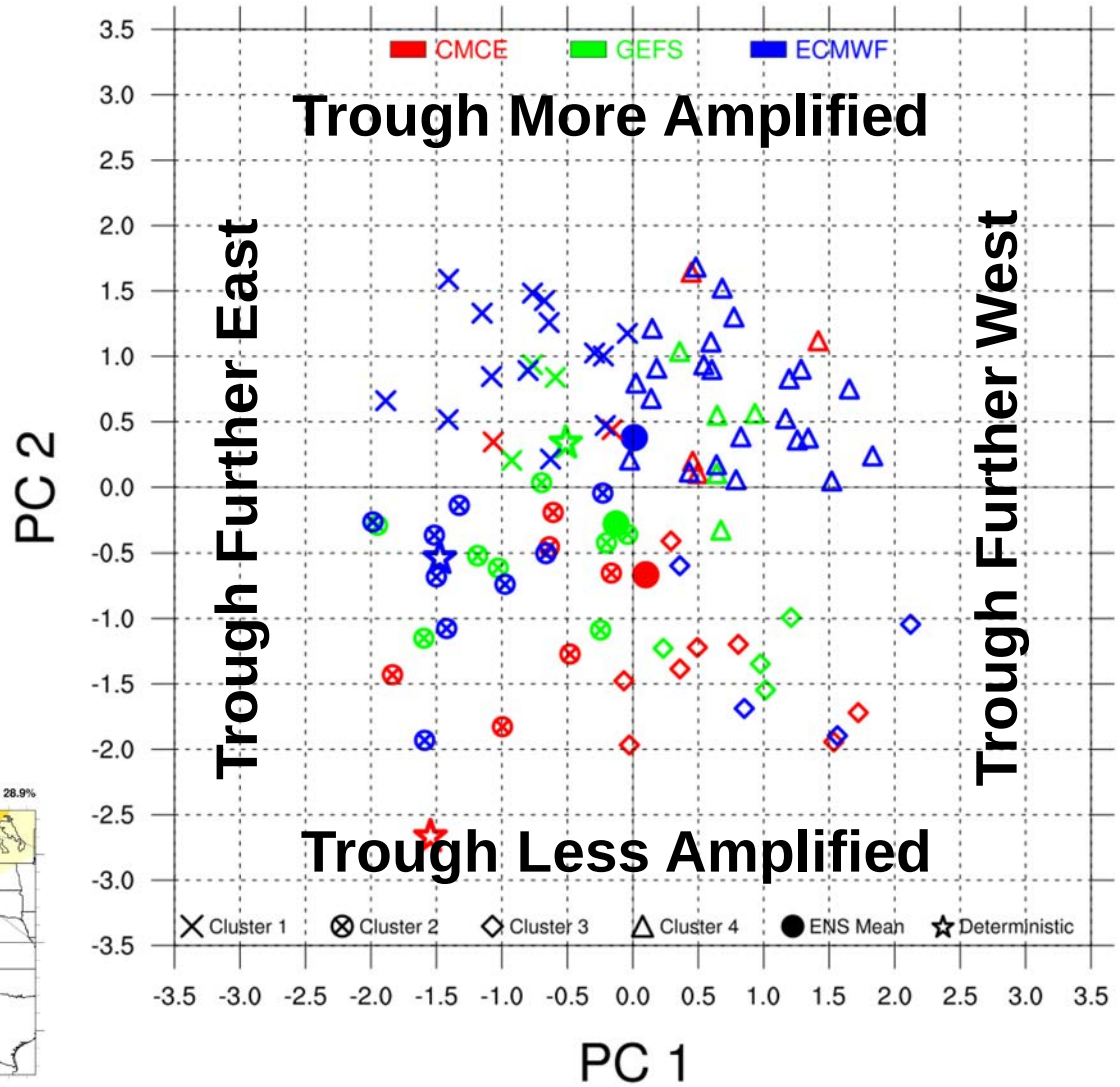


- EOF 1: Tells us that there is considerable (i.e., the most) uncertainty in the east-west position of the 500-hPa trough off the West Coast.
- EOF 2: Tells us that there is also uncertainty in the amplitude of the 500-hPa trough and the downstream ridge.

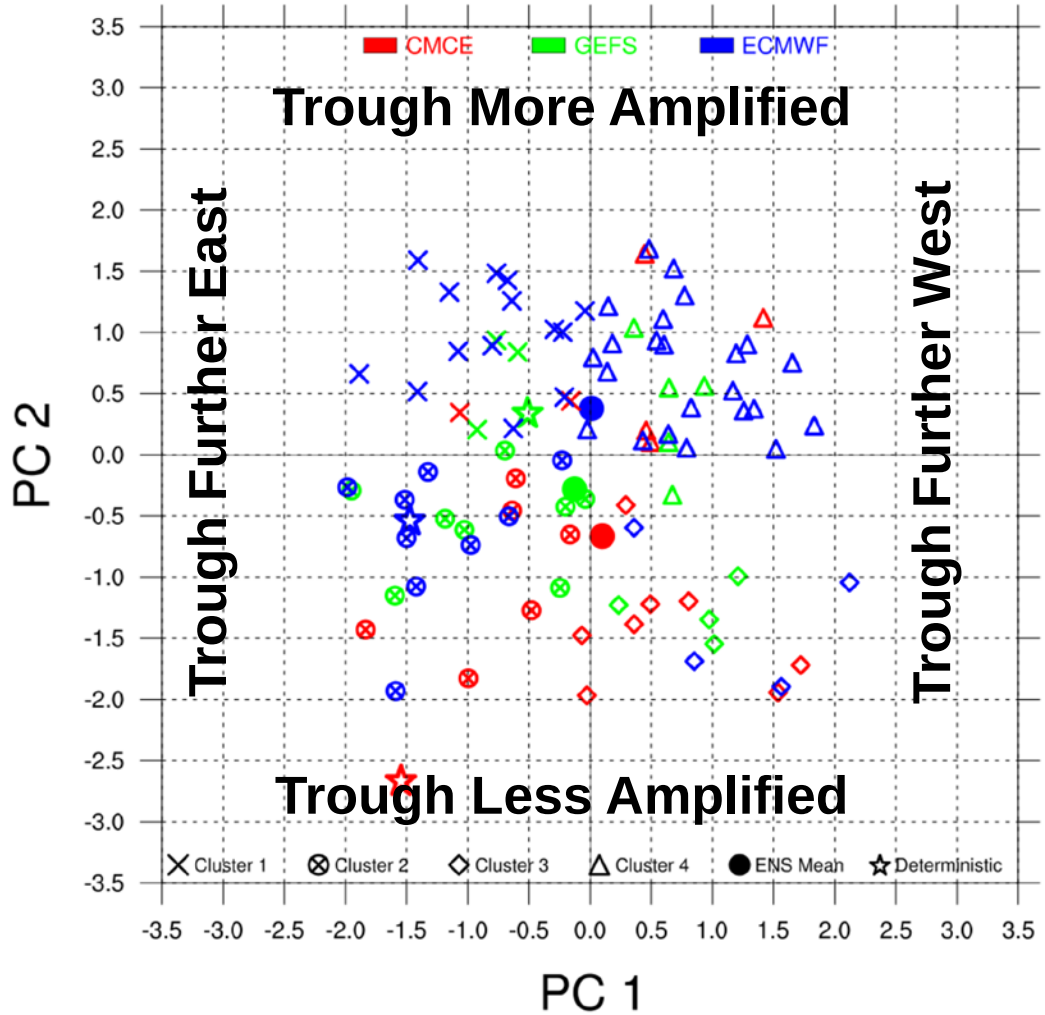
1. Calculate and Interpret First Two EOFs

- **EOF1: East-west differences in location of trough.**
- **EOF2: Differences in amplitude of trough and downstream ridge.**

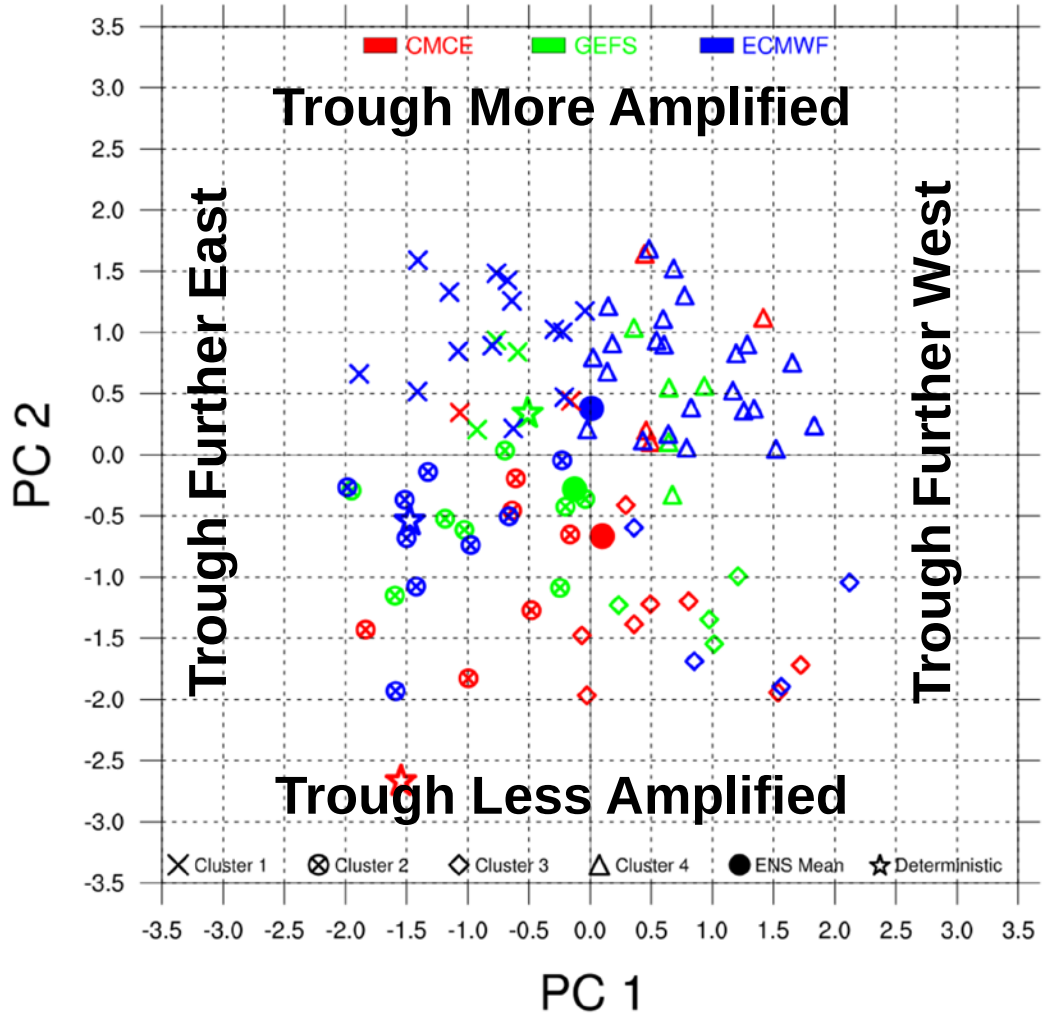




2. Use k-means Clustering

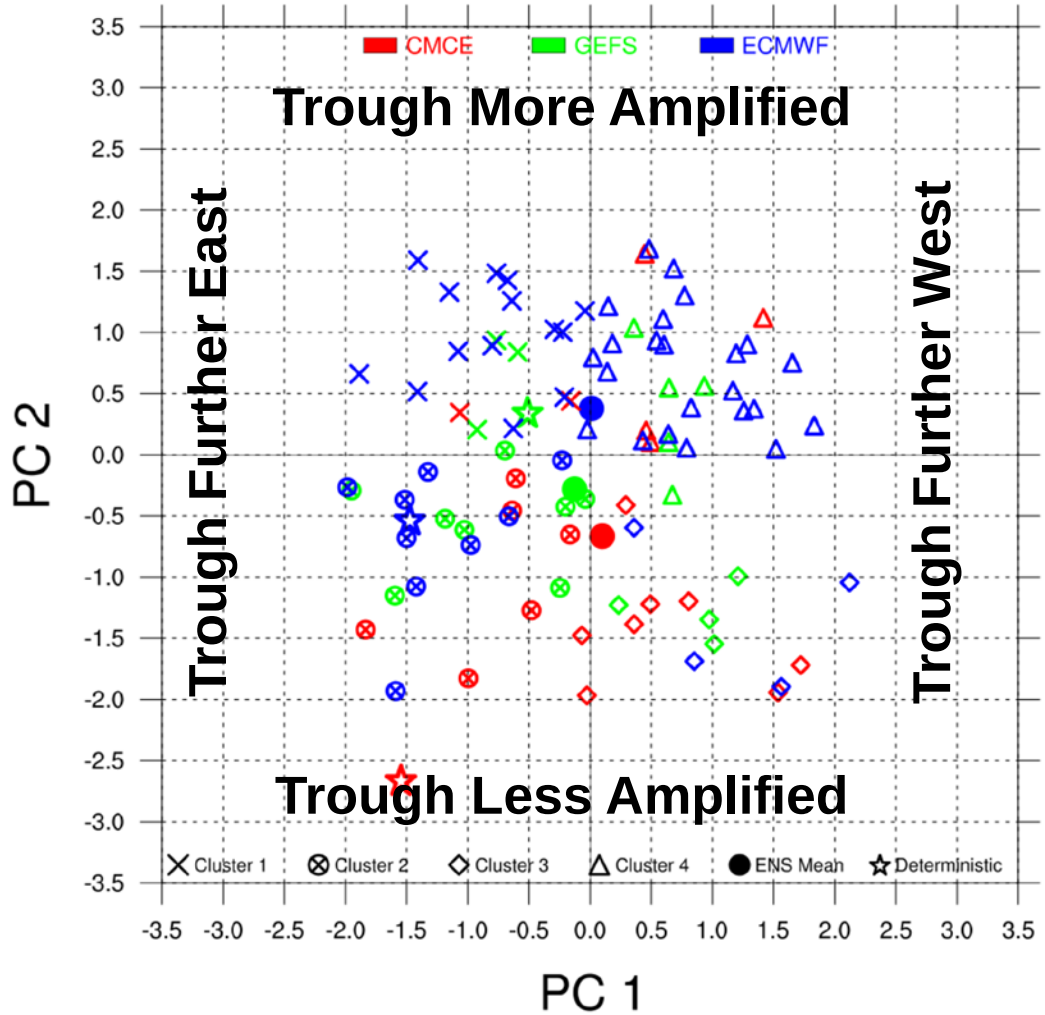


2. Use k-means Clustering



Cluster 1: Trough further **east** and **more amplified** than mean

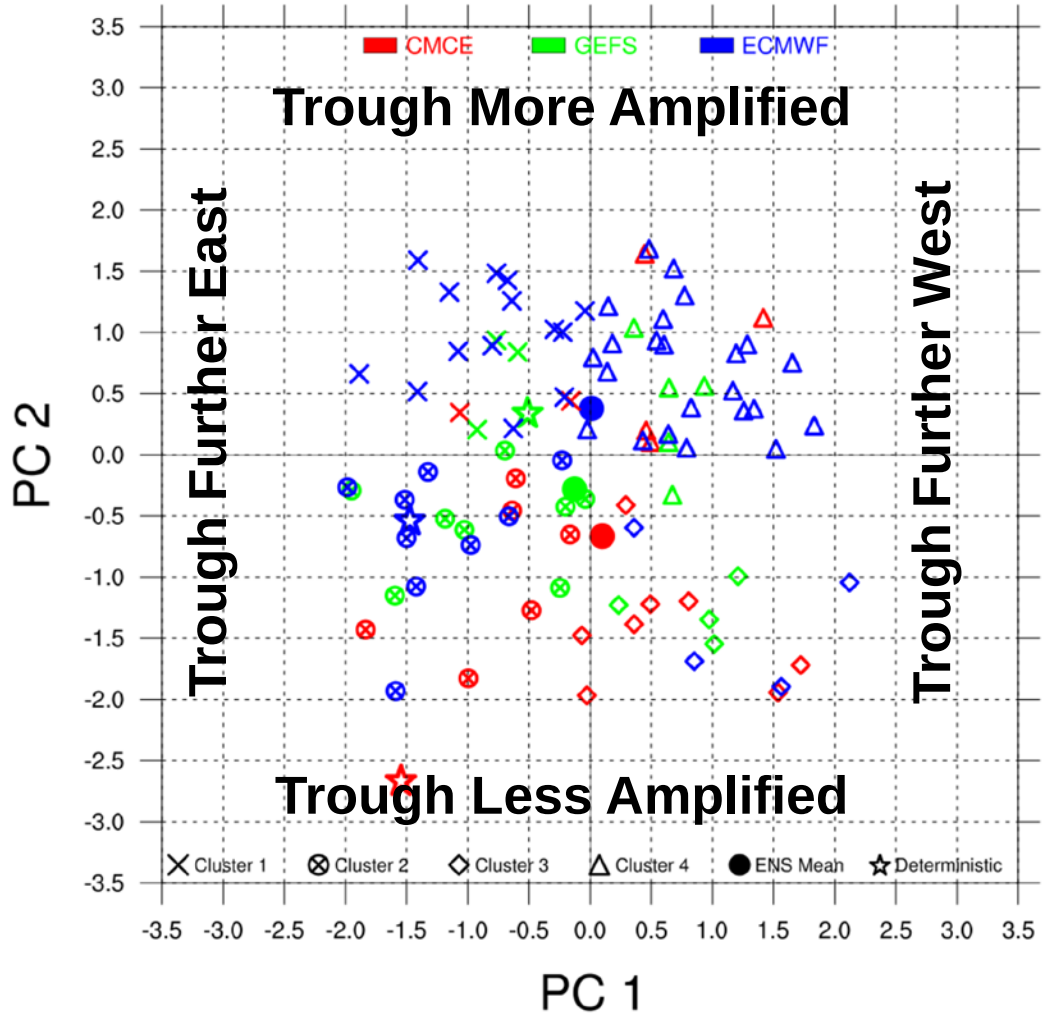
2. Use k-means Clustering



Cluster 1: Trough further **east** and **more amplified** than mean

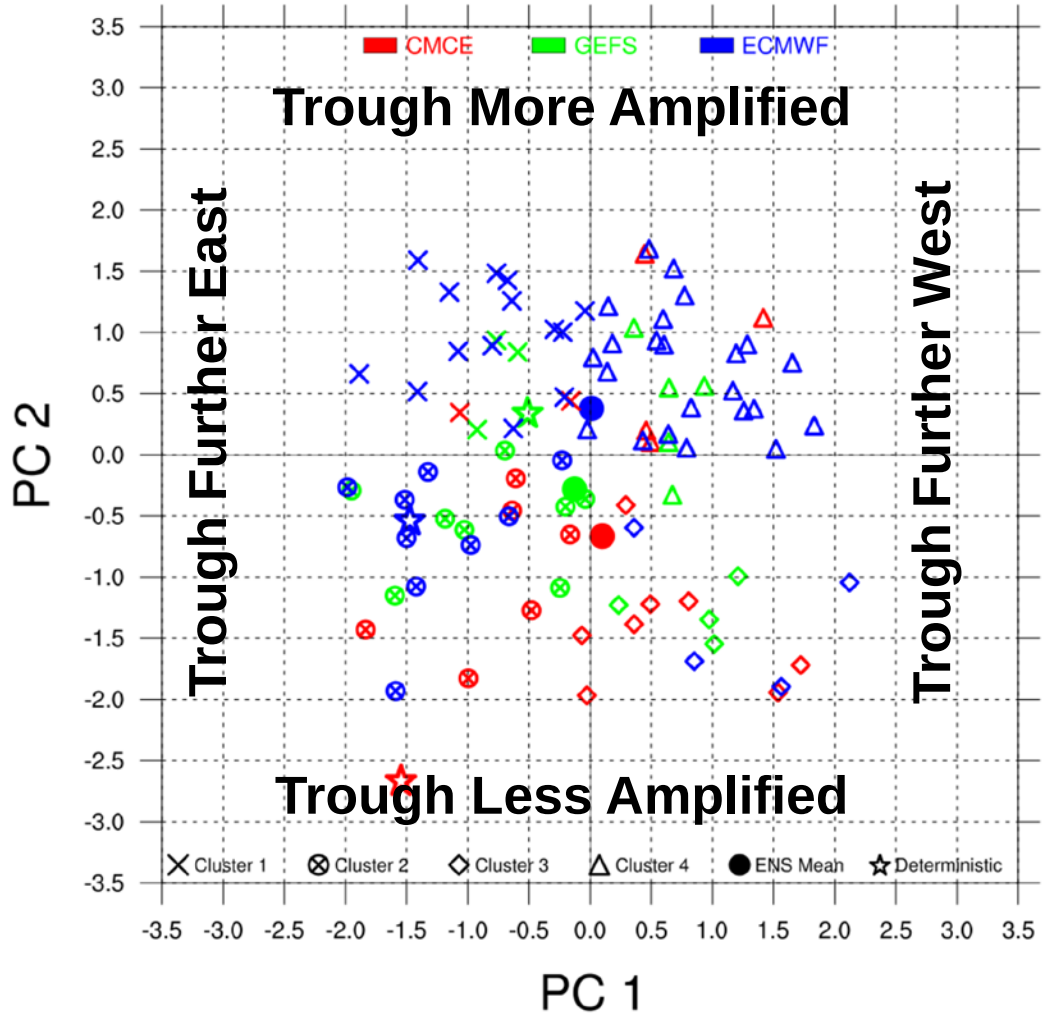
Cluster 2: Trough further **east** and **less amplified** than mean

2. Use k-means Clustering



Cluster 1: Trough further **east** and **more amplified** than mean
Cluster 2: Trough further **east** and **less amplified** than mean
Cluster 3: Trough further **west** and **less amplified** than mean

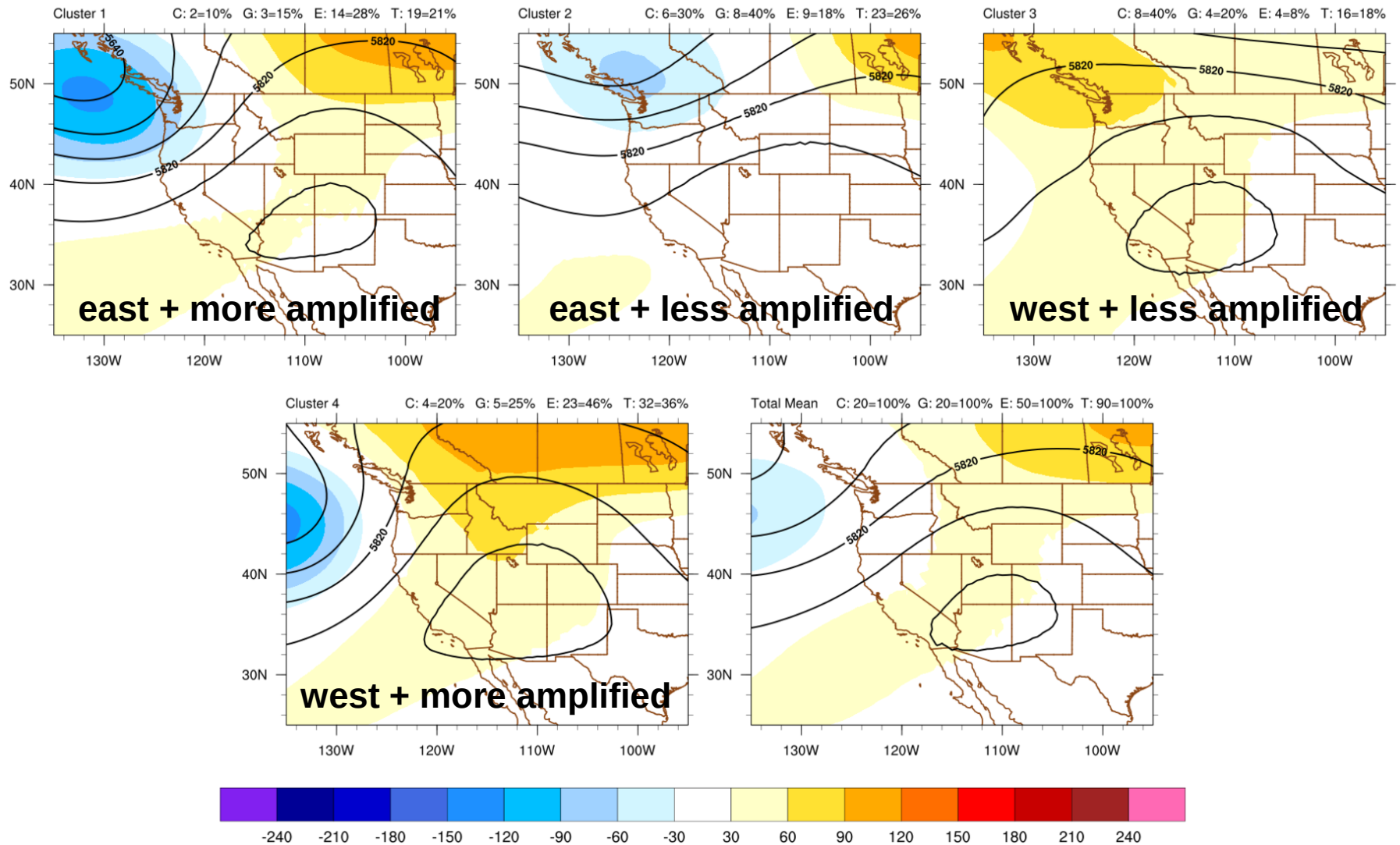
2. Use k-means Clustering



Cluster 1: Trough further **east** and **more amplified** than mean
Cluster 2: Trough further **east** and **less amplified** than mean
Cluster 3: Trough further **west** and **less amplified** than mean
Cluster 4: Trough further **west** and **more amplified** than mean

3. View the Clusters

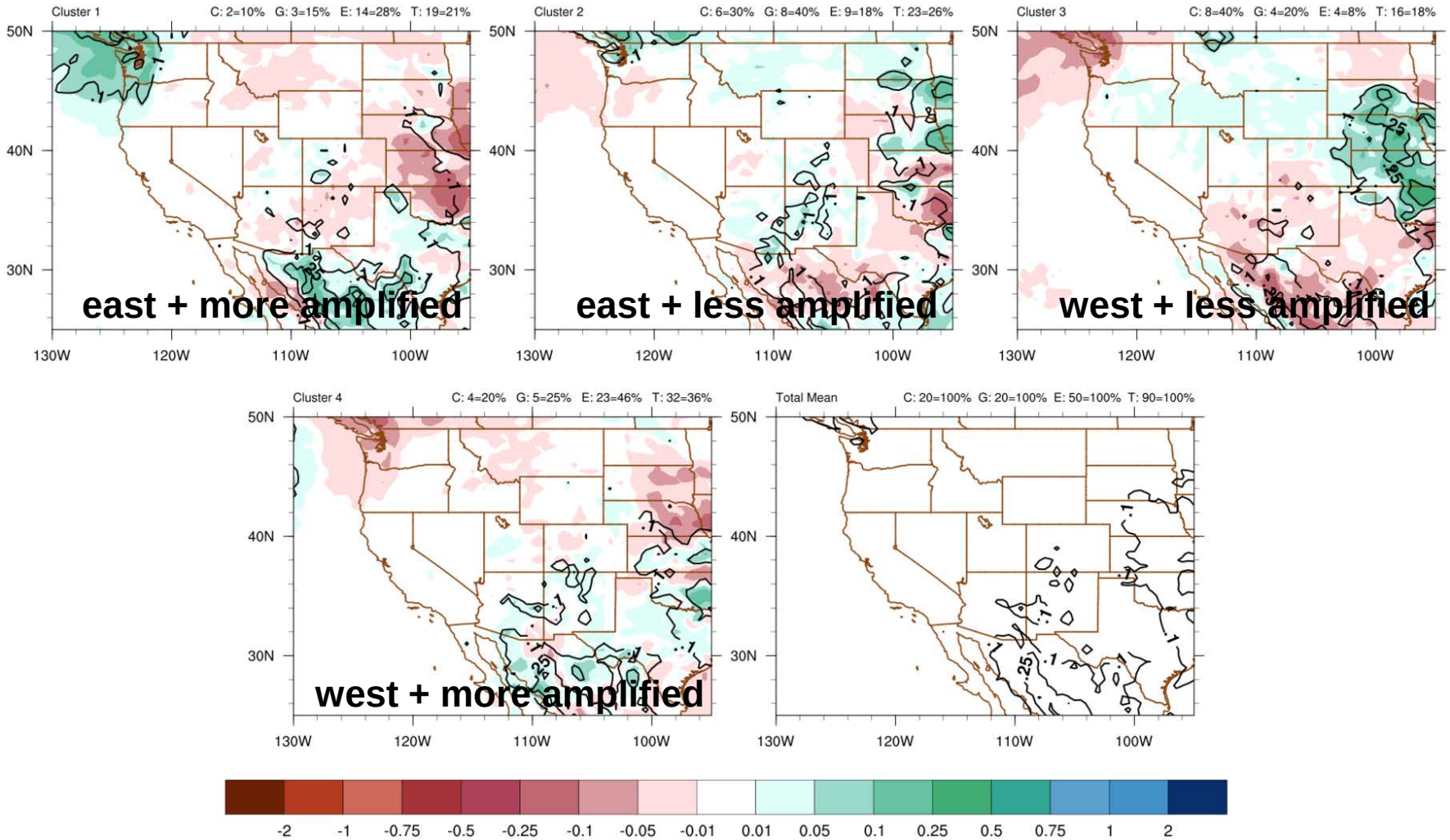
Init: 0000 UTC Thu Jul 23 2020 Valid: 24-hours Ending 0000 UTC Fri Jul 31 2020



Cluster mean 500 hPa heights (contours) and anomalies (color fill)

3. View the Clusters

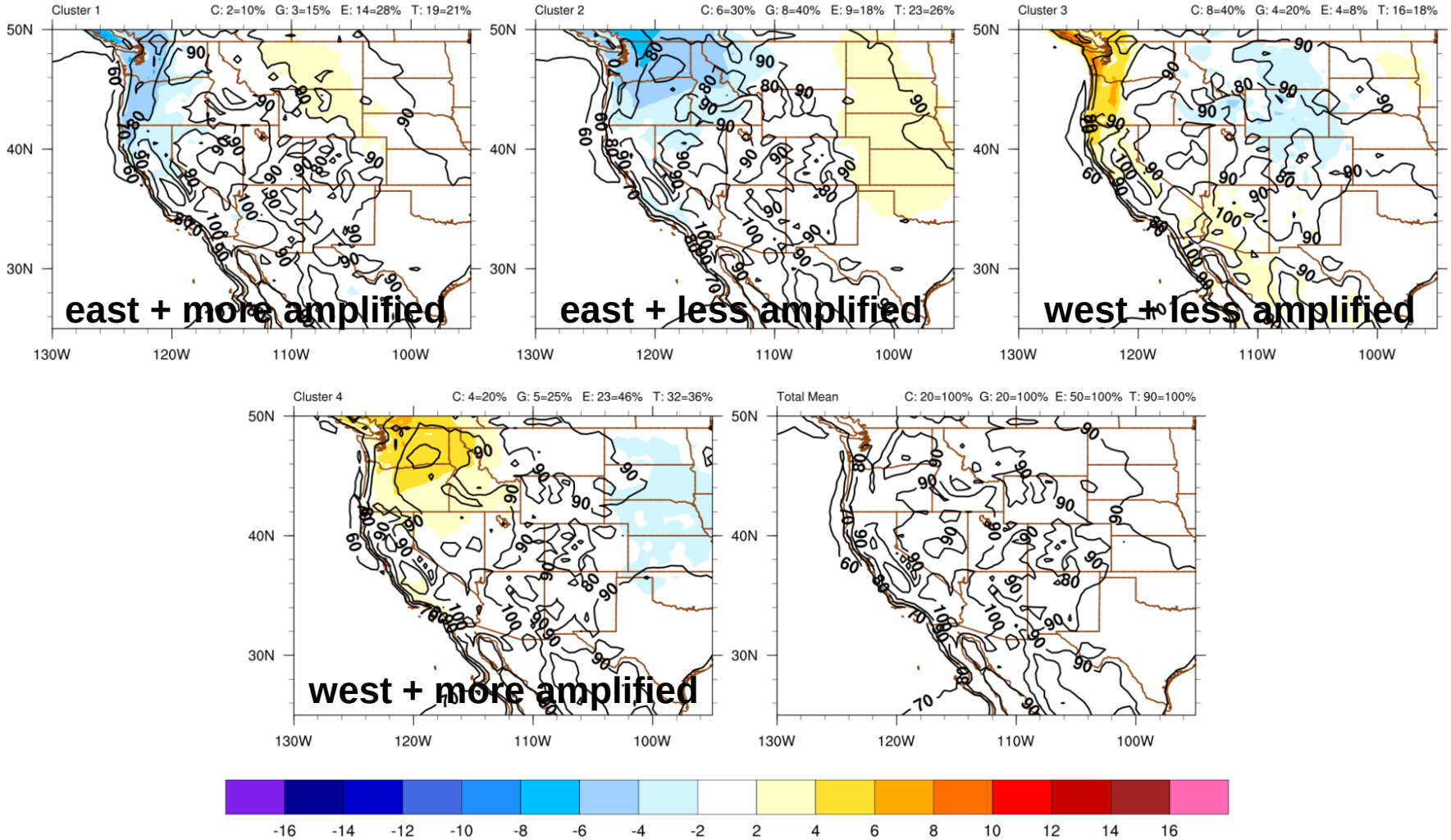
Init: 0000 UTC Thu Jul 23 2020 Valid: 24-hours Ending 0000 UTC Fri Jul 31 2020



Cluster mean QPF (contours) and the difference between each cluster mean and the ensemble mean of all 90 members (color fill)

3. View the Clusters

Init: 0000 UTC Thu Jul 23 2020 Valid: 24-hours Ending 0000 UTC Fri Jul 31 2020



Cluster mean TMAX (contours) and the difference between each cluster mean and the ensemble mean of all 90 members (color fill)

Putting it All Together

- A forecaster coming into work 7 days prior to Thursday could view the 500 hPa height EOFs and immediately see that the greatest uncertainty in the forecast was the location and amplitude of the trough offshore of the Western US.

- Viewing WPC's fuzzy cluster products would help them see how differences in the large-scale pattern affect precipitation amounts and temperatures over the Western US.

- They can also get an idea of how much ensemble support each scenario has.

- This is useful information when making and communicating a forecast and the possible ways it could unfold.

Future Work

- Incorporate bias corrected and downscaled forecasts from the global ensembles.
- Incorporate new GEFS when it becomes available (10 more members!)
- Applying fuzzy clustering to create forecast scenarios for QPF exclusively.

WPC Clustering Products on the Web

- Day 3-7 Conus Weather Clusters (Max temp, Min Temp, QPF): https://origin.wpc.ncep.noaa.gov/hmt/wk2/day_3_7/view.php

- Day 8-10 Conus Weather Clusters (Max temp, Min temp, QPF): https://origin.wpc.ncep.noaa.gov/hmt/wk2/clusters/view.php

Zheng, M., E. K. Chang, B. A. Colle, Y. Luo, and Y. Zhu, 2017: Applying fuzzy clustering to a multimodel ensemble for U.S. East Coast winter storms: Scenario identification and forecast verification. *Wea. Forecasting*, **32**, 881–903, https://doi.org/10.1175/WAF-D-16-0112.1.

Thanks!

- Day 3-7 Conus Weather Clusters (Max temp, Min Temp, QPF): https://origin.wpc.ncep.noaa.gov/hmt/wk2/day_3_7/view.php

- Day 8-10 Conus Weather Clusters (Max temp, Min temp, QPF): https://origin.wpc.ncep.noaa.gov/hmt/wk2/clusters/view.php

Questions or Comments?

bill.lamberson@noaa.gov