Diving into the Forecast with Ensemble Clustering and Sensitivity Analysis: Insights from the 2024 FFaIR Experiment and Applications to Winter Weather & ARs

Austin Coleman[1], Jim Nelson[2], Brian Colle[3], Travis Wilson[4], and Sai Yashwanth Brahmadevara[3]

WWE/PEAR Seminar Series | Thurs December 12th, 2024 | 1pm ET

[1]CIRES/WPC, [2]NOAA WPC, [3]Stony Brook University, [4]NOAA GSL

Motivation

As technology improves and NWS responsibilities expand

- Forecasters have access to **more data** with simultaneously **less time** to interrogate those data

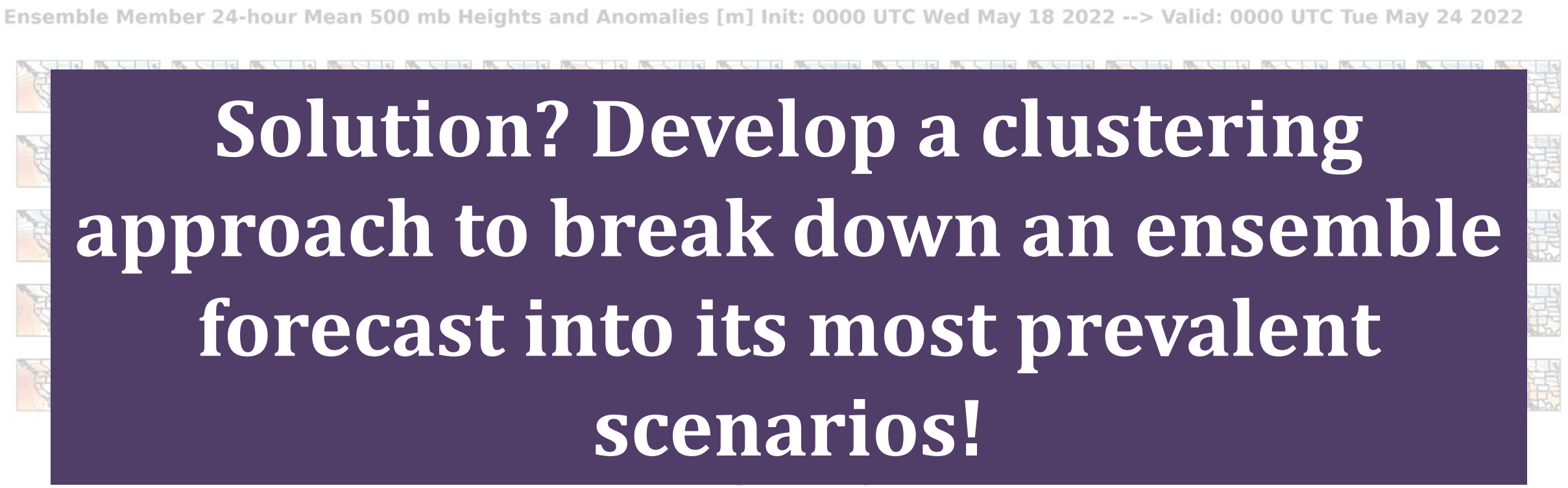
The National Blend of Models (NBM) is frequently used as a first-guess for said forecasts

- Blends a large amount of forecast data, but can be seen as a black box
- Forecasters desire more information about what makes up the NBM

Ensemble mean of NBM's sub-ensemble systems (CMCE, GEFS, and ECMWF) is one way to quickly summarize solutions

- **Problem:** it often washes out important nuance amongst ensemble membership

Motivation



Solution? Develop a clustering approach to break down an ensemble forecast into its most prevalent scenarios!

This isn't a viable answer either!

But first, how does clustering work?

Relies heavily on Empirical Orthogonal Function (EOF) Analysis, traditionally known in statistics as Principal Component Analysis (PCA)

Empirical Orthogonal Function (EOF) Analysis decomposes a series of spatial fields into mathematically-independent (orthogonal) modes

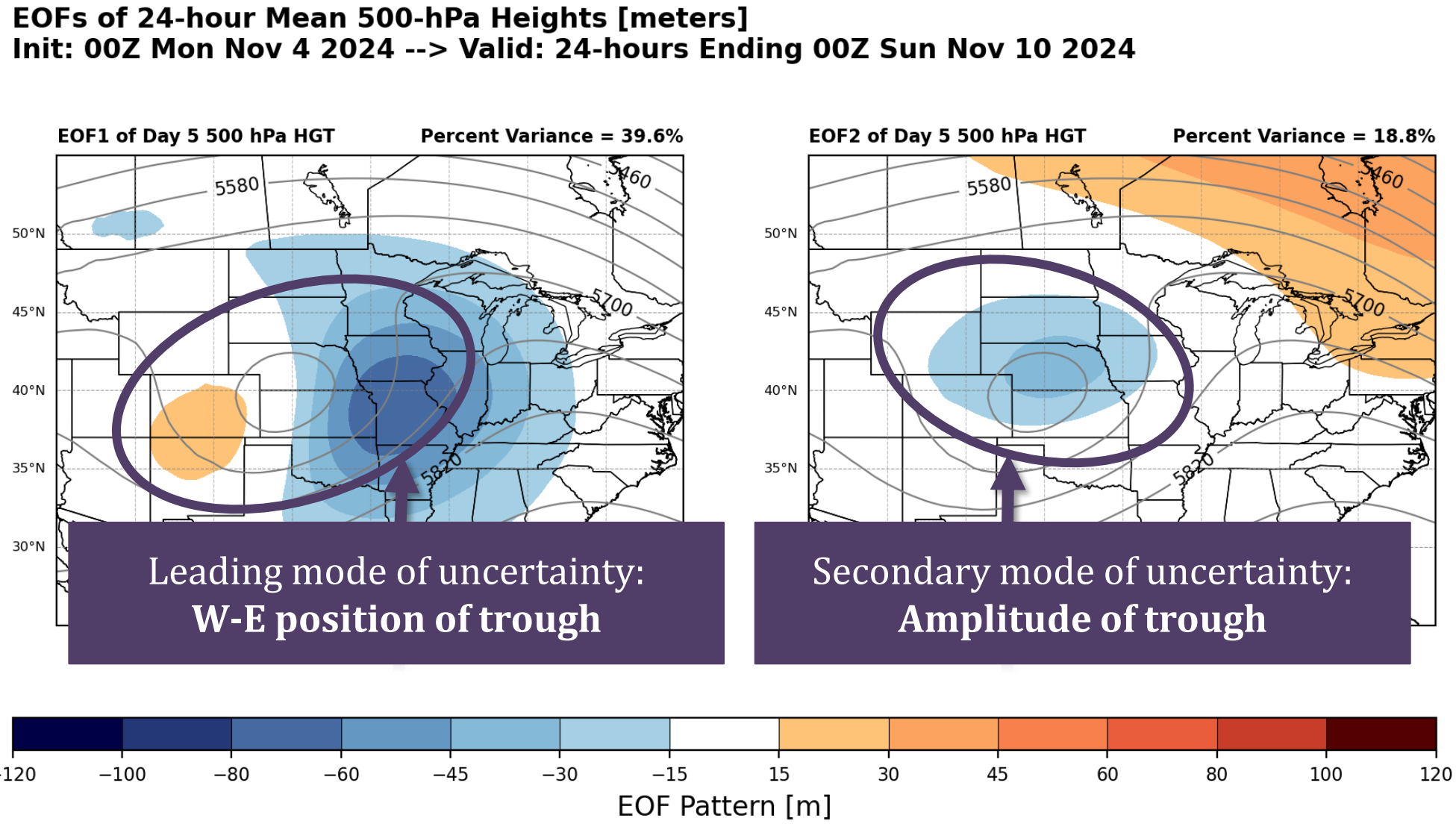
Method most often used by climatologists to understand leading spatial modes of variability in a time series

> For example, it could be used to answer the question: *What wind patterns are typically associated with various phases of the North Atlantic Oscillation?*

For ensemble clustering, we seek to identify the **leading spatial modes of variability amongst the ensemble membership**

("THE CLIMATE DATA GUIDE: EMPIRICAL ORTHOGONAL FUNCTION (EOF) ANALYSIS AND ROTATED EOF ANALYSIS", 2013)

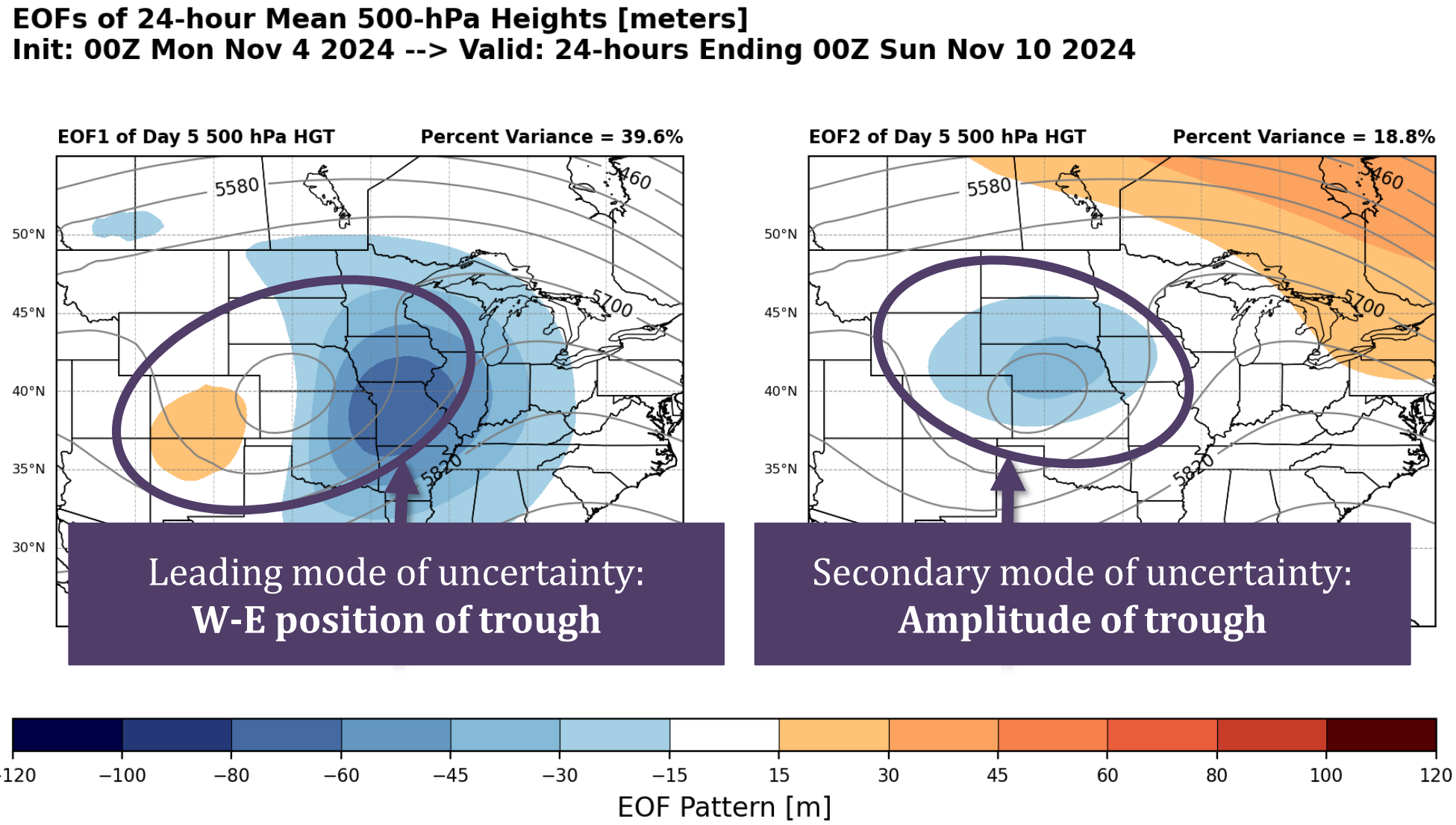
We break down the forecast (super-ensemble of CMCE, GEFS, & ECMWF) into its leading modes of variability via EOF Analysis



Common source of confusion:

What do the positives and negatives mean?

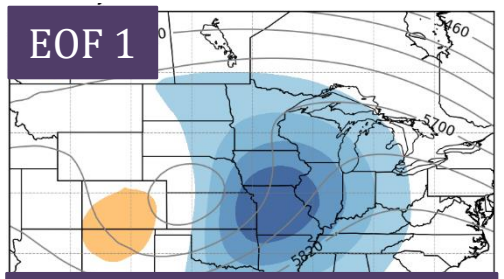
- Sign doesn't matter when looking at the EOFs themselves
- Only becomes important once we start clustering or looking at members in PC phase space



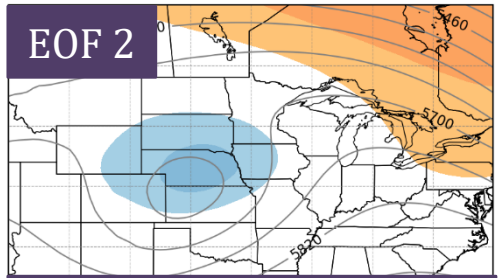
How does it work?

Next, we group similar ensemble solutions together with clustering

First two EOFs for reference



Leading uncertainty mode:
W-E position of trough



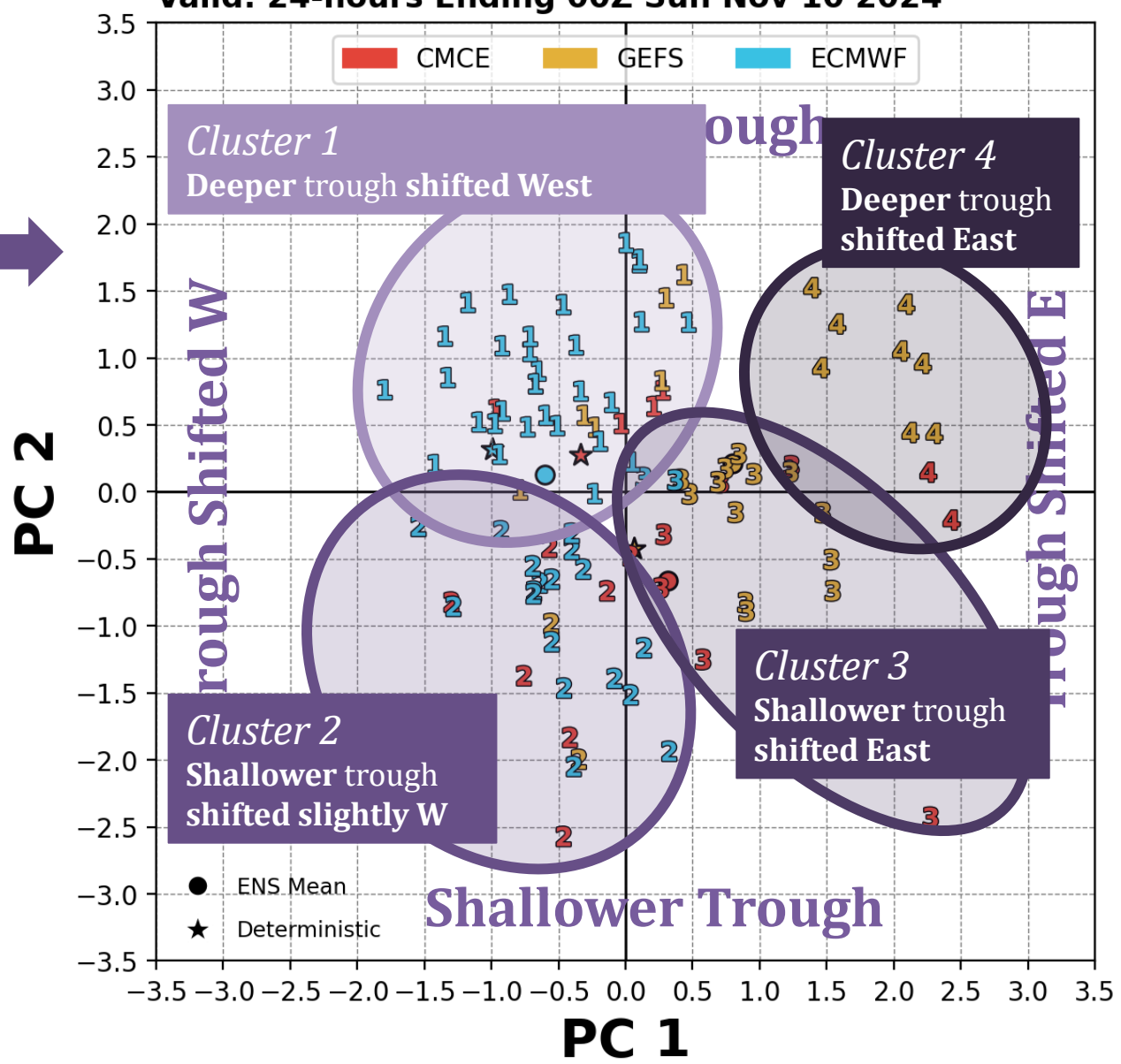
Secondary uncertainty mode:
Amplitude of trough

Projecting ensemble members into PC phase space shows us the forecast scenario for each member

Members with positive PC1 will look more like EOF1 *(trough shift E)*

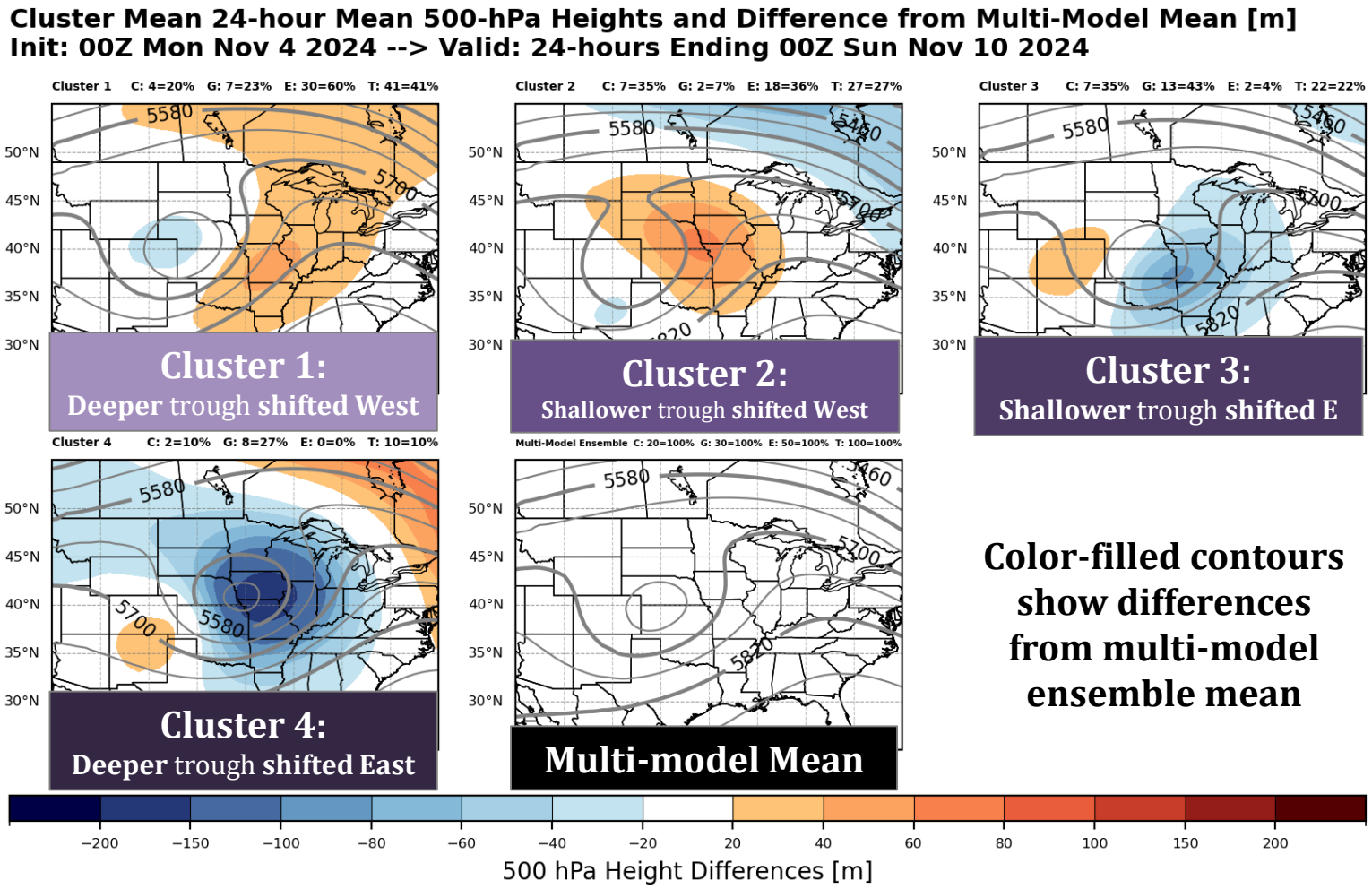
whereas members with negative PC1 will look opposite EOF1 *(trough shift W)*

K-means Clustering groups members with similar solutions based on our leading modes of uncertainty



Don’t even need to look at EOFs or PCs to use!

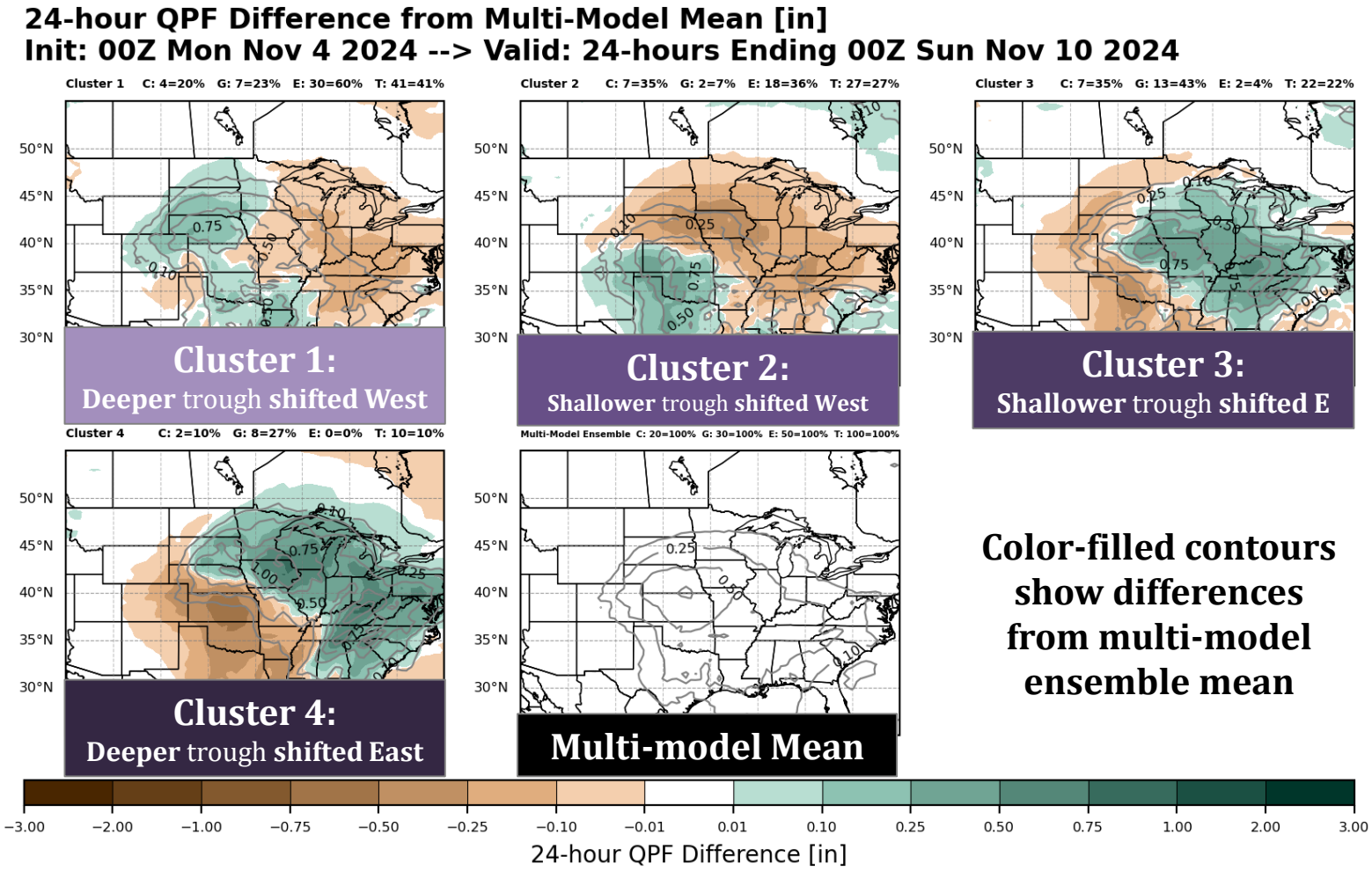
Can skip straight to the cluster forecasts (of 500-hPa heights in this case)



Color-filled contours show differences from multi-model ensemble mean

Can use 500-hPa height clusters to predict other fields

24-hr QPF

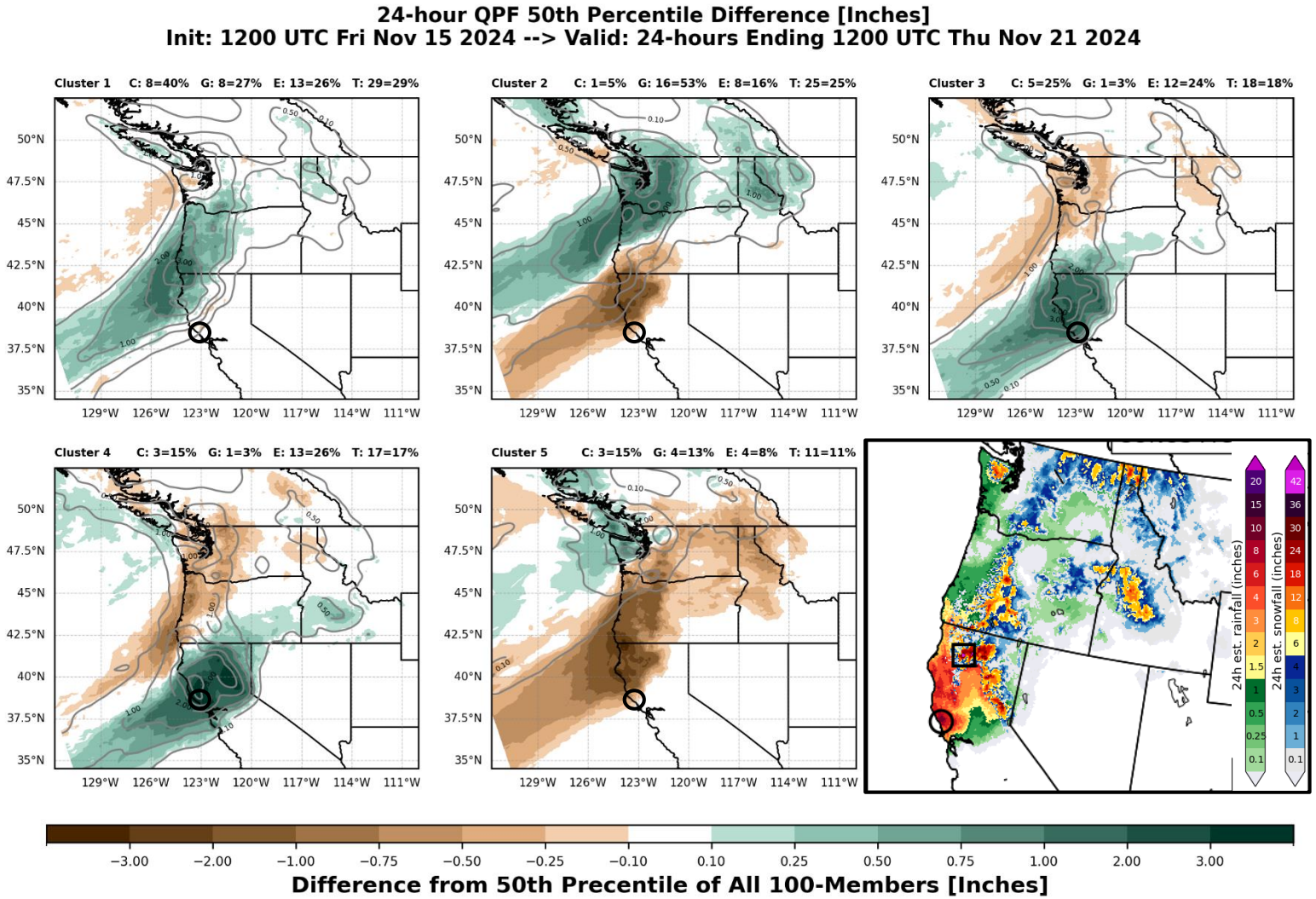


We additionally have a WPC page that clusters directly on QPF!

Algorithm

- ✓ Identifies QPF objects (0.50” Day 5 QPF object shown here)
- ✓ Picks a varying number of clusters based on the silhouette score (5 clusters picked here)
- ✓ **Provides more skillful QPF scenarios than using QPF derived from 500-hPa height clusters!**

(Kiel and Colle 2024)



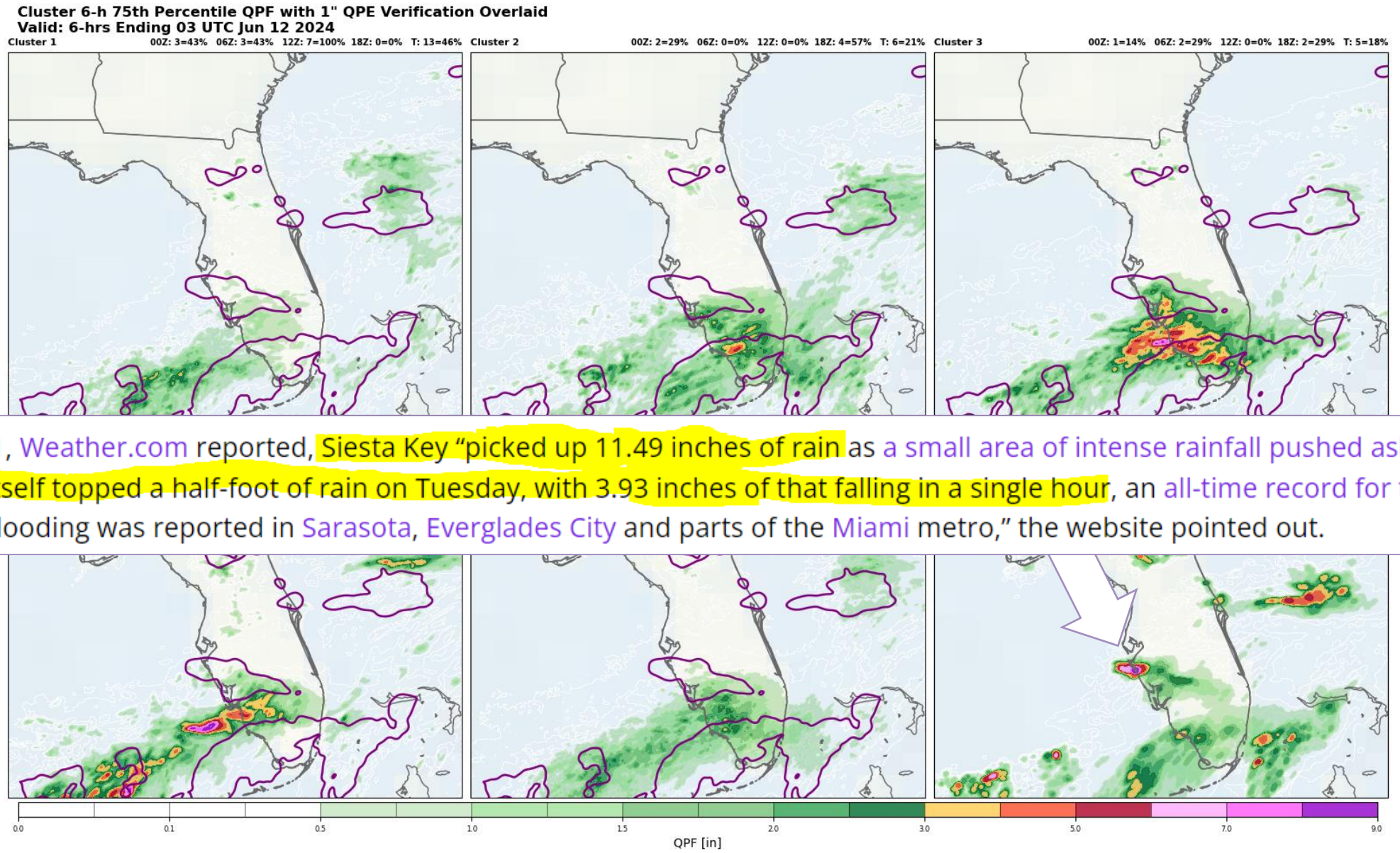
But what about QPF from Convection-allowing Ensembles?

We experimentally applied this method to a time-lagged RRFS ensemble system for the 2024 FFaIR

REFS Configuration used throughout most of FFaIR

06 UTC Cycle	00 UTC Cycle	18 UTC (D-1) Cycle	12 UTC (D-1) Cycle
m1 (ctrl) 06Z	m8 (m1 00Z)	m15 (m1 18Z Day - 1)	m22 (m1 12Z Day - 1)
m2 06Z	m9 (m2 00Z)	m16 (m2 18Z Day - 1)	m23 (m2 12Z Day - 1)
m3 06Z	m10 (m3 00Z)	m17 (m3 18Z Day - 1)	m24 (m3 12Z Day - 1)
m4 06Z	m11 (m4 00Z)	m18 (m4 18Z Day - 1)	m25 (m4 12Z Day - 1)
m5 06Z	m12 (m5 00Z)	m19 (m5 18Z Day - 1)	m26 (m5 12Z Day - 1)
m6 06Z	m13 (m6 00Z)	m20 (m6 18Z Day - 1)	m27 (m6 12Z Day - 1)
m7 HRRR 06Z	m14 HRRR 00Z	m21 HRRR 18Z (Day – 1)	m28 HRRR 12Z (Day – 1)

The next day, participants subjectively verify MRTP clusters with MRMS observations

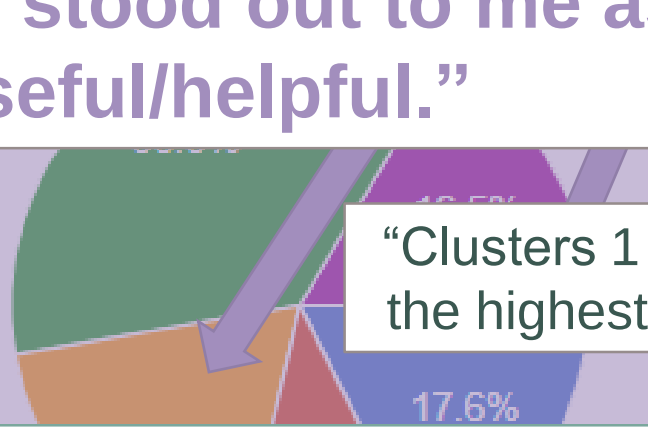


On June 11, Weather.com reported, Siesta Key "picked up 11.49 inches of rain as a small area of intense rainfall pushed ashore. Sarasota itself topped a half-foot of rain on Tuesday, with 3.93 inches of that falling in a single hour, an all-time record for that location. Flooding was reported in Sarasota, Everglades City and parts of the Miami metro," the website pointed out.

Summary of Subjective Evaluation Results!

Cluster 3 and Cluster 1 tend to follow as
scenarios
le, Cluster 1 (which
ld theoretically

"Clusters 1 and 4 were consistently the best performing clusters throughout the week. I don't know what is special about those two particular clusters, but they stood out to me as the most useful/helpful."



ster 1
ster 2
Cluster 3

"Clusters 1 and 4 looked the best, but cluster 1 had the highest magnitudes a little too far to the south."

"Clusters 1 and 4 looked the best to me, but I thought cluster 1 had a slight edge as far as the structure and magnitude."

Why does Cluster 4 tend to subjectively outperform the full ensemble most often?

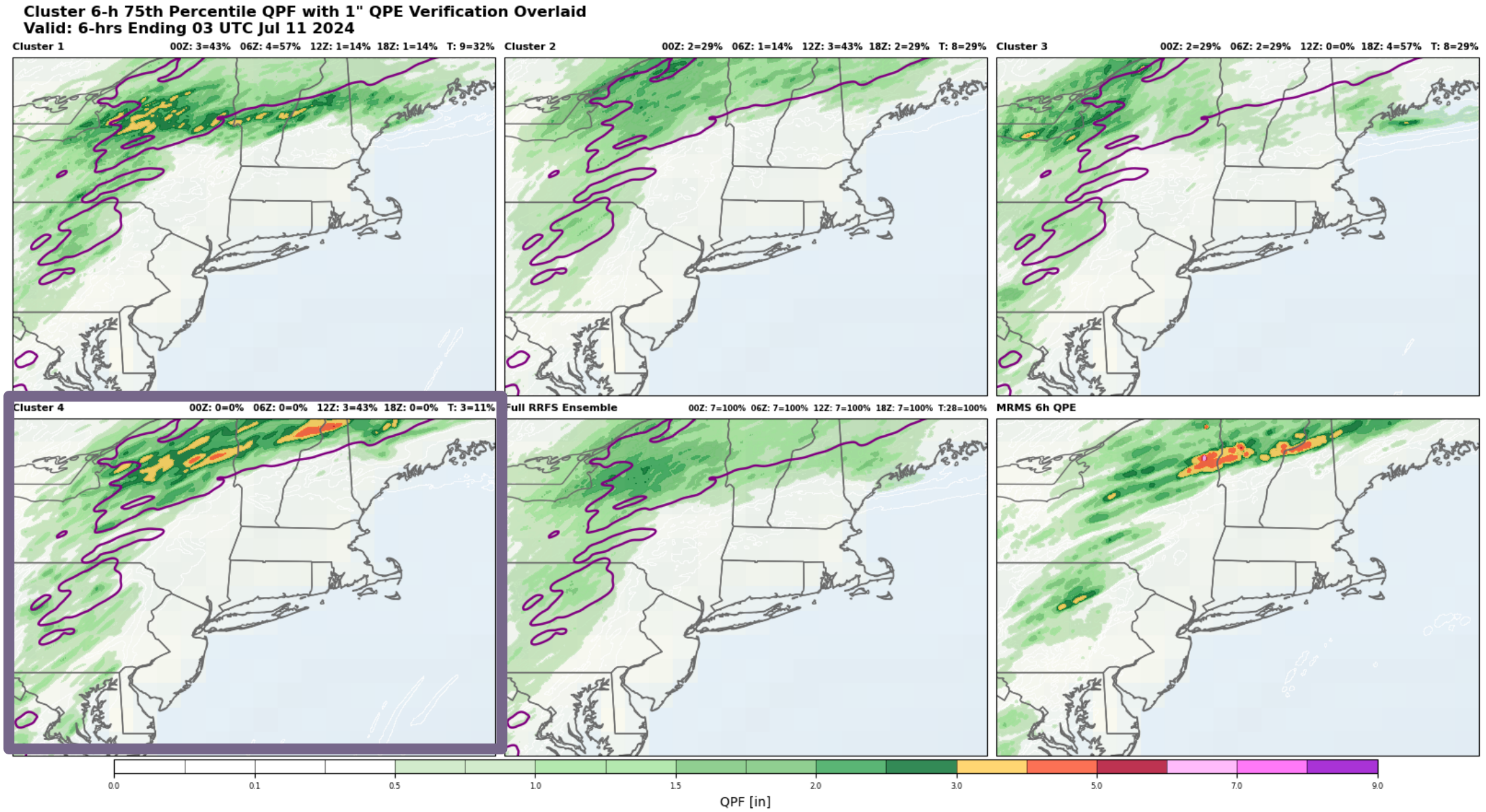
The 14-member REFS composed of the two most recent cycles (half the size of the RRFSe system used here) depicts more binary probabilities than the HREF*

- Suggests under dispersive ensemble

*from Matt Pyle's FFaIR Seminar

Assuming our 28-member RRFSe system is similarly under dispersive, it makes sense that the "outlier" clusters might have a better chance of capturing the correct outcome

Relatively low spatial and magnitude variability of QPF among clusters
- Cluster 4 consists of three "extreme" members, depicting higher QPF amounts than the remaining members



Why does Cluster 4 tend to subjectively outperform the full ensemble most often?

The 14-member … ere) depicts more binary probabili…

Suggests und…

*from Matt P…

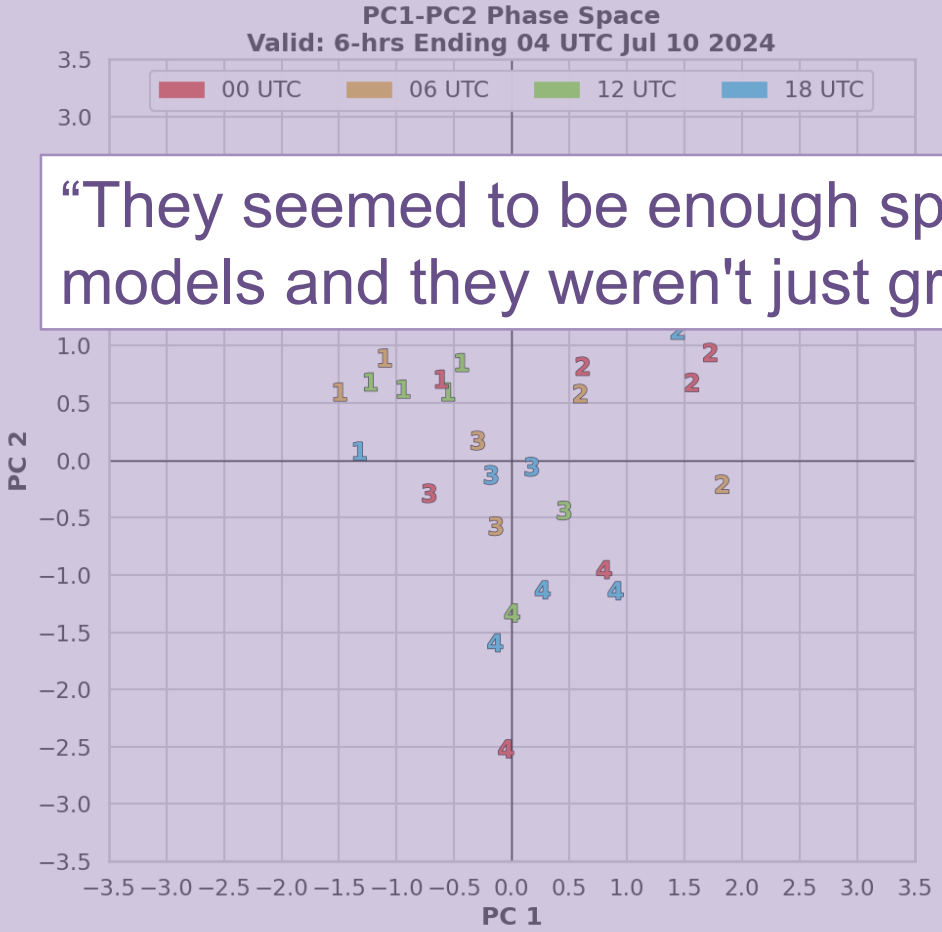
Assuming our 2… lusters might have a better ch…

"I think the finer resolution detail of cluster 4 bumps it ahead of the full ensemble, in my opinion. I want to see that higher resolution detail from a deterministic member and if it can be captured in a cluster of the ensemble, even better!"

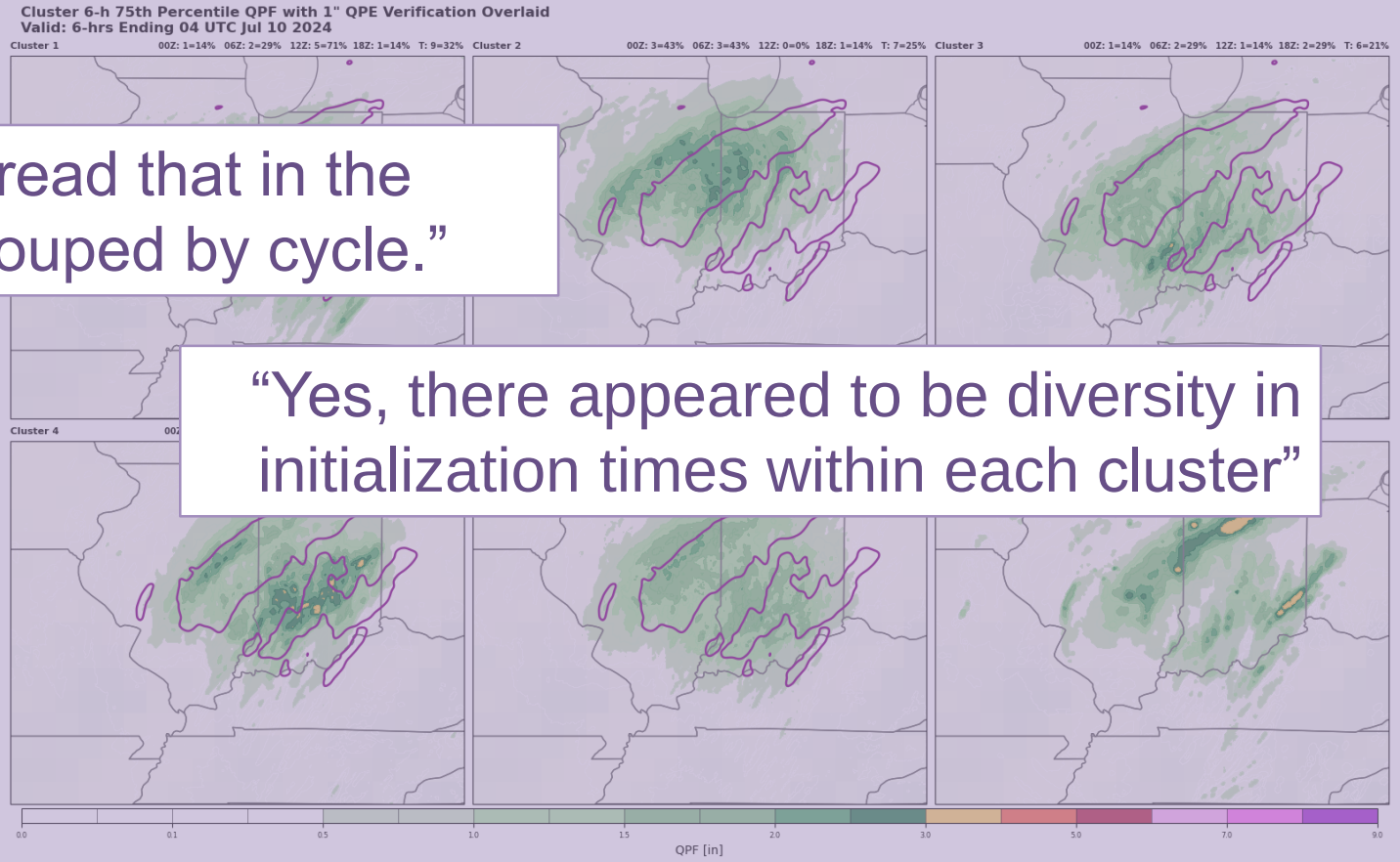
Alternatively, participants may prefer the visualization of Cluster 4 to the others due to its more deterministic appearance!

Suggests visualization could be improved for remaining clusters!

While under dispersive as a whole, the time-lagged ensemble members seem to be represented equally amongst clusters!



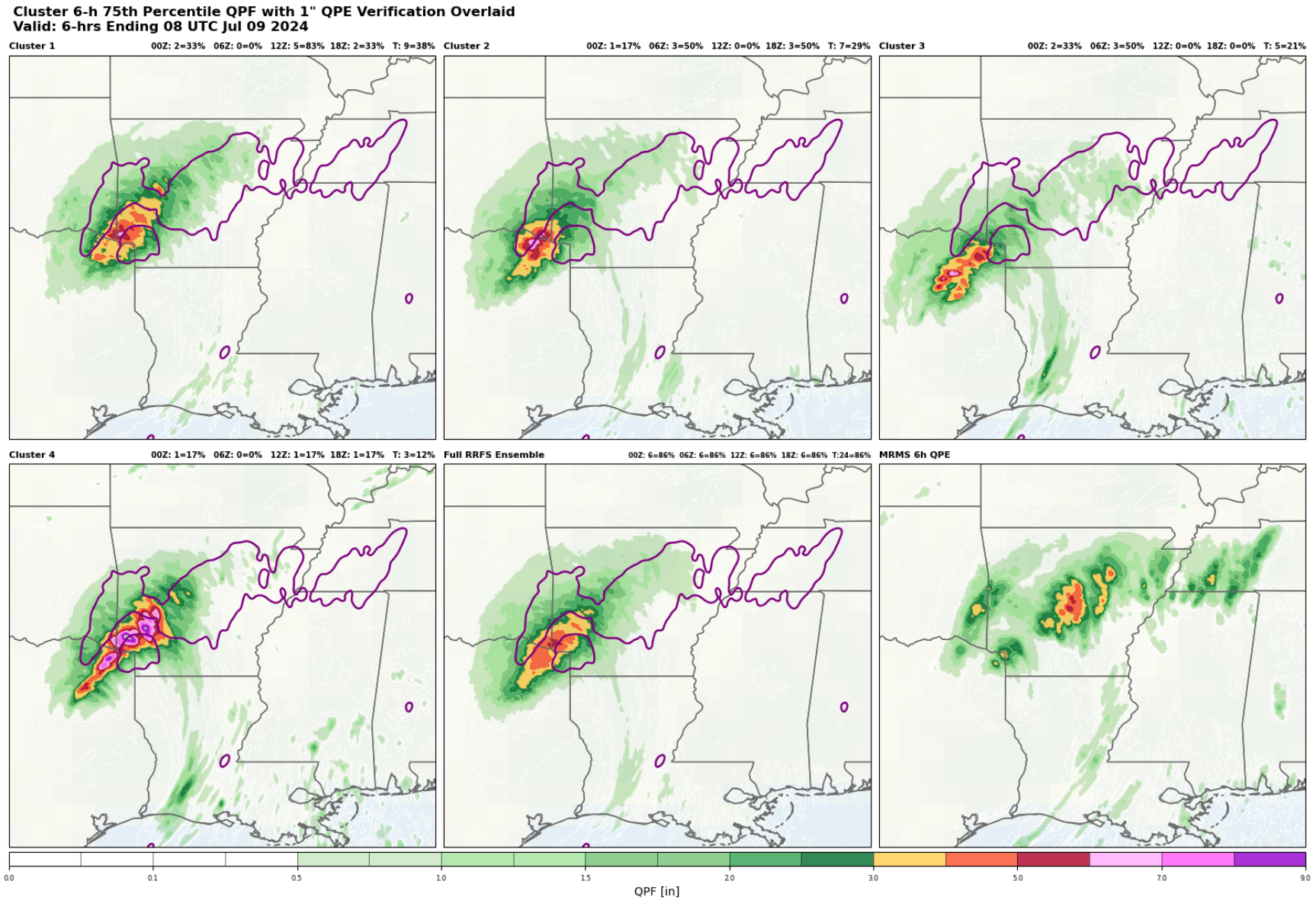
"They seemed to be enough spread that in the models and they weren't just grouped by cycle."



"Yes, there appeared to be diversity in initialization times within each cluster"

Having diversity of cycles within each cluster did not always lead to diversity of scenarios!

- For post-landfall Beryl clusters, we had spread in QPF magnitudes but little spread in placement of QPF Max

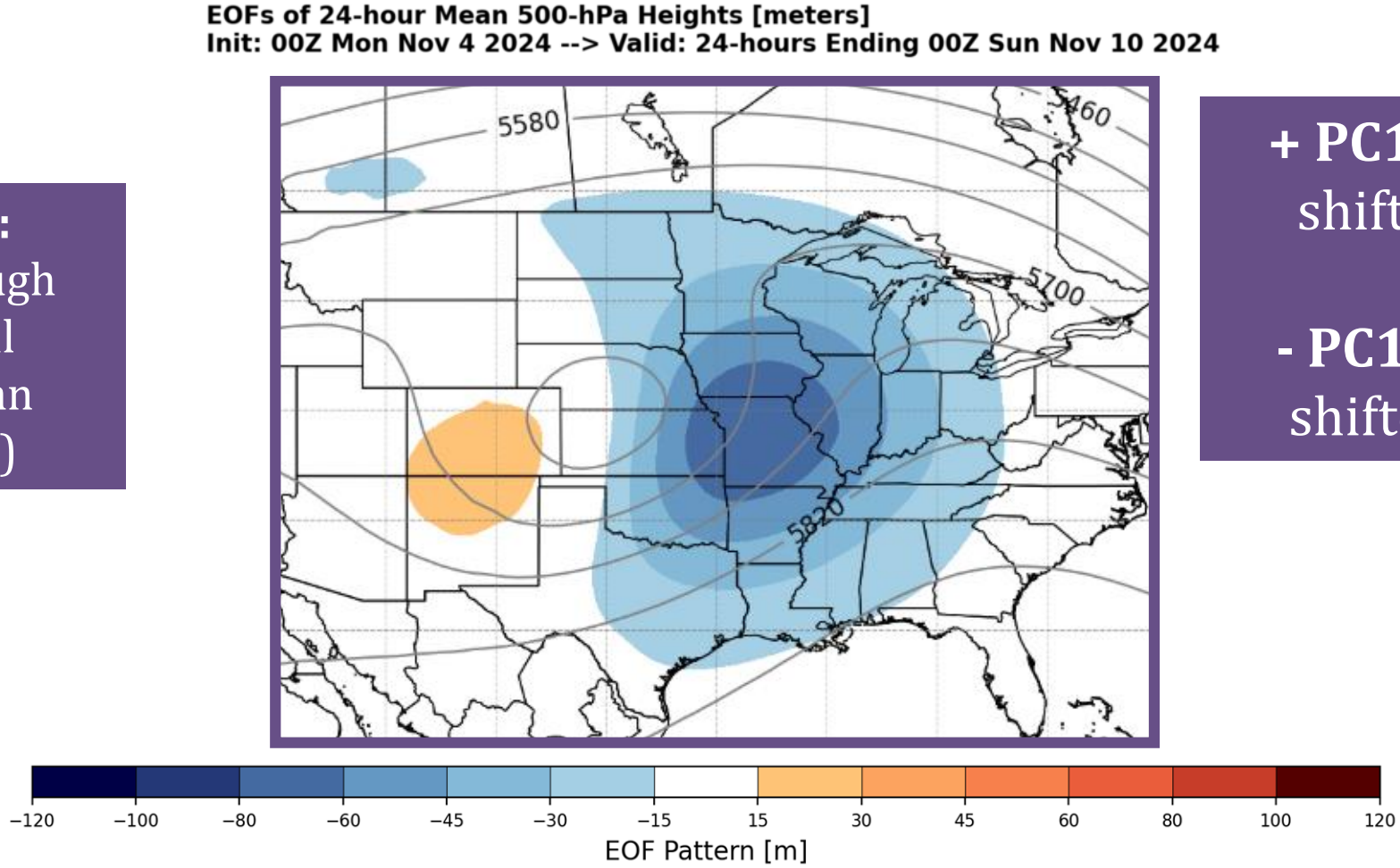


What about when you ***do*** have a lot of diversity in your scenarios? Where does that forecast uncertainty come from?

Ensemble Sensitivity Analysis (ESA) tells us how the atmosphere needs to evolve early on in order to look like a given EOF!

Let's correlate the phase speed uncertainty of the pattern with the early 500-hPa height field

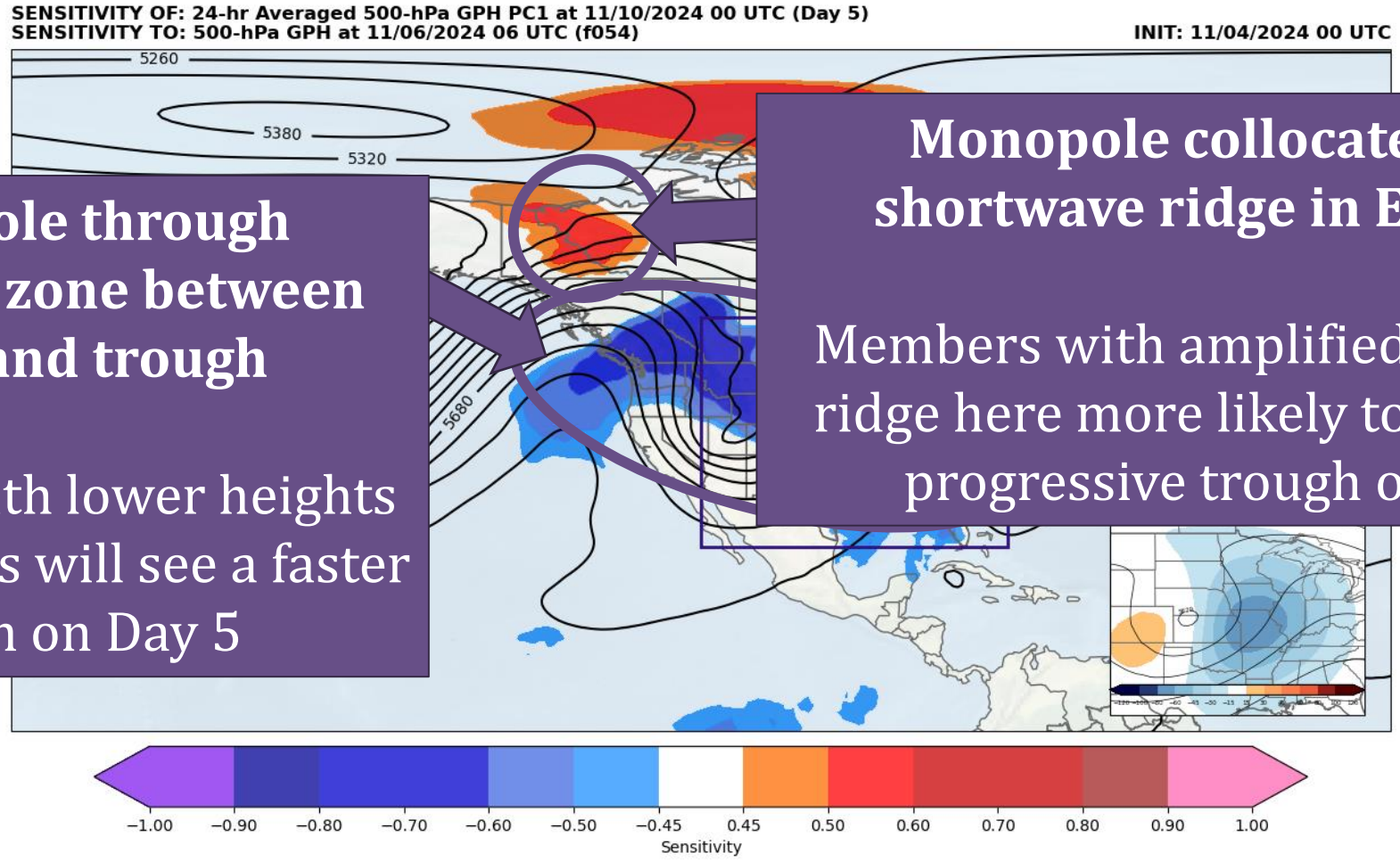
Uncertainty:
Position of trough relative to full ensemble mean (phase speed)



+ PC1 means trough shifted to the East

- PC1 means trough shifted to the West

ESA shows us what the ensemble "cares about" most when predicting the position of the trough at Day 5



Monopole through transition zone between ridge and trough

Members with lower heights in these areas will see a faster trough on Day 5

Monopole collocated with shortwave ridge in ESA Field

Members with amplified shortwave ridge here more likely to have more progressive trough on Day 5

Future Clustering & ESA Work

- Adopting a cluster consistency approach in DESI similar to that used by the Japan Meteorological Agency

  (expected Jan 2025 DESI release pending positive test results next week)

- Creating a verification dashboard with bulk long-term cluster statistics

  (expected summer 2025)

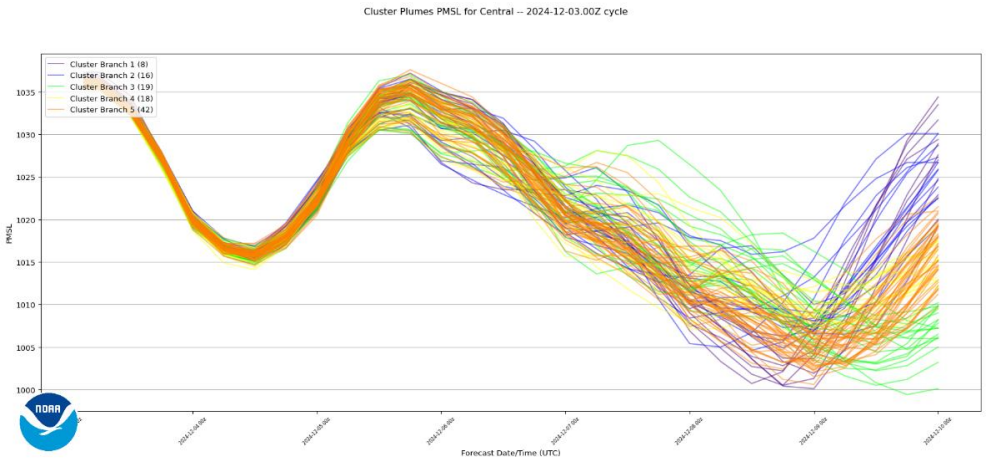
- Exploring idea of a cluster blender product that allows for weighting of the forecast based on clusters

  (playing with idea spring & summer 2025)

- Collaborating with Greg Mann (SOO of NWS Detroit) to potentially develop a temporal clustering application

  Idea is to group similar members in time rather than space

  Plan view clusters also allowed to merge and branch off in time when “distance” between clusters crosses a certain threshold



Take-Home Points

Ensemble clustering is a quick way to distill an ensemble forecast down to its prevalent scenarios

Ensemble sensitivity analysis (ESA) provides context on how the atmosphere must evolve to lead to different cluster scenarios

Testament to the potential of data mining ensemble systems

- As we continue to build techniques that extract information from these datasets, need to keep forecaster needs at the forefront
- Lots of room for O2R/R2O in these spaces

Links!

Day 3-9 500Z Clusters + Ensemble Sensitivity Analysis Page

FFaIR 2024 RRFSe QPF Clusters

Days 1-6 CONUS QPF (NBM 4.2 QMDs) Cluster Page

Days 3-9 500Z (CMCE+GEFS+ECMWF) Cluster Page

Days 3-9 MSLP (CMCE+GEFS+ECMWF) Cluster Page