Ensemble Prediction and Predictability of Extreme Weather via Circulation Regimes

46th Annual Climate Diagnostics and Prediction Workshop

Greg Jennrich[1]

David Straus[2], Muthuvel Chelliah[3], and Cory Baggett[1]

[1] Innovim LLC

[2] George Mason University, AOES Department

[3] NOAA Climate Prediction Center

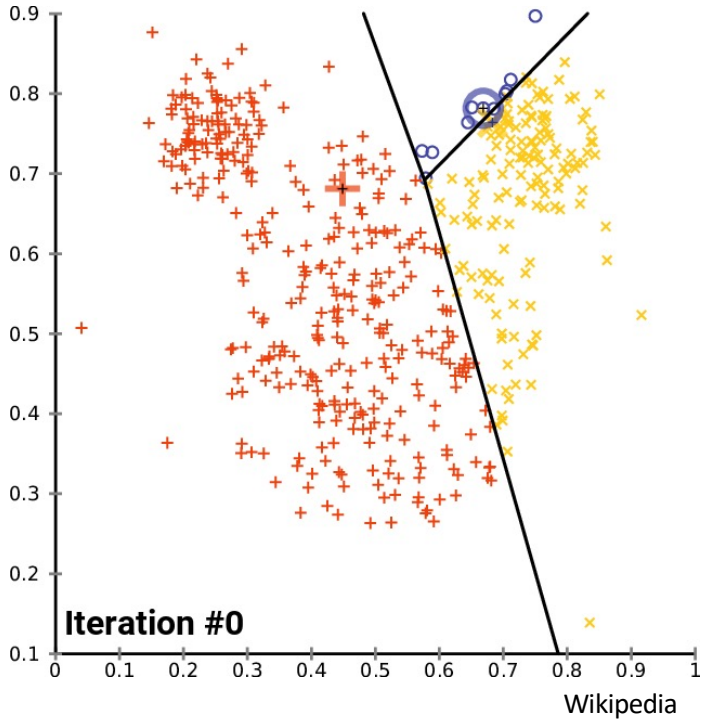
Background

- Forecasting the Week 3/4 period presents many challenges, resulting in a subsequent need for improvements to forecast skill.
- **Goal**: Categorize the forecasted 500 hPa geopotential height field and make a forecast based on historical patterns associated with the assigned pattern (cluster).
- **Step 1**: Identify the 500 hPa cluster patterns in reanalysis
- **Step 2**: Use the cluster assignments to composite other variables (temperature, precipitation, etc.)
- **Step 3**: Make forecasts based on ensemble model assignments of clusters.

Categorizing the circulation pattern via K-means clustering

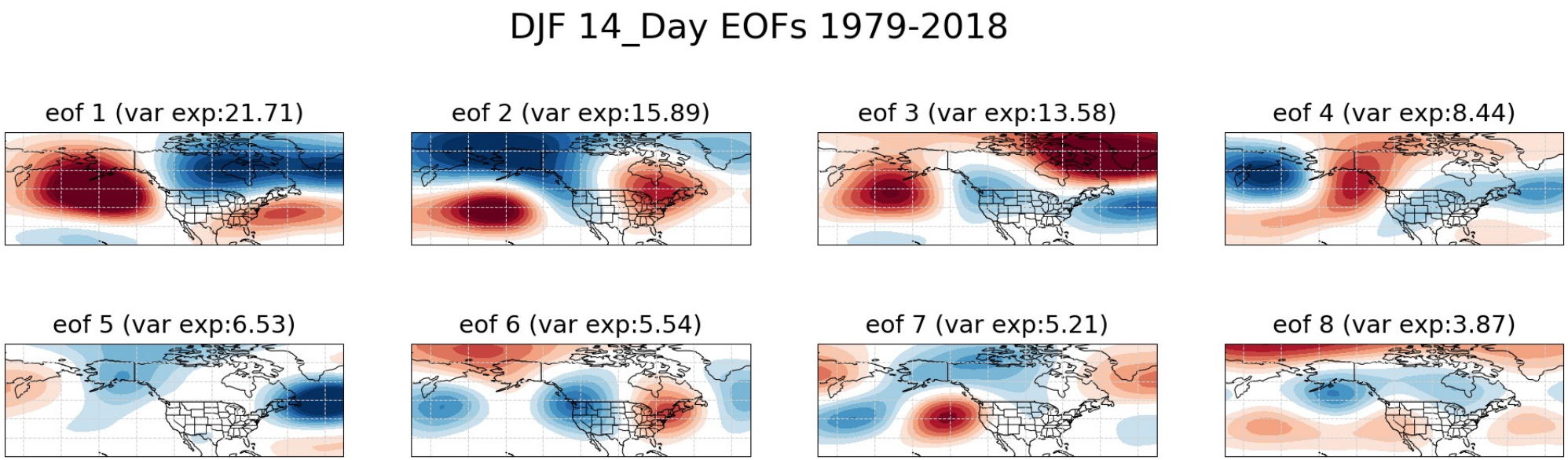
- “This partition is accomplished in such a way that the ratio of the variance between cluster *centroids* (a centroid being the average over all states in a cluster) to the average intracluster variance is maximized. This is equivalent to minimizing the intracluster variance since the total variance of the dataset is fixed. The variances are computed using an Eulerian distance metric in the state space of the PCs” (Straus et al. 2007)

- Data: ERA-Interim z500 anomalies
- Domain: Pacific and W. Atlantic (150 to 330 E, 20 to 80 N)
- Period: 1979/80-2018/19 (40 years)
- Season: DJF- Nov 16- Mar 16 (Mar 15 on leap years), always 121 days

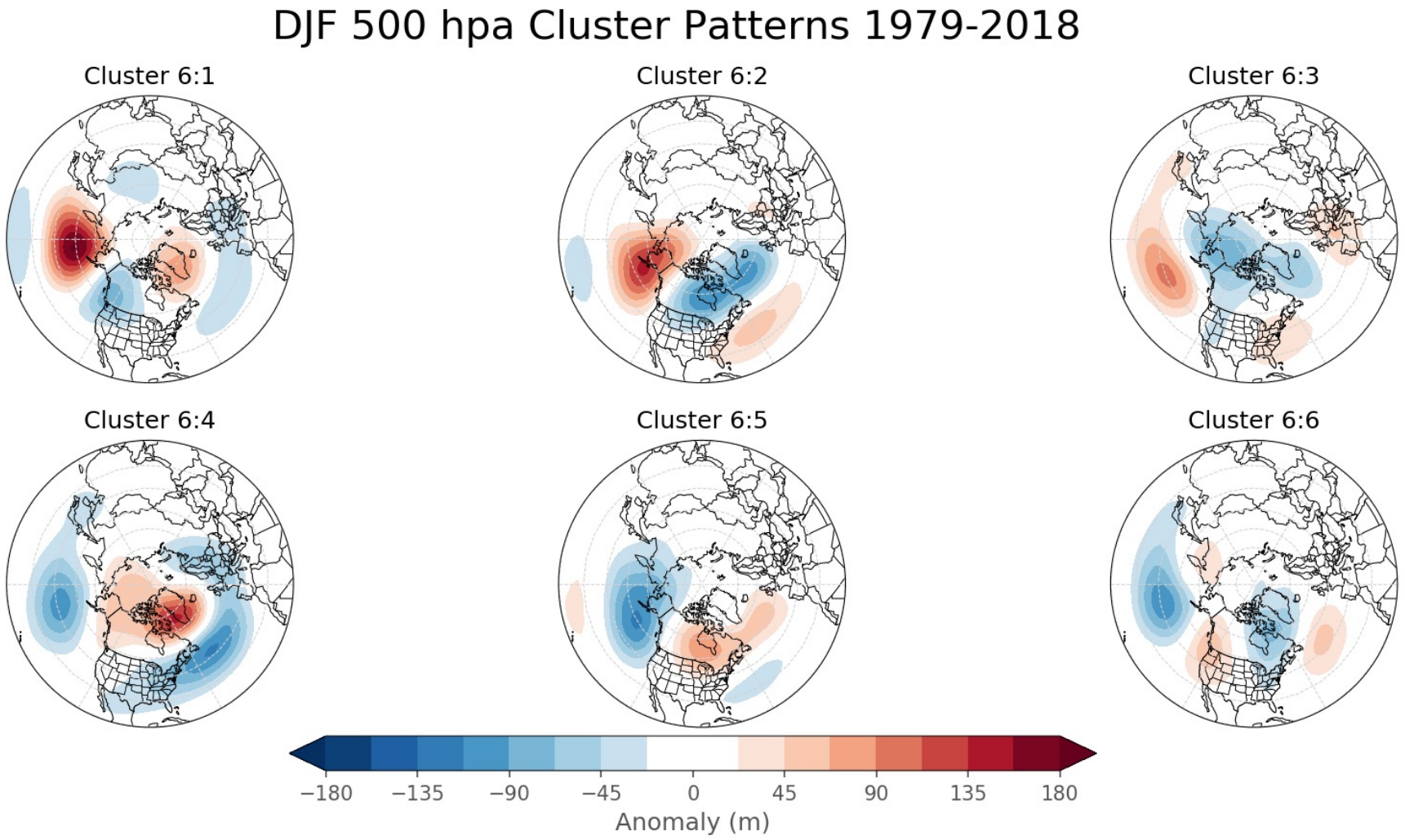


Categorizing the circulation pattern via K-means clustering

- Calculate the leading EOF patterns and the associated principal component (PCs) values
- Use the top 10 PCs for K-means clustering, about 85% of the total variance explained
- K-means clustering assigns clusters to each period. We then composite z500 anomalies, resulting in the cluster patterns
- Through sensitivity testing, a total of 6 clusters (K=6) optimized forecast skill



Categorizing the circulation pattern via K-means clustering



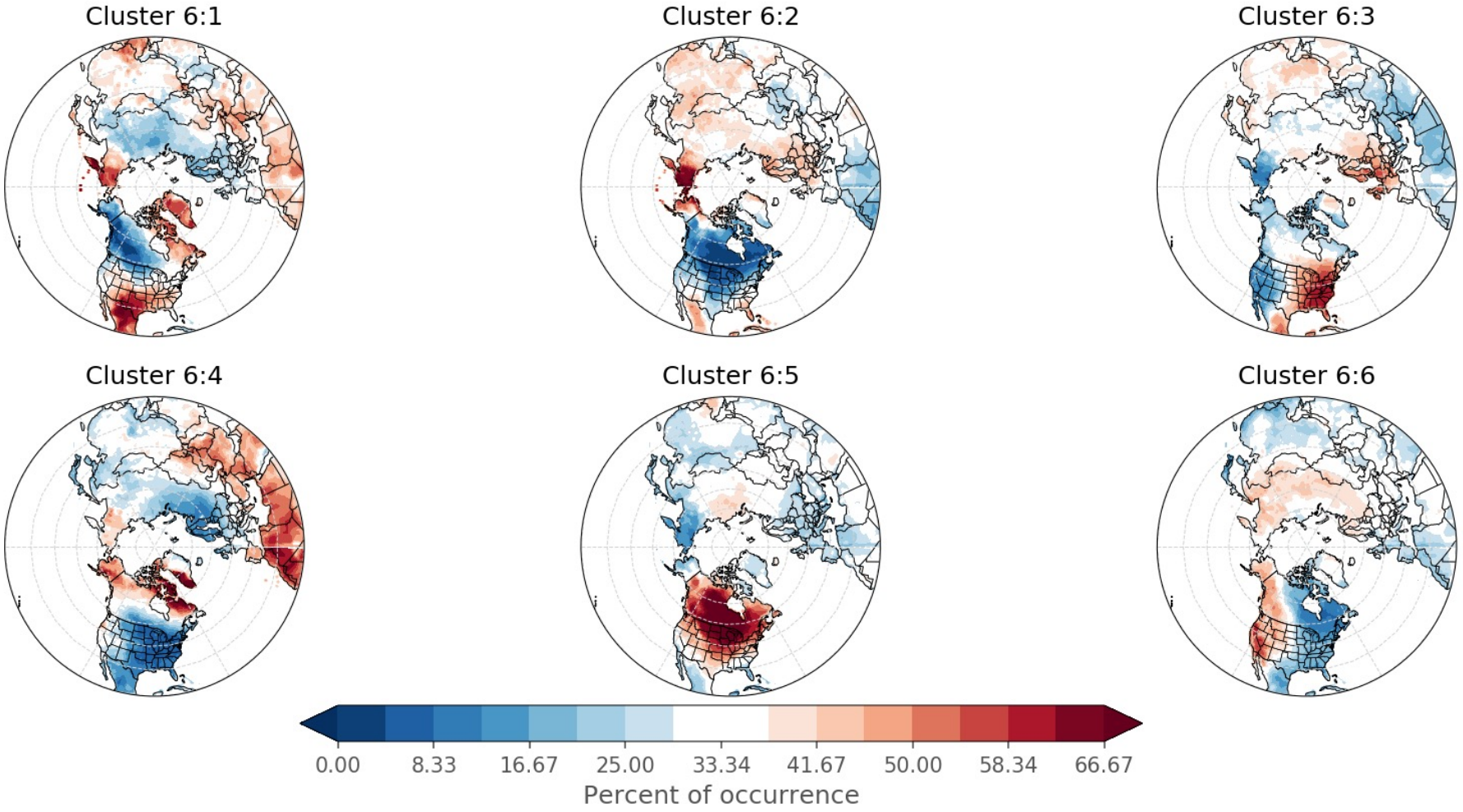
*No significance in the order of clusters!

Associated Cluster Composites

- Now that we have a cluster assignment for each 14-day period, we can find the associated anomaly composites for any variable. We first investigate temperature and precipitation terciles.
- Data:
  - Temperature: CPC 2-meter Daily Reanalysis
  - Precipitation: CPC Global Unified Gauge-based Daily Precipitation Reanalysis
  - 14-day running mean anomalies (temperature) or sum anomalies (precipitation) to match cluster periods
- Method:
  - Calculate terciles (33^{rd} and 67^{th}) for each running 14-day period
  - Smooth terciles (15 period smoother for temperature, 31 for precipitation)
  - Each period now can be classified as above, near, or below normal
  - For each of the 6 clusters, calculate the occurrence of each tercile
  - For example, a given point for Cluster 1 may have 70% occurrence of above, 20% near, and 10% below normal temperatures

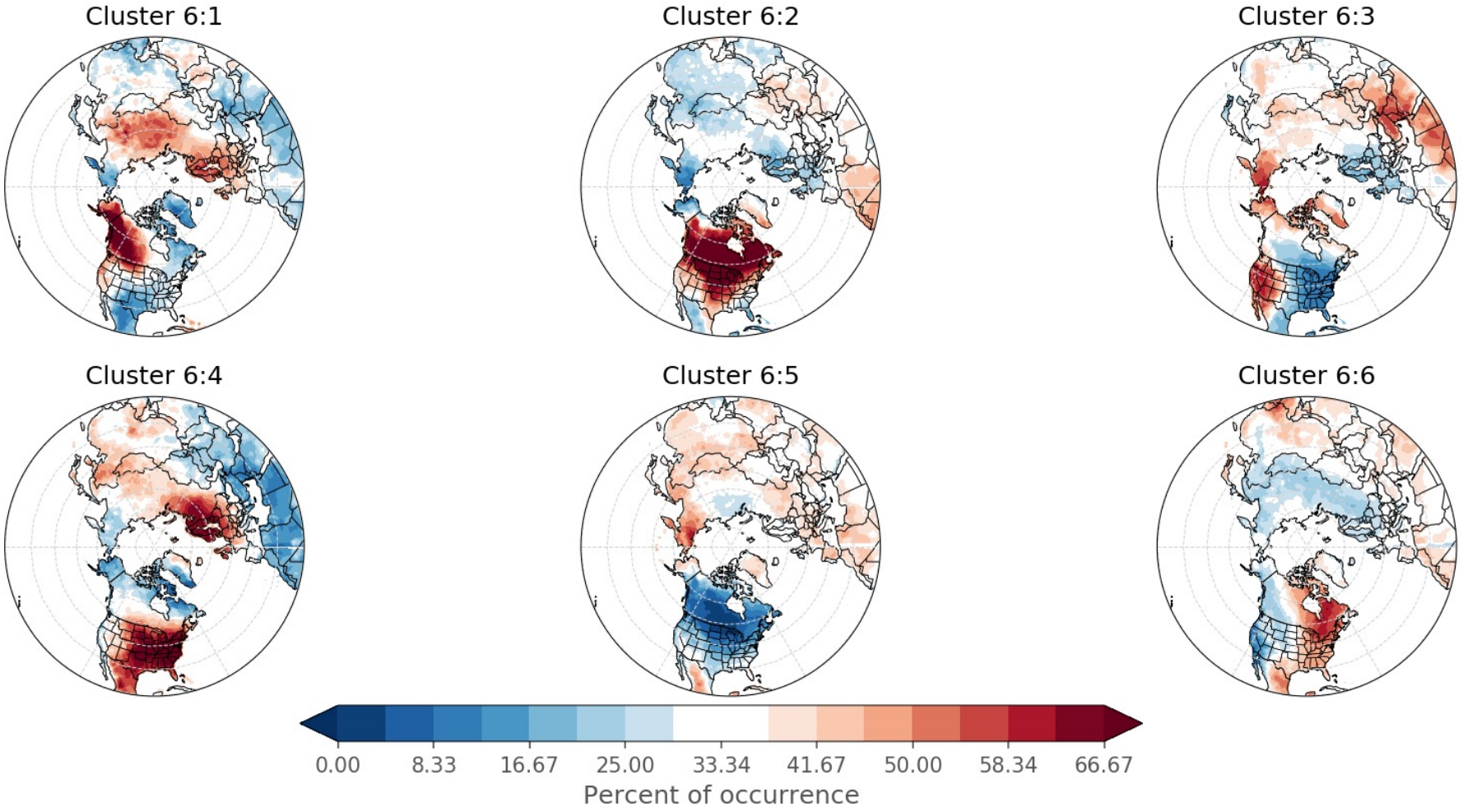
Associated Cluster Composites

Probability of **Above** Normal Temperature



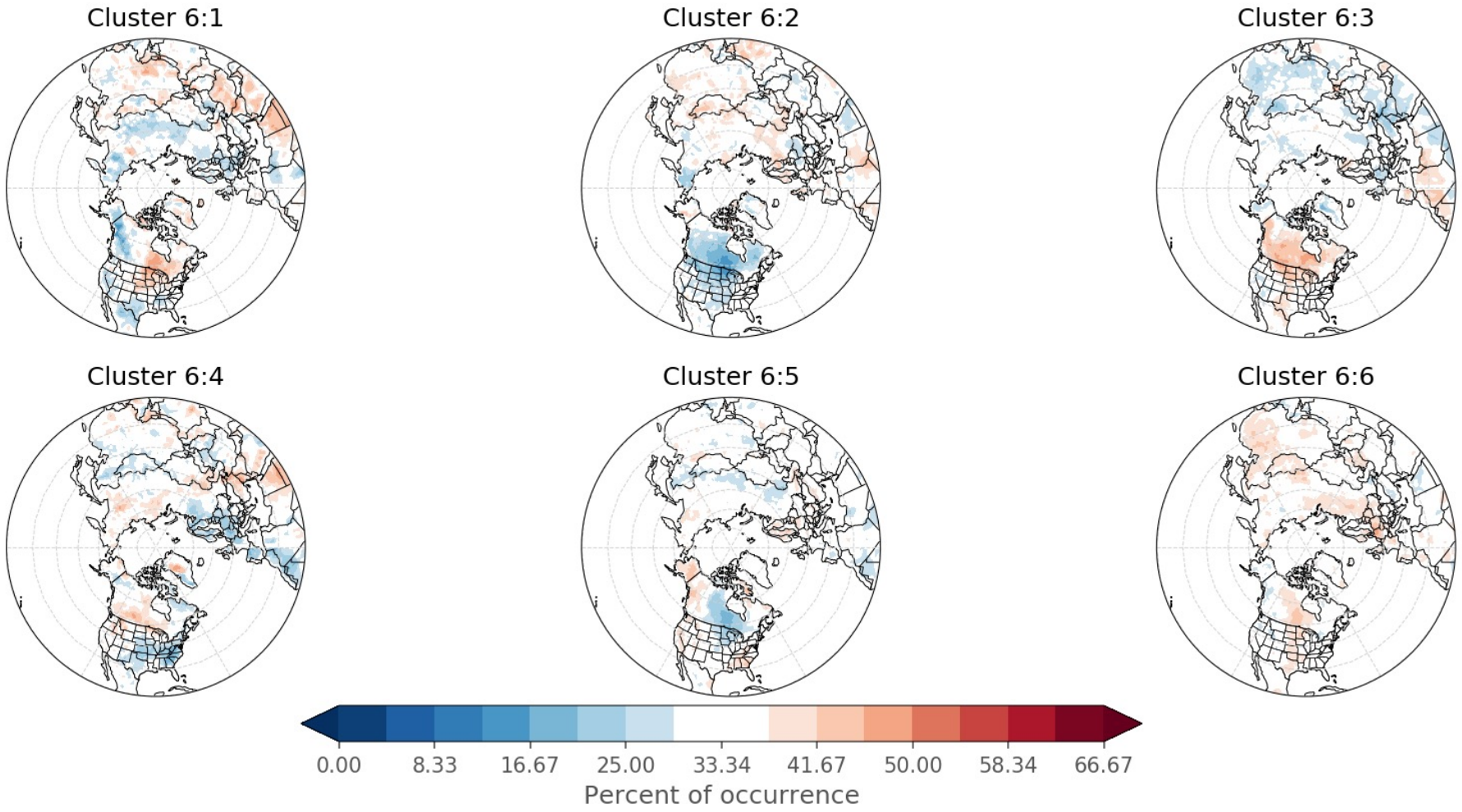
Associated Cluster Composites

Probability of **Below** Normal Temperature



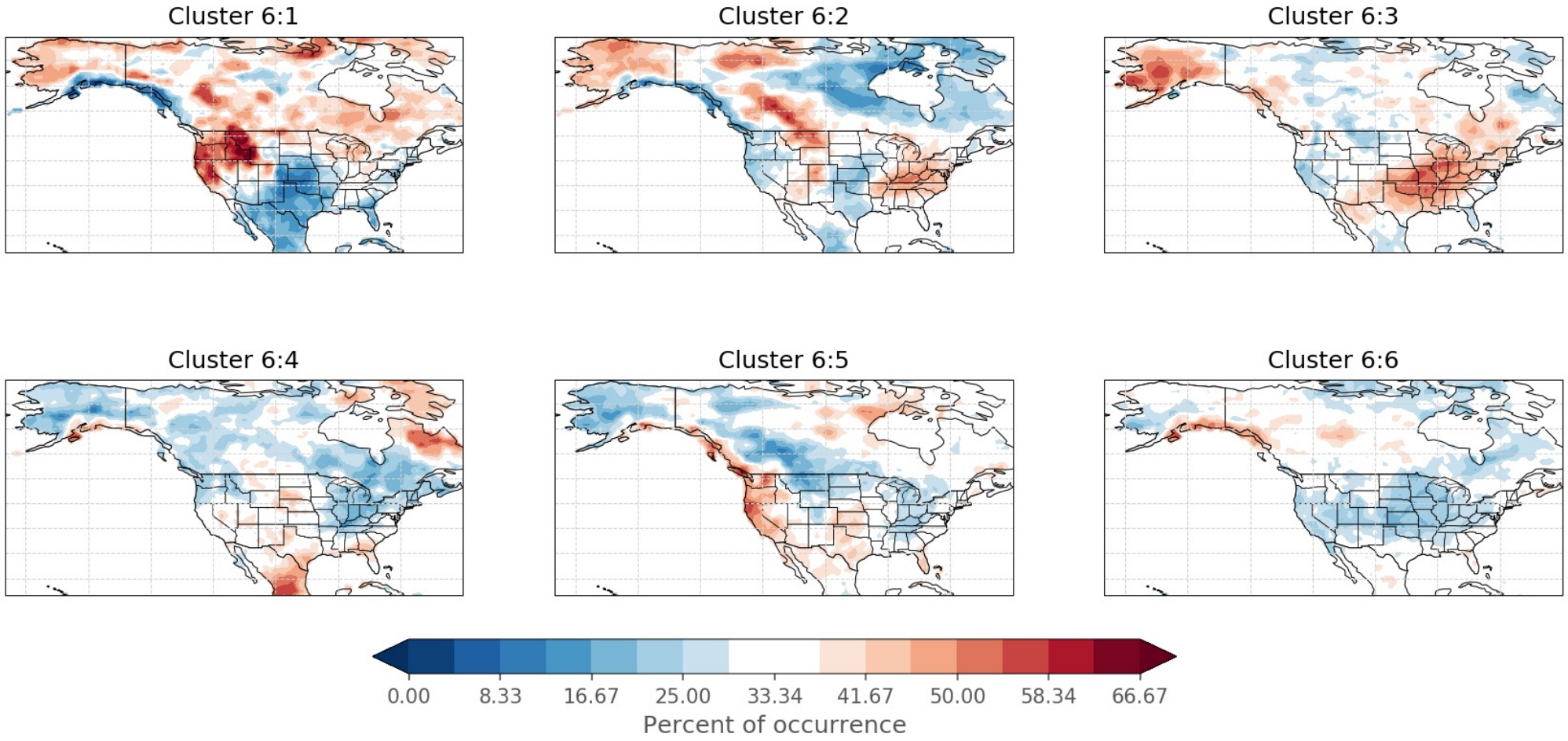
Associated Cluster Composites

Probability of **Near** Normal Temperature



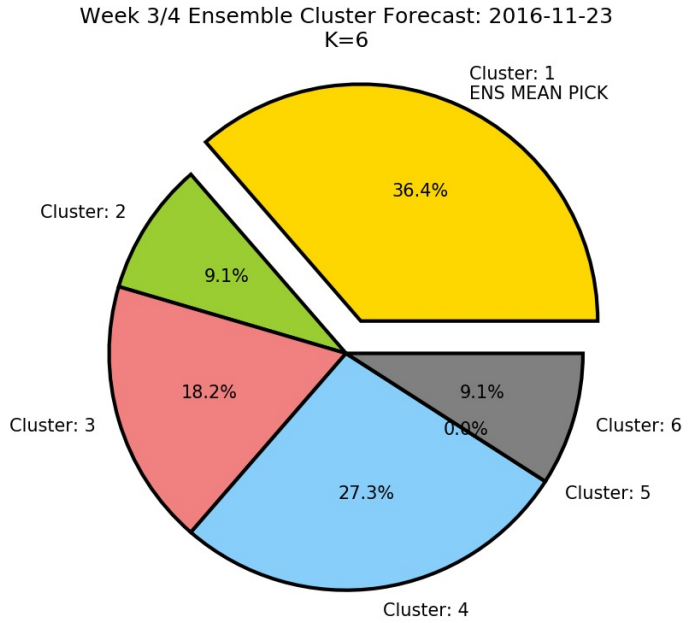
Associated Cluster Composites

Probability of **Above** Normal Precipitation

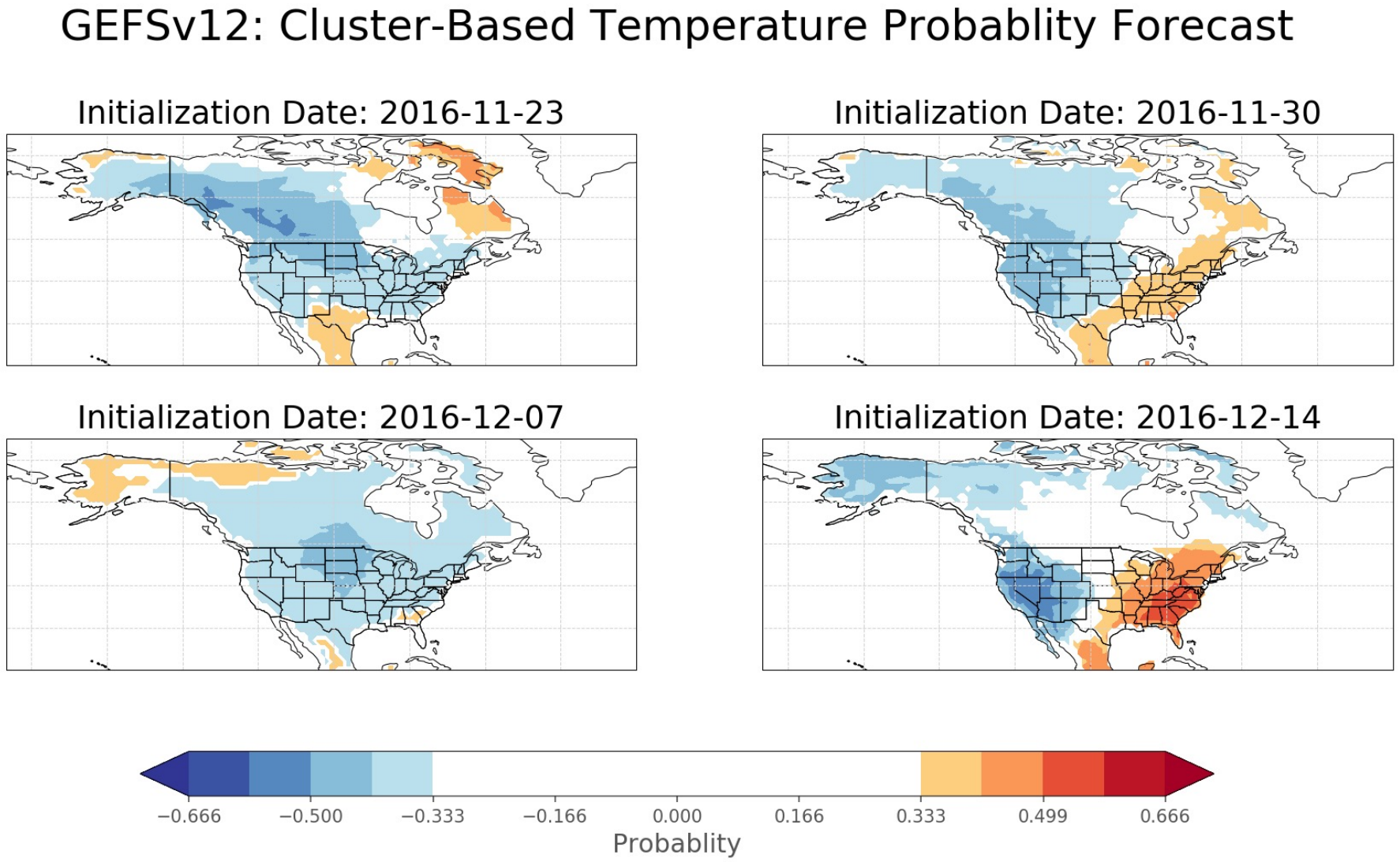


Cluster-Based Forecasts

- Method:
  - We can make forecasts using this analysis by assigning model ensemble members' Week 3/4 forecast to a cluster
  - Composite maps are weighed by the number of ensemble assignments and composited to give a final probabilistic forecast, comprising of chances of above, near and below average
- Data:
  - GEFSv12 hindcast z500 forecasts
  - 11 ensemble members
  - Weekly Initialized between 11/15-2/15
  - For years 2000-2019
  - 251 samples



Cluster-Based Forecasts



Cluster-Based Forecasts

- Initial skill scores:
  - Scoring all 251 GEFSv12 hindcast samples with the same observational datasets
  - Score using the Heidke Skill Score (HSS)
  - Scoring only the CONUS/AK region

Forecast Type	Forecast Category	Observation Category	Temperature	Precipitation
Maximum Probability Category	Above/Near/Below	Above/Near/Below	8.1	6.2
Maximum Probability Category	Above/Near/Below, Near=Ignore	Above/Below	23.6	10.0
GEFSv12 raw ENS Mean Anomaly	Above/Below	Above/Below	17.4	10.2

Conclusions

- The dominant circulation regimes over North America are combinations of teleconnections.
- These clusters have relationships to temperature and precipitation anomalies, and likely other fields.
- Based on hindcast testing, the cluster framework provides skillful forecasts, particularly when the z500 forecast is skillful and in regions of strong anomaly correlation.
- This framework is computationally efficient, and adaptable for any field.

Ongoing/Future Work

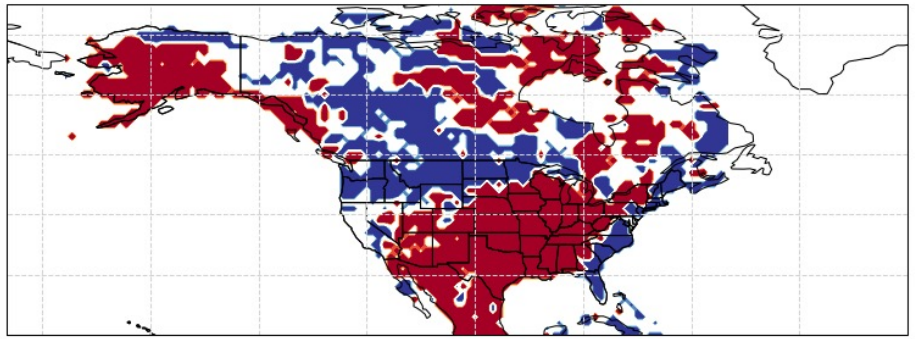
- Increase the number of ensembles used by including hindcasts from ECMWF, Subx, JMA, etc.
- An investigation into extremes and other fields (storm tracks, diabatic heating, etc.)
- Investigate forecasts of opportunity
- Realtime forecasts/scoring
- Prepare manuscript

Contact:
Greg Jennrich
Gregory.Jennrich@noaa.gov

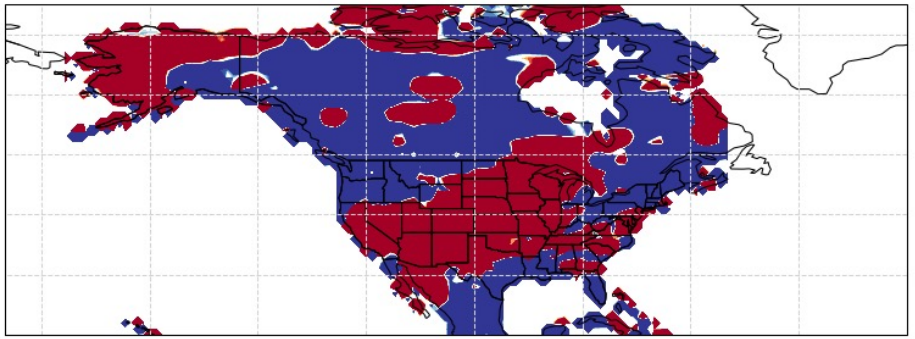
End slide

GEFSv12: Cluster Precip Cluster 2-cat Probablity Forecast

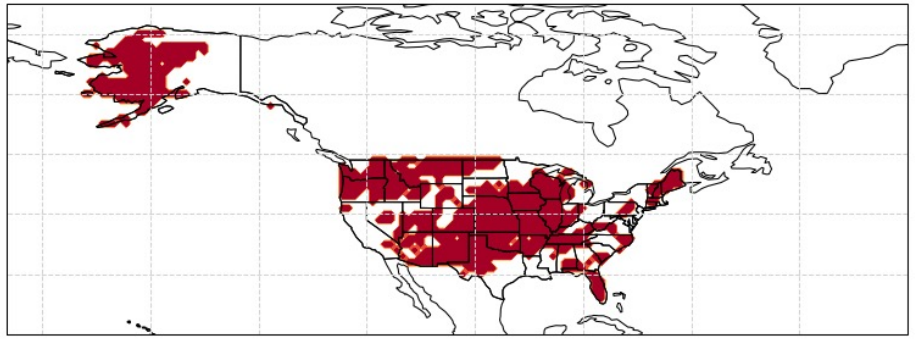
Forecast, HSS: 31



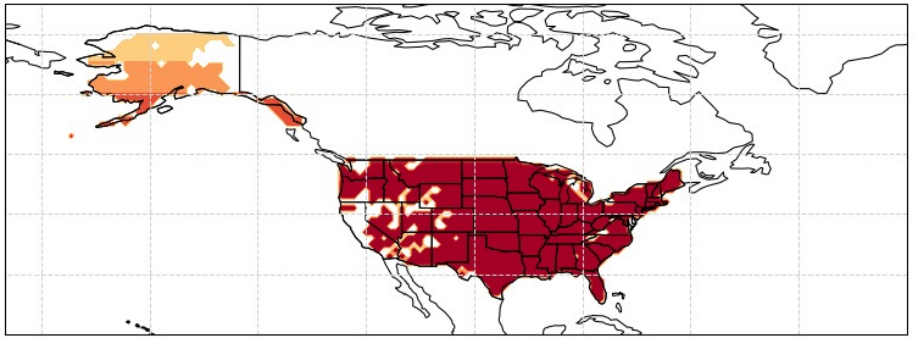
Observation

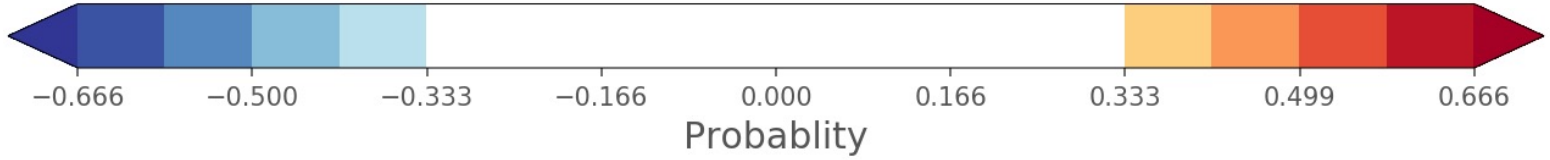


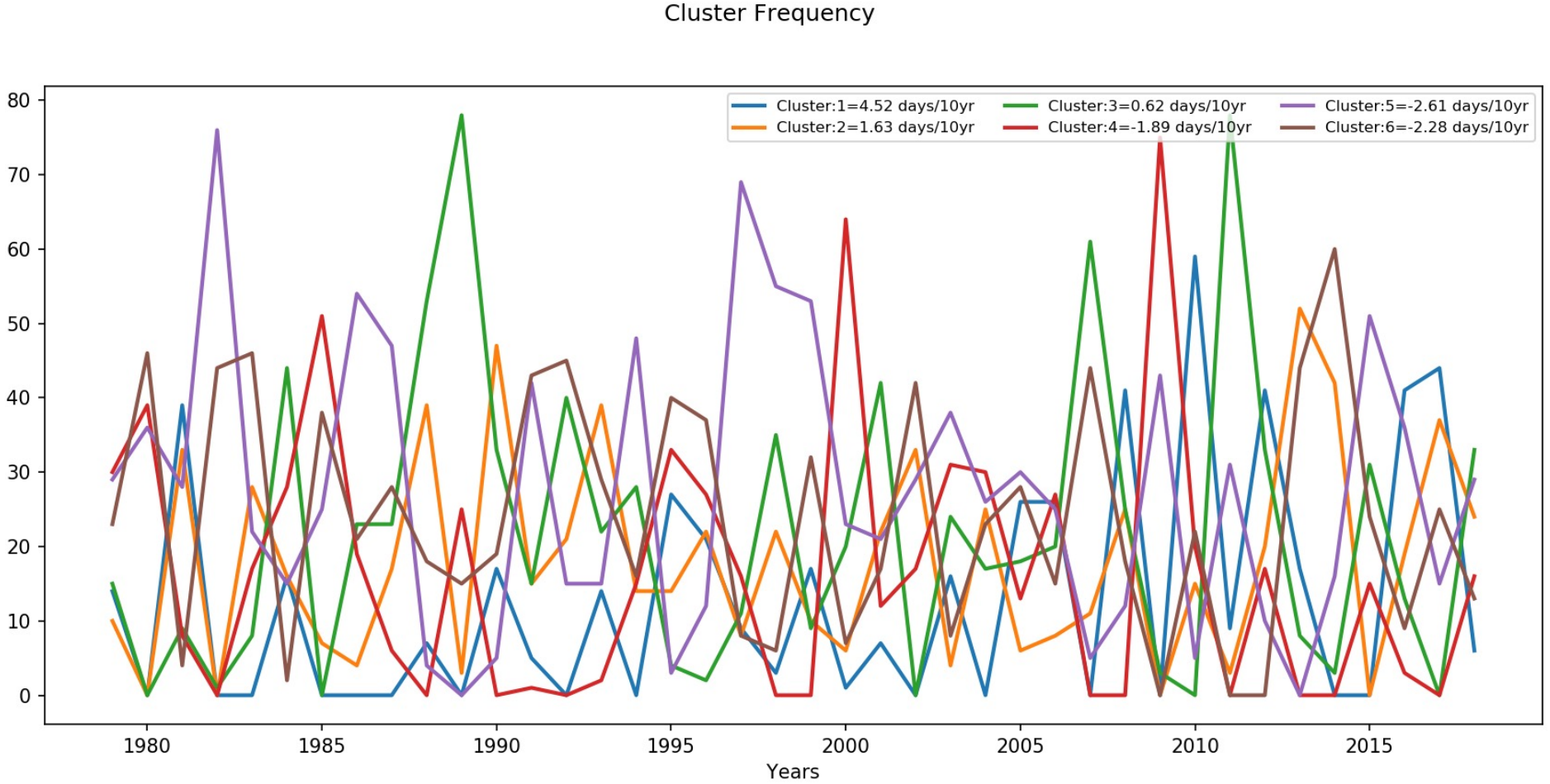
Correct Points



Cosine weighting for valid points



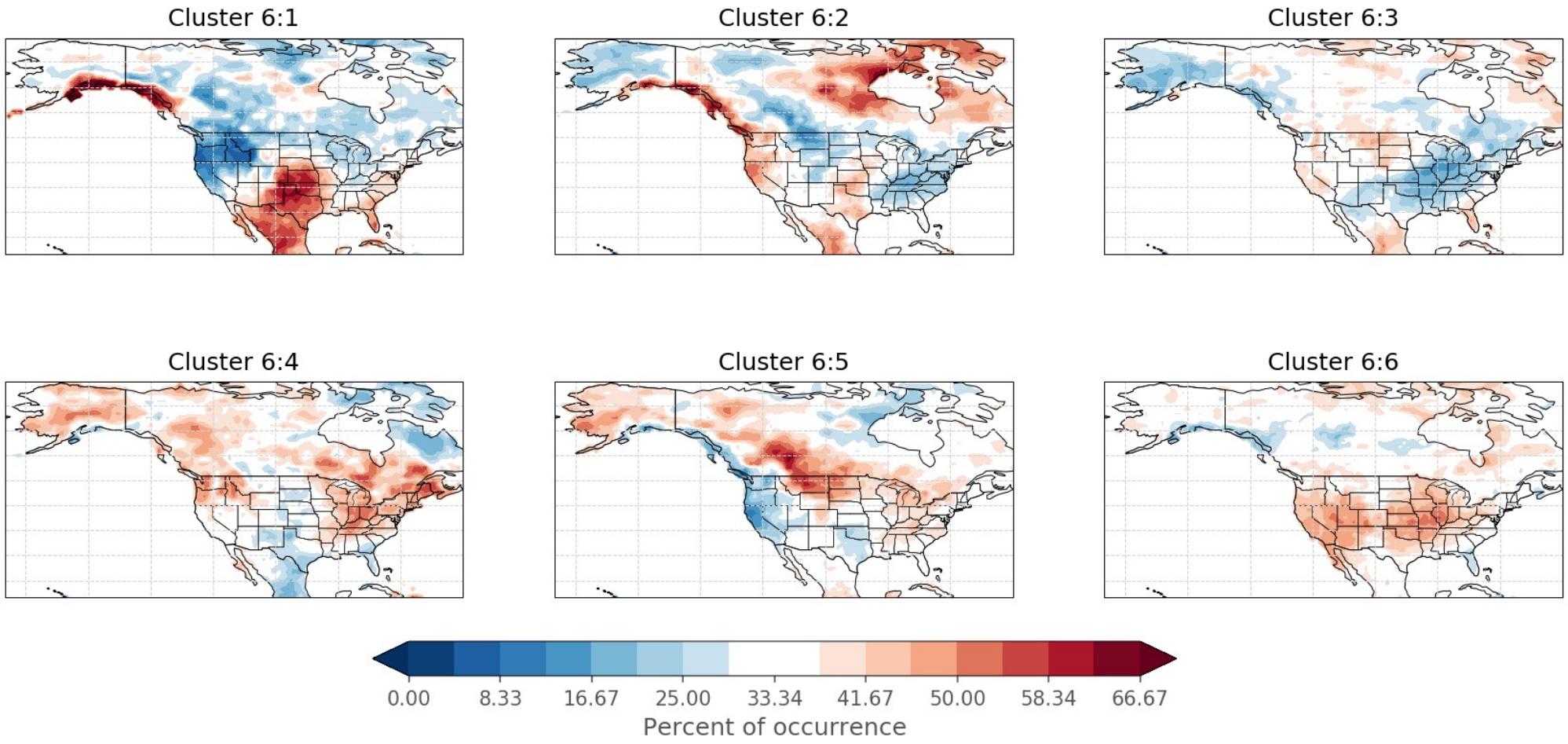




Cluster	1	2	3	4	5	6
Count	523	721	882	657	1098	959

Forecast period/model	Forecast Method	Z500	T850	t2m	Precip
Reanalysis as forecast (perfect forecast)	RMSE	38.4	37.6	36.1	15.5
	Correlation	39.2	39.6	38.1	15.1
GEFS Week 1/2	Ens. combo	34.8	33.7	31.2	13.2
	Ens. mean	28.0	28.7	27.5	11.0
GEFS Week 3/4	Ens. combo	9.3	11.2	11.1	6.1
	Ens. mean	8.6	8.5	8.9	8.0

Probablity of Below Normal Precip (1979-2018 percentiles)



Probablity of Near Normal Precip (1979-2018 percentiles)