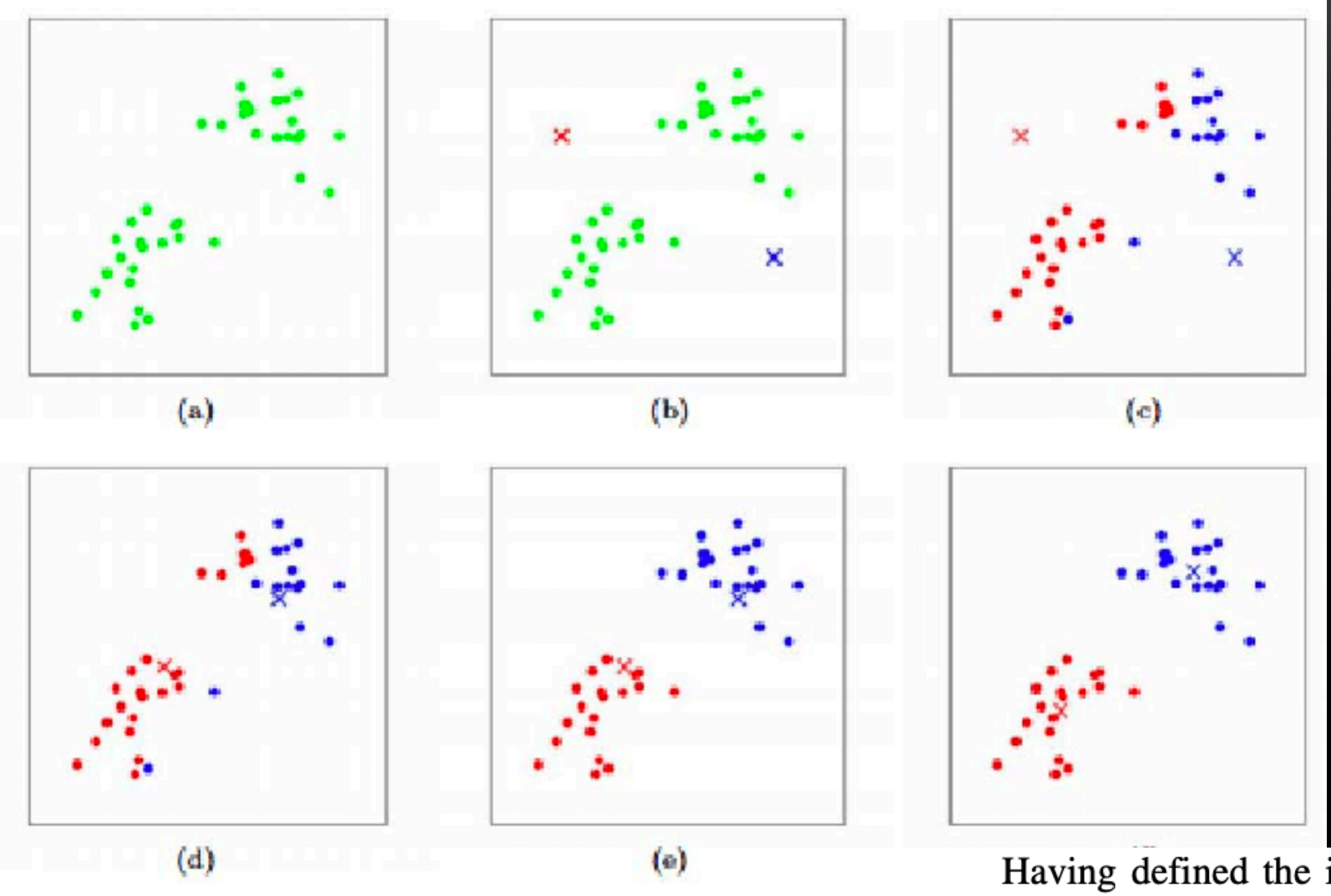
Weather Regimes: Part II

- What are weather regimes?
  - Interpretation of weather regimes
  - Implication for climate prediction
- How do we identify weather regimes?
- Application of weather regimes to S2S prediction

How do we identify weather regimes?

- Weather regimes are often identified using the k-means cluster analysis and daily mean geopotential height or SLP data:
  - The seasonal cycle is first removed
  - High-frequency variability is removed by applying a low-pass filter (such as taking the 5-day running mean) or/and by reconstructing the data with leading EOF modes. The latter is performed on the basis that higher-order EOF modes typically have smaller spatial scale and are often regarded as "noises". Dimension reduction will also reduce the computational cost of a large dataset.
  - Apply the K-means cluster analysis: the number of clusters needs to be prescribed.
  - Each data point (e.g., a daily H500 pattern) is assigned to a specific regime based on the nearest distances.
    - Or a "fuzzy" classification method can be adopted in which weather regimes may overlap and each particular daily weather map can be assigned a probability of belonging to one or another regime.

K-Means Cluster Analysis



https://www.researchgate.net/publication/334995028_Metamorphic_Exploration_of_an_Unsupervised_Clustering_Program

Wilks 2006

Having defined the initial membership of the G groups in some way, the K-means algorithm proceeds as follows:

1) Compute the centroids (i.e., vector means) $\overline{\mathbf{x}}_g$, $g = 1, \ldots, G$; for each cluster.
2) Calculate the distances between the current data vector $\boldsymbol{x}_i$ and each of the G $\overline{\mathbf{x}}_g$s. Usually Euclidean or Karl-Pearson distances are used, but distance can be defined by any measure that might be appropriate to the particular problem.
3) If $\boldsymbol{x}_i$ is already a member of the group whose mean is closest, repeat step 2 for $\boldsymbol{x}_{i+1}$ (or for $\boldsymbol{x}_1$, if $i = n$). Otherwise, reassign $\boldsymbol{x}_i$ to the group whose mean is closest, and return to step 1.

How many clusters?

- Optimal Ratio: the number of clusters can be determined by maximizing the optimal ratio

$$Optimal\ Ratio = \frac{Intercluster\ variance}{Intracluster\ variance},$$

  - the intercluster variance is the variance between the cluster centroids, and a large value implies that the clusters are well separated.
  - the intracluster variance refers to the average variance of the differences between a cluster centroid and the data points of that cluster, and a small value implies that data points tightly cluster around the associated centroid
  - A larges optimal ratio is thus associated with a more robust regime structure with tightly clustered data (Strommen et al. 2019).
- Bayesian information criterion (BIC):

$$BIC = \sum_{n}^{N} \underset{K}{\mathrm{argmin}} \left(\|X_K - x_n\|^2\right) + KD \cdot \log(N).$$

The first term is the intracluster variance, and the second term penalizes a large cluster number N (Dorrington and Strommen 2020).

- The elbow method: the intracluster variance (also known as distortion) is plotted as a function of cluster numbers. The turning point of the curve is chosen as the optimal cluster number.

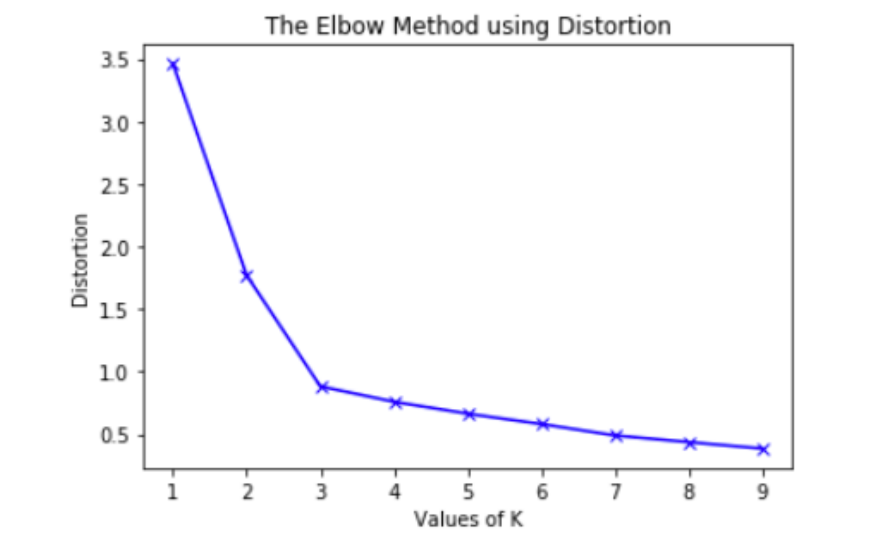
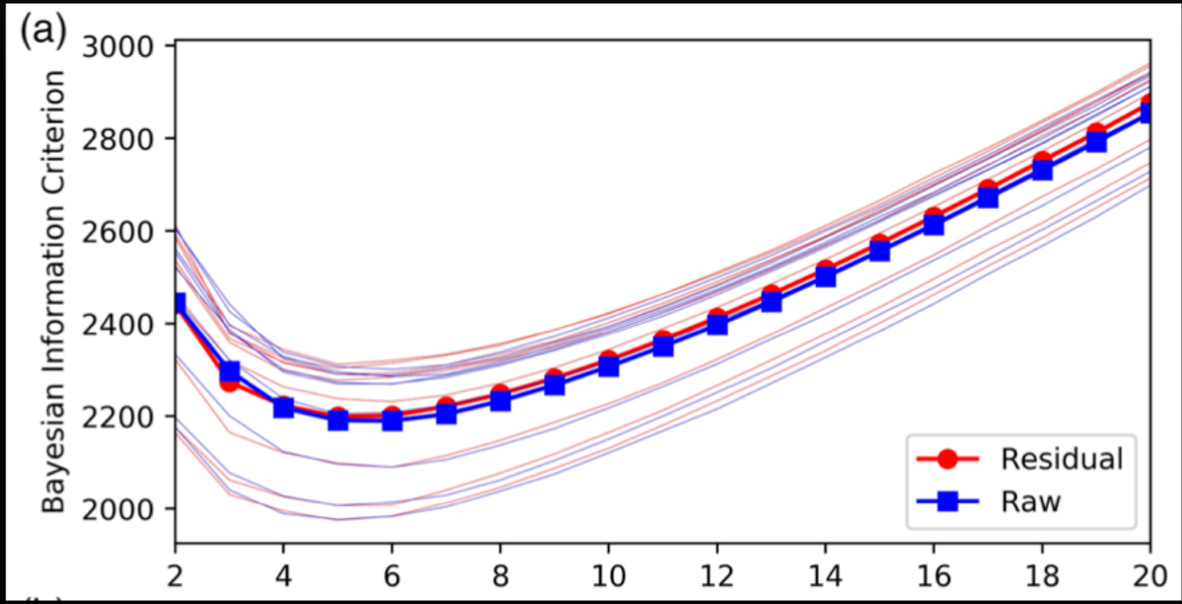


Figure from https://www.geeksforgeeks.org/elbow-method-for-optimal-value-of-k-in-kmeans/

How many clusters? (cont'd)

- A metric may not produce an unambiguous cluster number (see the figure on the right; 5 and 6 have similar BIC values and both look reasonable).
- Different metrics may yield different optimal cluster numbers.
- The "optimal" number may vary on the decadal time scale (Dorrington and Strommen 2020).

A larger number of weather regimes can reveal finer structure in the atmosphere's phase space but often at the cost of lower statistical confidence.



BIC for different numbers of K-means clusters (Dorrington and Strommen 2020)

How many EOFs in dimension reduction?

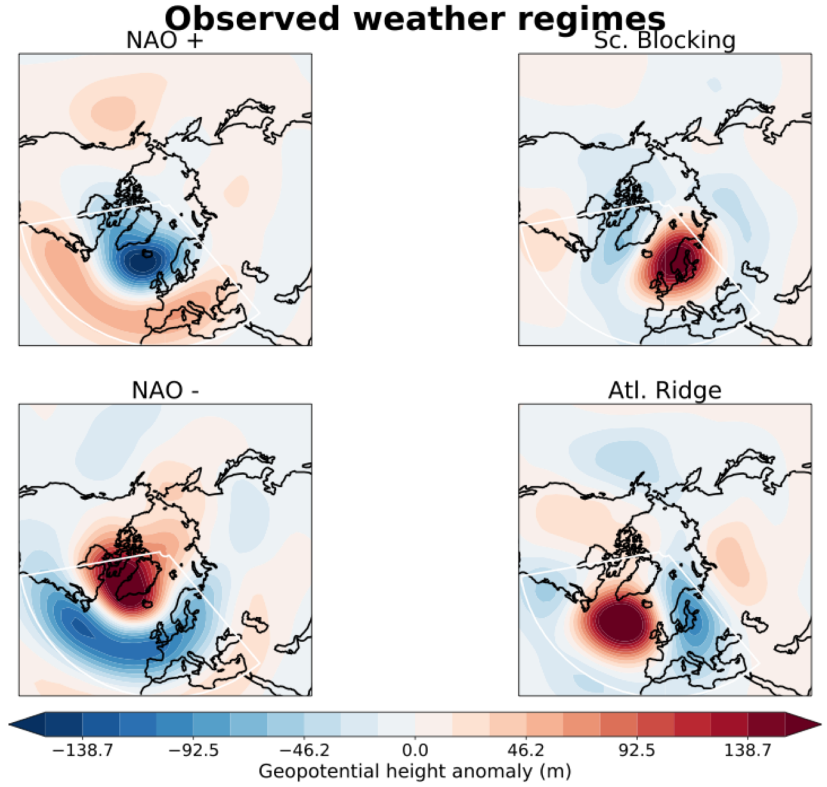


Figure 2. Spatial patterns of the four regimes defined by the cluster centroids for ERA-Interim (1979-2010), when using 10 EOFs in the phase space decomposition.

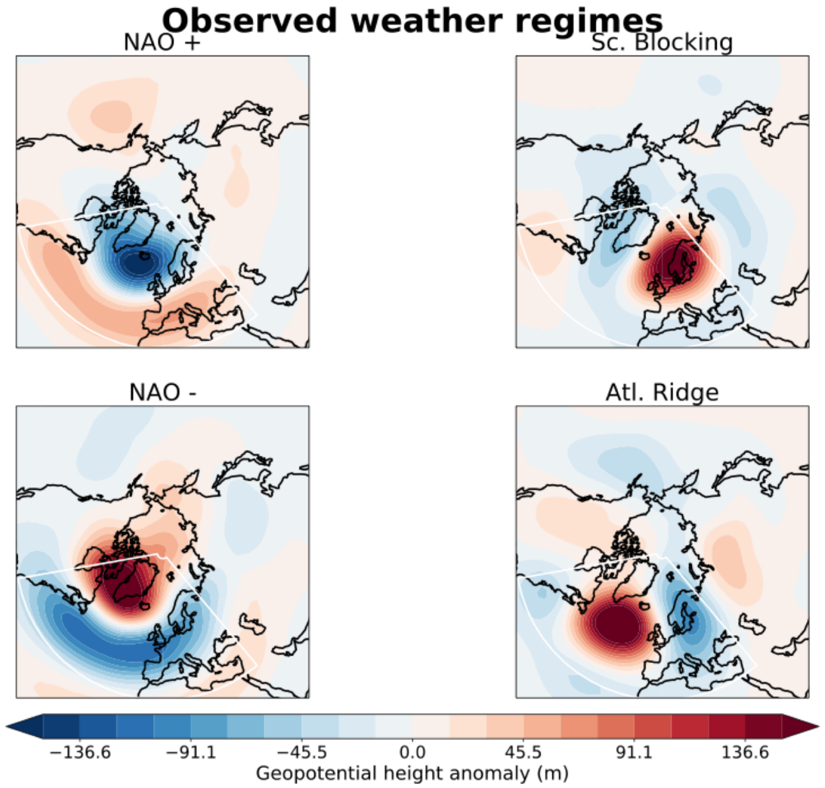


Figure 1. Spatial patterns of the four regimes defined by the cluster centroids for ERA-Interim (1979-2010), when using 4 EOFs in the phase space decomposition.

Strommen et al. 2019

- Cluster analysis is rather insensitive to small changes in the number of EOFs in dimension reduction (Strommen et al. 2019).
- Some studies suggested that the EOF-based dimension reduction is not necessary (Falkena et al. 2020)

Typical Weather Regimes over North America

Four regimes can be identified using daily 500-hPa geopotential height over the domain (10–70N, 150–40W)

- West coast ridge
- Greenland high
- Pacific trough
- Pacific ridge

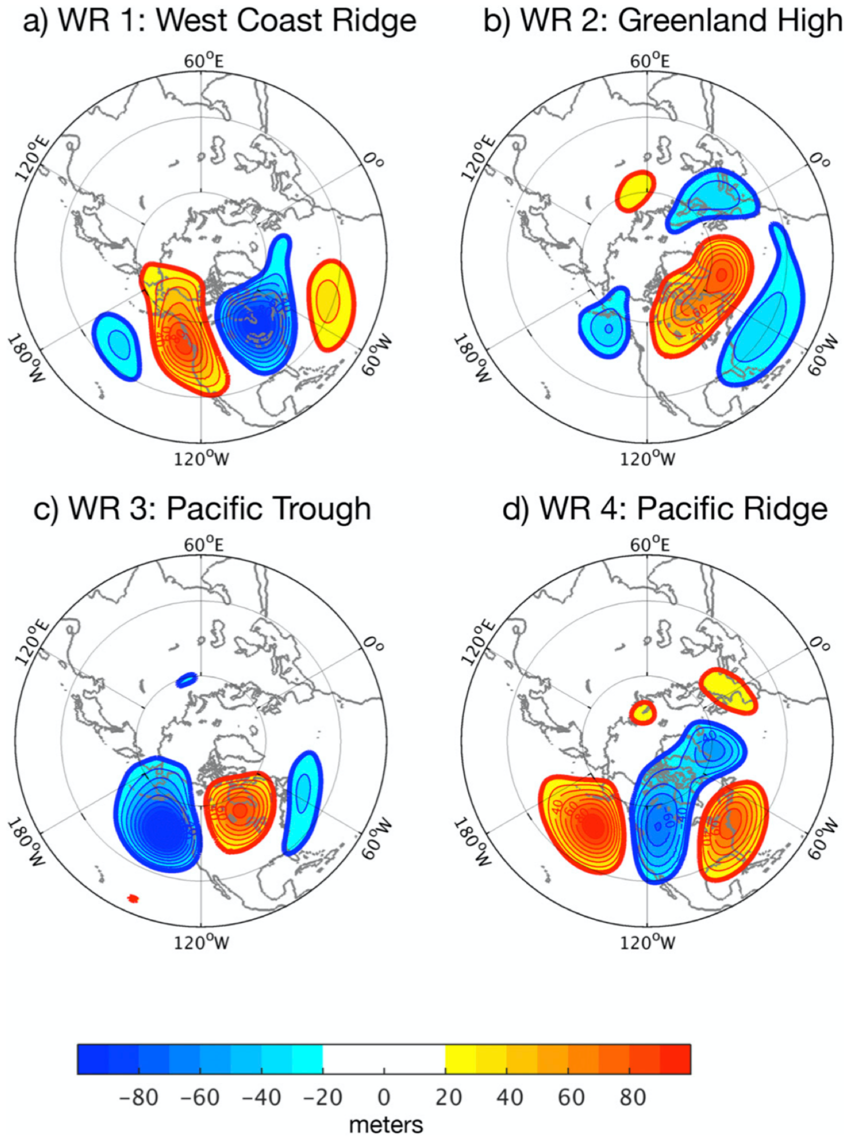


FIG. 1. Regime composites of 500-hPa geopotential height anomalies (in gpm), computed from MERRA reanalysis data. Shaded anomalies are significant at the 5% level of a Student's t test.

Impacts of Weather Regimes

- Weather regimes have strongly impacts on precipitation and near-surface temperature.
    - For example, fire risk is substantially increased in WR1 (west coast ridge) and reduced in WR3 (Pacific trough)

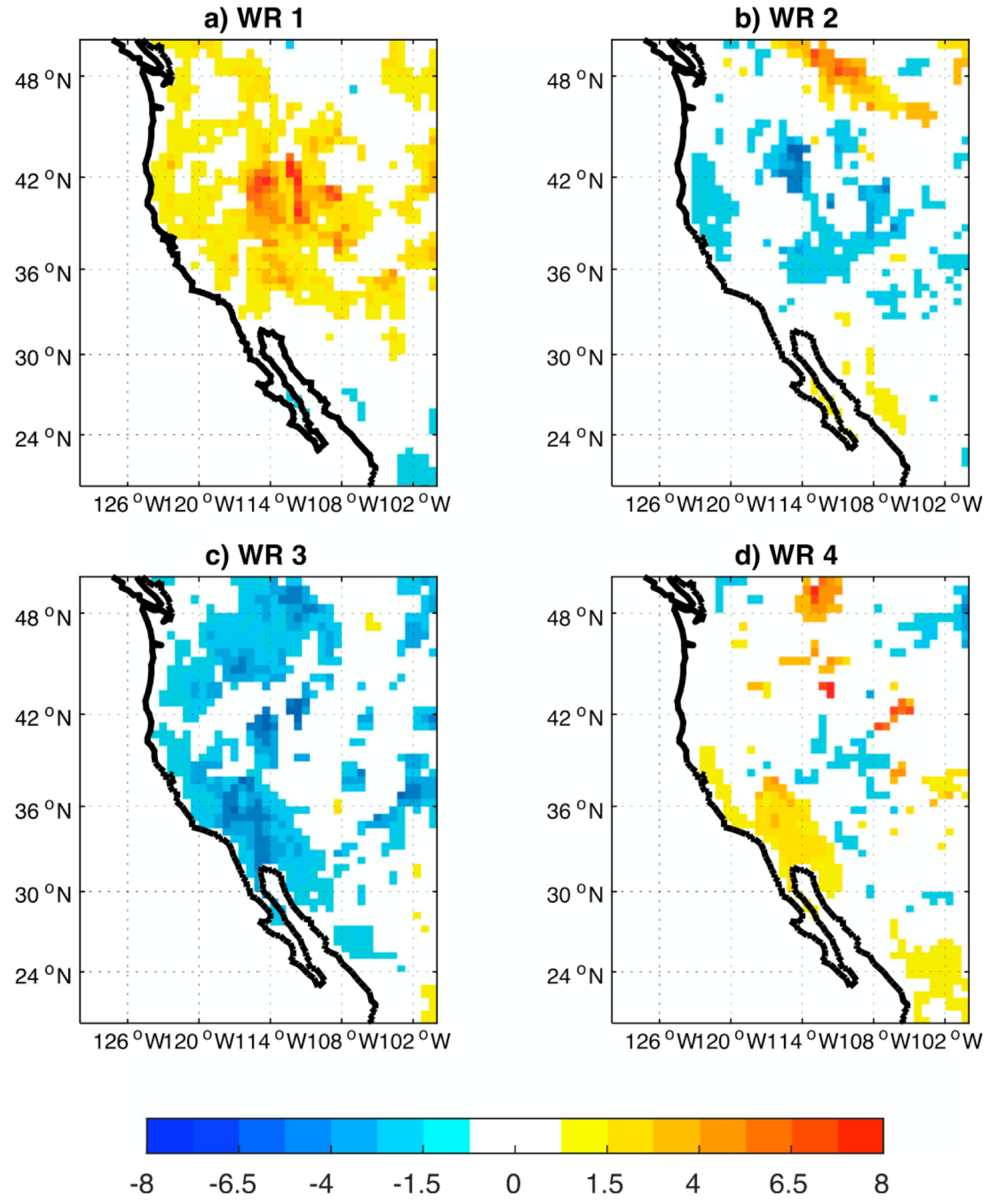


FIG. 2. Regime composites of fire weather index (dimensionless), expressed as deviations from the 1982–2014, October–March long-term average.

WRs in CFSv2 Hindcasts

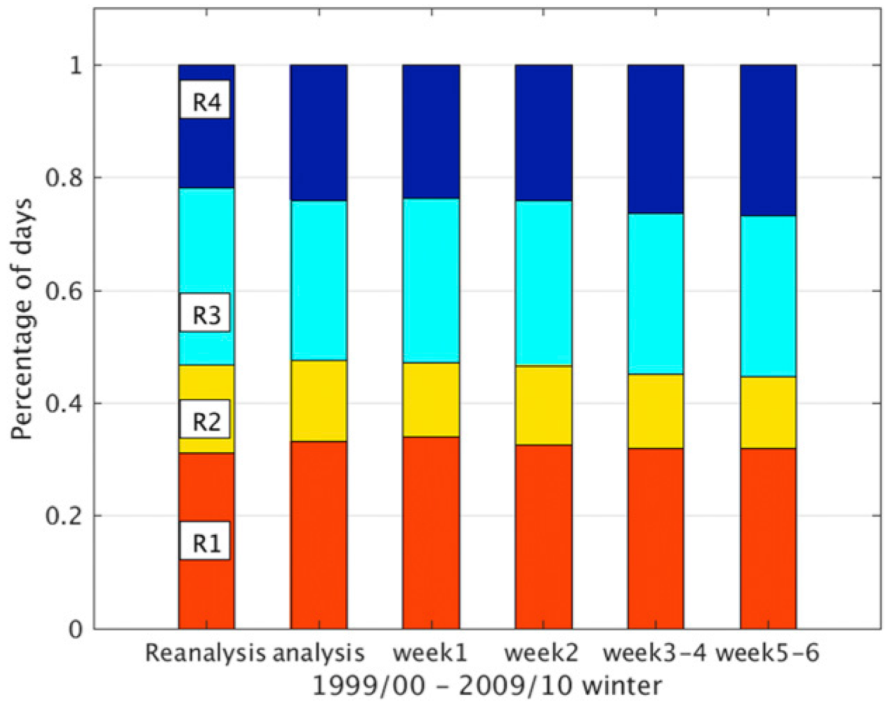


FIG. 3. Percentage of days assigned to each MERRA reanalysis regime (labeled R1 to R4) during the (left column) 1999–2010 reanalysis period, and for the 1999/00–2009/10 four-member ensemble mean CFSv2 hindcasts as a function of lead time.

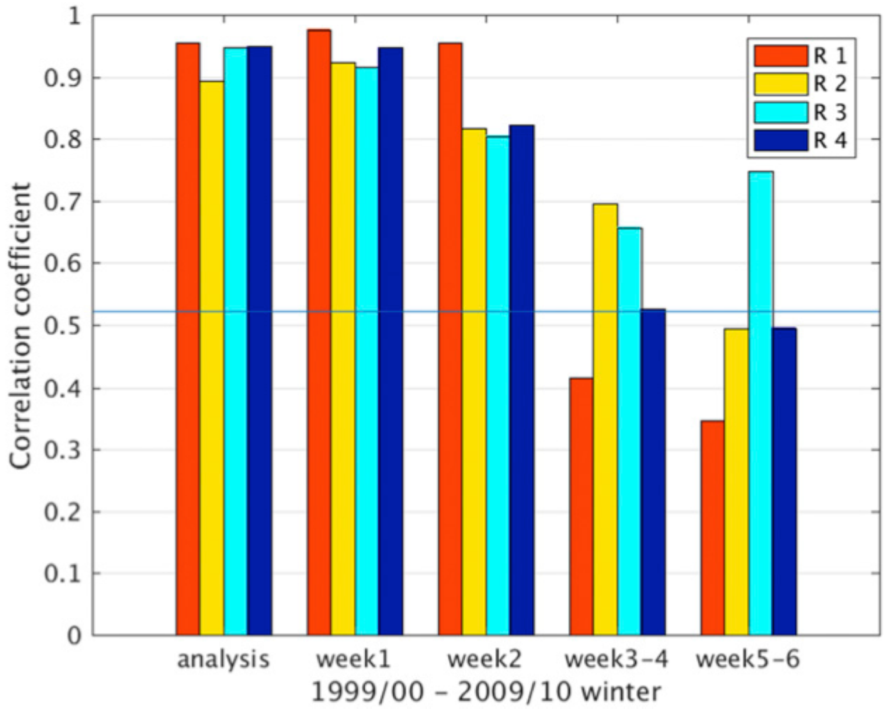


FIG. 4. Interannual correlations between the frequency of each regime in the CFSv2 hindcasts vs the MERRA reanalysis, as a function of lead time. The analysis refers to the hindcast initial condition (lead zero). The $r = 0.521$ 95% significance threshold is indicted.

- The long-term mean frequency of occurrence of WRs are well represented in CFSv2, especially for week 1&2
- CFSv2's can capture seasonal mean frequency of WRs with lead times up to two weeks, but a substantial drop of skill occurs from week 2 to week 3-4.

Prediction Skill of Weather Regime

ACC of 5-day moving means of regime counts, between MERRA and the forecast ensemble-mean

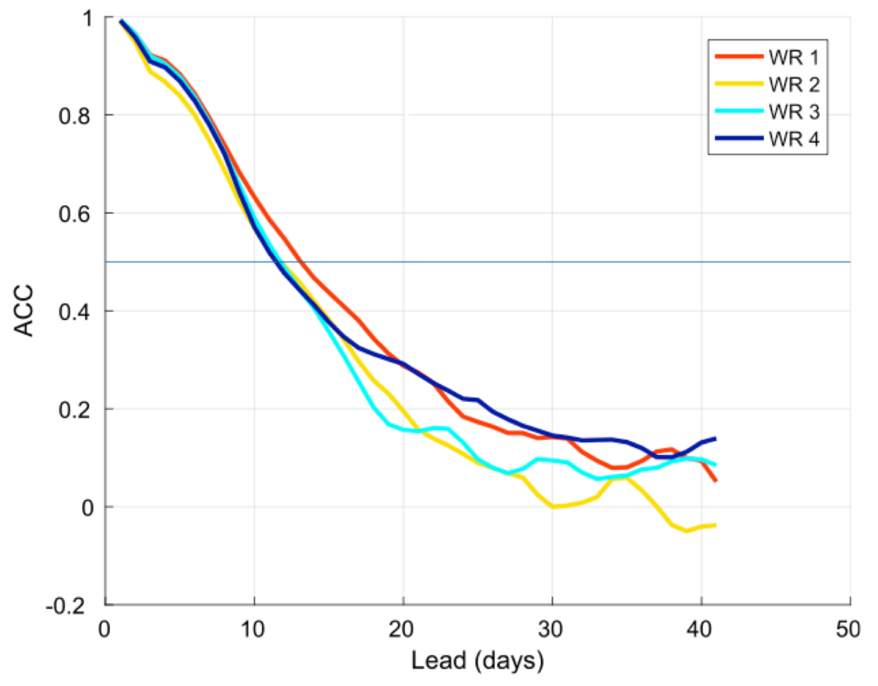
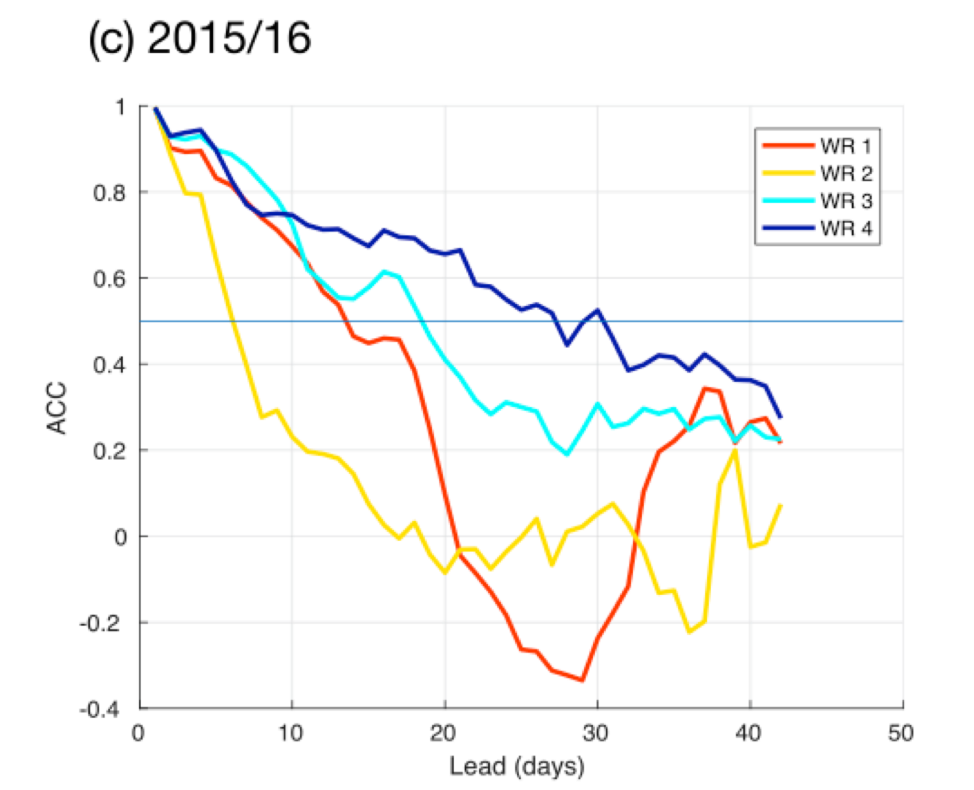


FIG. 12. Anomaly correlation skill for each regime, based on the 1999–2010 hindcast period. The 0.5 correlation value is indicated.



- When averaged over 1999-2010, hindcast skill of regime frequency is limited to 10–15 days.
- In some years and for some regimes, skillful prediction can be achieved up to ~30 days ahead. This represents "forecasts of opportunity" when subseasonal forecast skill is much higher than the average.

References

1. Corti, S., Molteni, F. & Palmer, T. Signature of recent climate change in frequencies of natural atmospheric circulation regimes. Nature 398, 799–802 (1999). https://doi.org/10.1038/19745
2. Dorrington, J., & Strommen, K. J. (2020). Jet speed variability obscures Euro-Atlantic regime structure. *Geophysical Research Letters*, *47*, e2020GL087907. https://doi.org/ 10.1029/2020GL087907
3. Ghil, M., & Robertson, A. W. (2002). "Waves" vs. "particles" in the atmosphere's phase space: A pathway to long-range forecasting? Proceedings of the National Academy of Sciences of the United States of America, 99(SUPPL. 1), 2493–2500. https://doi.org/10.1073/ pnas.01258089
4. Plaut, G., & Vautard, R. (1994). Spells of Low-Frequency Oscillations and Weather Regimes in the Northern Hemisphere, Journal of Atmospheric Sciences, 51(2), 210-236
5. Robertson, A. W., Vigaud, N., Yuan, J., & Tippett, M. K. (2020). Toward Identifying Subseasonal Forecasts of Opportunity Using North American Weather Regimes, Monthly Weather Review, 148(5), 1861-1875.
6. Strommen, K., & Palmer, T. N. (2019). Signal and noise in regime systems: A hypothesis on the predictability of the North Atlantic Oscillation. *Quarterly Journal of the Royal Meteorological Society*, *145*, 147–163. https://doi.org/10.1002/qj.3414
7. Wang, B., Barcilon, A., & Fang, Z. (1999). Stochastic Dynamics of El Niño–Southern Oscillation, Journal of the Atmospheric Sciences, 56(1), 5-23.Wilks, 2011, "Statistical Methods in the Atmospheric Sciences", Chapter 7