Clustering Analysis of Autumn Weather Regimes in the Northeast U.S.

David Coe[1], Mathew Barlow[1], Laurie Agel[1], Frank Colby[1], Christopher Skinner[1], Jian-Hua Qian[2]

[1]University of Massachusetts Lowell, [2]Centre for Climate Research Singapore (CCRS)

1. Motivation and Method

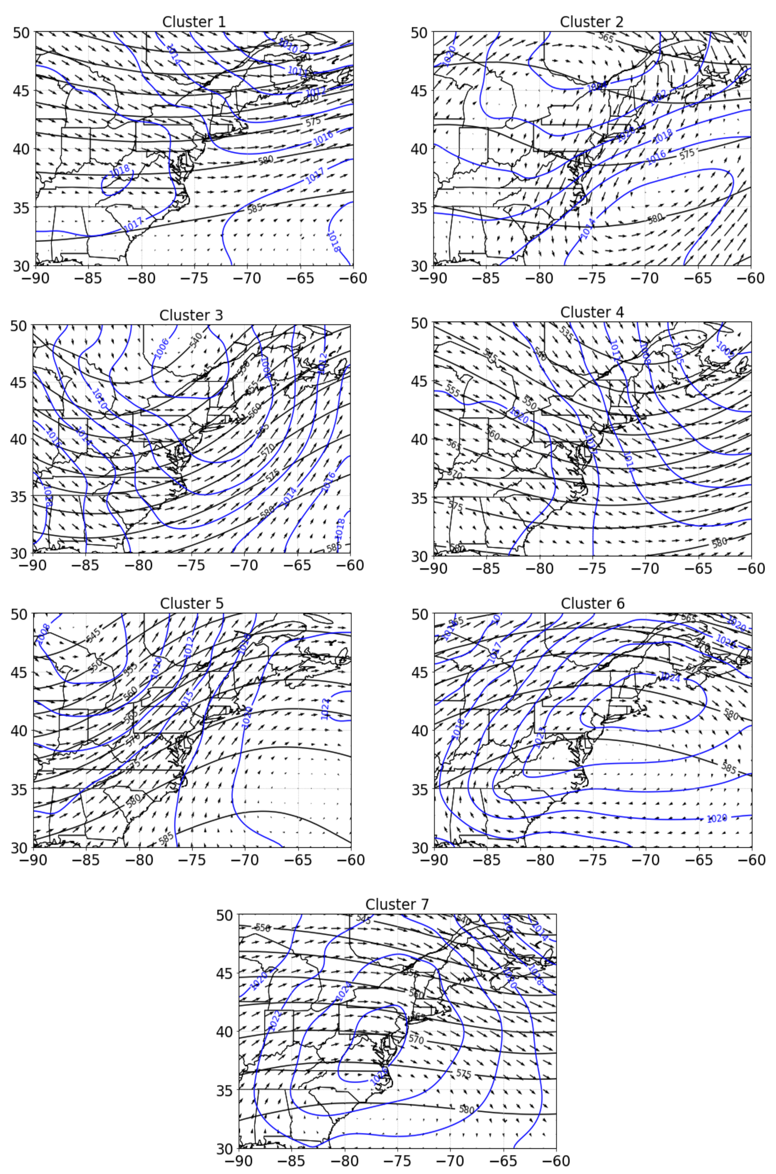
A k-means clustering method is applied to daily ERA5 500 hPa height, MSLP, and 850 hPa wind data to identify characteristic Weather Types (WTs) for Sep-Nov from 1979-2018 over the Northeast U.S. The resulting circulation clusters are analyzed in terms of their frequency of occurrence throughout the season, typical progressions between clusters, and whether differences between early and late fall were discernible.

Here we performed an analysis of Autumn Season circulation patterns in the Northeast U.S. through a K-Means Cluster analysis in section 2, monthly variations in frequency of clusters in section 3, identification of early and late season clusters in section 4, and a look at a preliminary time series analysis on the clusters over the 40 year study period in section 5.

The data features the seasonal mean removed rather than the daily mean. Both daily mean and seasonal mean removed datasets provided a similar number of clusters, 7 each, and showed similar early and late season patterns. However, the patterns were more clearly defined in the seasonal mean removed data and showed stronger correlations.

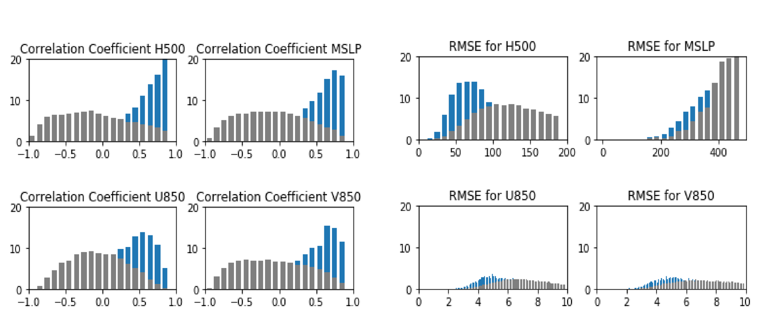
2. K-Means Clustering

A k-means clustering method is applied to daily ERA5 500 hPa height, MSLP, and 850 hPa wind data to identify characteristic Weather Types (WTs) for Sep-Nov from 1979-2018 over the Northeast U.S. The results for a 7-cluster solution are shown to the right. Below these results are figures showing the Correlation Coefficient and Root Mean Square Error analysis of individual days to their WTs.



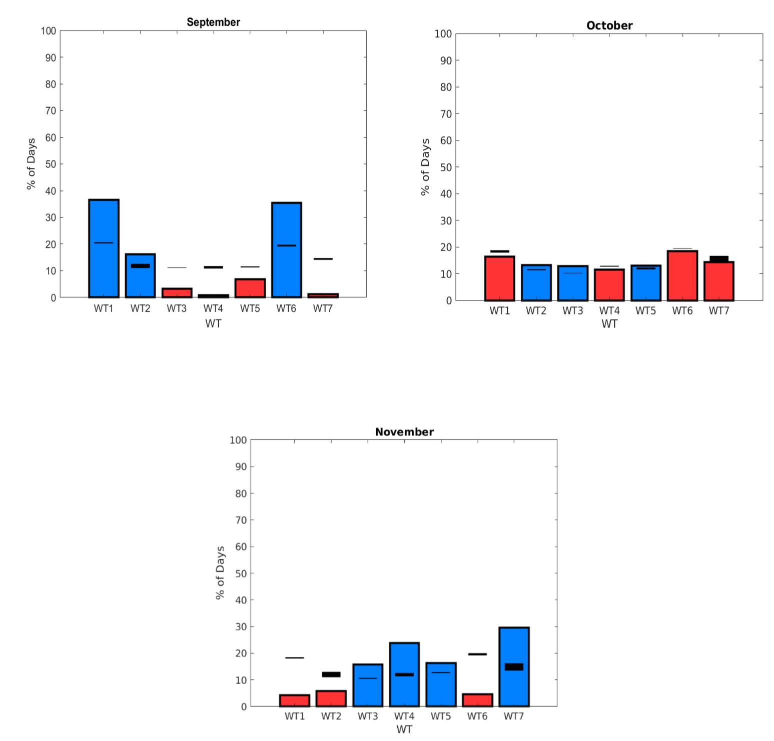
Each WT is a representative circulation pattern. WTs 1 and 2 feature a weak upper-level trough over the Northeast. WTs 3 and 4 have a strong upper-level trough over the Northeast. WTs 5 and 6 feature a strong upper-level ridge over the Northeast WT7 shows a weak upper-level ridge over the Northeast.

The two statistical analyses presented to the right represent how well each day was matched into its respective WT. Blue bars represent days within their respective WT, while grey bars indicate days to all other WTs.



3. Variations of WT Frequency Throughout the Season

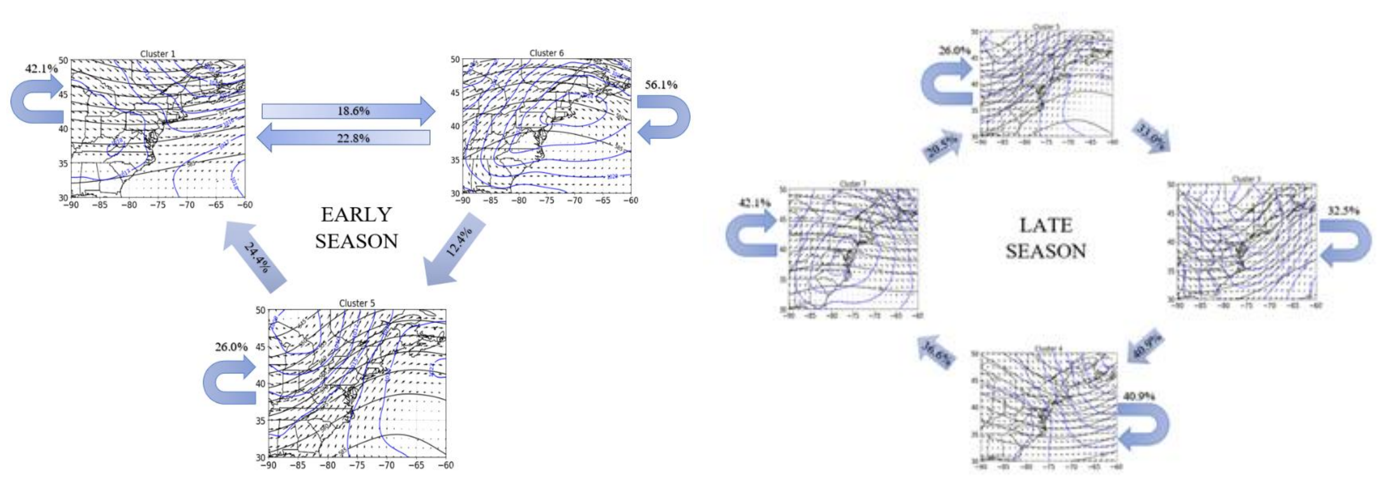
The monthly likelihood of occurrence of each WT is presented to the right. A Monte Carlo analysis is used to determine the 95% confidence interval (black bars). Blue and red shaded bars indicate statistical significance with likely to occur in Blue and unlikely to occur in Red.



Note that multiple WTs see a decrease/increase in frequency throughout the season, demonstrating the seasonality of the WTs.

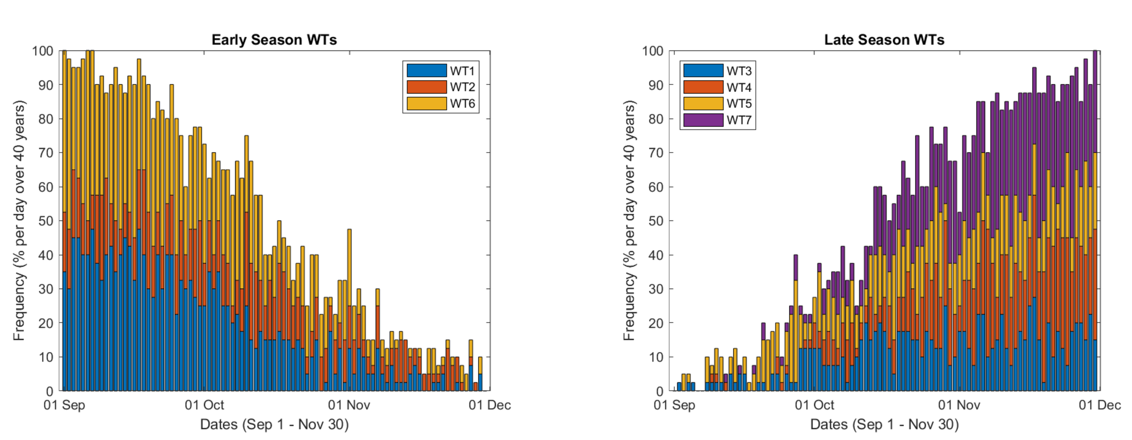
4. Autumn Early Season vs Late Season

Progression here is defined as which WT was likely to occur the next day. As weather systems move through the region, they are represented by a sequence of WTs. Results are shown to the right with arrows showing significant values of next day progression. Two main sequences emerged containing 6 of the 7 WTs (WT2 showed no statistically-significant progression to any of the WTs on the next day).

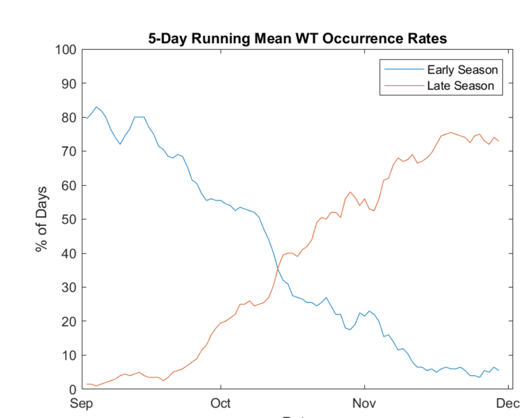


Based on the monthly frequency of the WTs, they can be separated into "Early Season" (WT1, WT2, WT6) and "Late Season" (WT3, WT4, WT5, WT7).

Averaging the daily frequency over all years for Early Season (left figure) and Late Season (right figure), we found that October 16th is the date of the average shift from Early to Late Season WTs.

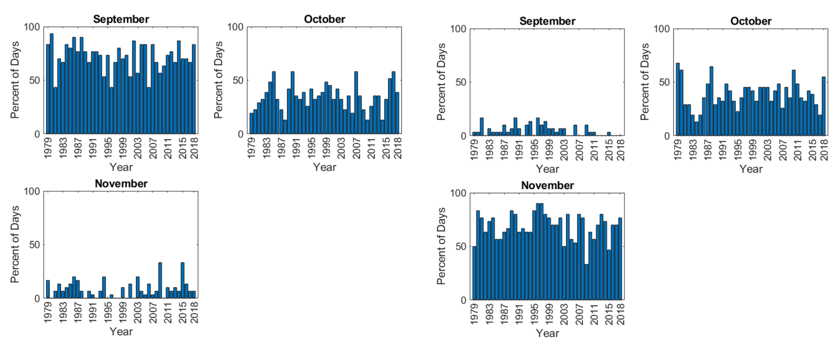


Note that overall, Late Season WTs are unlikely to occur in September and Early Season WTs are unlikely to occur in November (< 10% chance of occurrence for each day of the month)



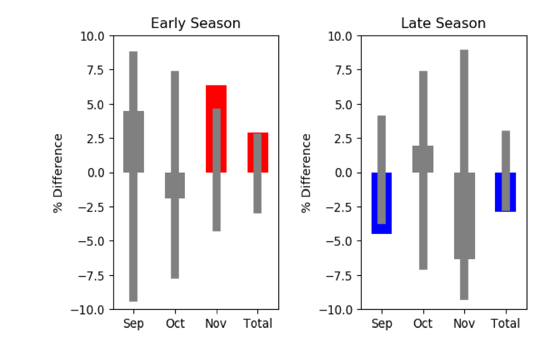
5. Preliminary Time Series Analysis of Autumn WTs

After identifying Early and Late season WTs, they can be used their evolution within the Fall season for the 40-year time period of the study. Here we show a preliminary analysis, which will be the basis of further study.



While there is ample evidence of changes in the timing of the season in terms of temperature, the current approach allows a consideration of changes in the timing of daily circulation patterns. Comparing the first 20 years of the period with the last 20 years, there is some suggestion that the daily circulations are changing as well: the early season WTs have increased in November by 36% and the late season WTs have decreased in September by 46% (significance of 95% via Monte Carlo).

Difference in Monthly Occurrence from 1979-1998 to 1999-2018



6. Discussion

7 distinct WTs were identified for the autumn season in the Northeast U.S through use of the k-means clustering algorithm applied to ERA5 500 hPa height, MSLP, and 850 hPa u- and v- wind data from 1979-2018.

An analysis of WT frequency per month showed seasonality between the WTs with 3 WTs (WTs 1, 2, and 6) likely to occur in September and not in November and 4 WTs (WTs 3,4,5, and 7) likely to occur in November and not in September.

The WTs showed characteristic progressions between each other. Their frequency of occurrence varies by months, allowing a definition of "Early Season" WTs "Late Season" WTs.

Preliminarily, there is some suggestion that the timing of the season itself is changing in terms of these daily circulation patterns. Over the period considered, there has been a statistically-significant increase of early season WTs in the last month of the season, and decrease of late season WTs in the first month of the season. That is, in addition to the well-known warming of the fall season, there appears to be a corresponding circulation change.

7. Acknowledgements

This work was supported by NASA NESSF grant NNX15AN91H S03 and NSF grant 1623912 as well as by the use of EMCWF's ERA5 data set.

For more information, contact: David_Coe@student.uml.edu, @Daviology38, github.com/Daviology38