3-2020

Characterizing Regime-Based Flow Uncertainty

John L. Fioretti

CHARACTERIZING REGIME-BASED FLOW UNCERTAINTY

THESIS

John L. Fioretti, Captain, USAF

AFIT-ENP-MS-20-M-093

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT-ENY-MS-15-M-093

CHARACTERIZING REGIME-BASED FLOW UNCERTAINTY

THESIS

Presented to the Faculty

Department of Engineering Physics

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Atmospheric Science

John L. Fioretti, B.S.

Captain, USAF

March 2020

AFIT-ENY-MS-15-M-093

CHARACTERIZING REGIME-BASED FLOW UNCERTAINTY

THESIS

John L. Fioretti, B.S.

Captain, USAF

Committee Membership:

Lt Col Robert C. Tournay, PhD
Chair

Lt Col Andrew J. Geyer, PhD
Member

Dr. Astrid Suarez-Mullins
Member

Abstract

The goal of this work is to develop a regime-based quantification of horizontal wind field uncertainty utilizing a global ensemble numerical weather prediction model. In this case, the Global Ensemble Forecast System Reforecast (GEFSR) data is utilized. The machine learning algorithm that is employed is the mini-batch K-means clustering algorithm. 850 hPa Horizontal flow fields are clustered and the forecast uncertainty in these flow fields is calculated for different forecast times for regions across the globe. This provides end-users quantified flow-based forecast uncertainty.

Acknowledgments

I would like to express my sincere appreciation to my faculty advisor, Lt Col Robert Tournay, for his guidance and support throughout the course of this thesis effort. The insight and experience was certainly appreciated, and for which this work would not have been possible without. I would, also, like to thank my sponsor, Dr. Astrid Suarez-Mullins, from the Air Force Technical Applications Center for both the support and latitude provided to me in this effort. Lastly, I would like to thank my family, friends, and classmates for their unwavering emotional and technical support in this endeavor.

John L. Fioretti

Table of Contents

Page

List of Figures

List of Abbreviations

CLARA Clustering Large Applications

CLM Cluster Likelihood Match

ECA&D European Climate Assessment and Dataset

ECWFM European Centre for Medium-Range Weather Forecasts Model

ESRL/PSD NOAA Earth System Research Laboratory's Physical Sciences Division

FH Forecast Hour

GEFS Global Ensemble Forecast System

GEFSR Global Ensemble Forecast System Reforecast

NCAR National Center for Atmospheric Research

NCEP National Centers for Environmental Prediction

NOAA National Oceanographic and Atmospheric Administration

NWP Numerical Weather Prediction

STE Source Term Esitmation

T-PCA T-mode Principal Component Analysis

CHARACTERIZING REGIME-BASED FLOW UNCERTAINTY

I. Introduction

Most people view weather as the current state of their observable atmosphere, e.g. sunny and warm. When observed from just the basic sense not much can be learned with respect to *why* it is sunny and warm. Atmospheric scientists learned very early on that if they were to understand the "why" behind the weather, they would have to analyze phenomena on a much larger temporal and spatial scale. Scientists like Baur and Rex studied such large-scale patterns in the 1950's in order to better predict weather (Franzke et al.2008).

The advent of computers began a new era of weather prediction through numerical weather prediction (NWP) models. NWP models enabled forecasting at a level not possible through hand-analysis due to the shear amount of processing power needed to make such forecasts. However, NWP models are not flawless. They are algorithms run by the governing laws forecasters provide. One issue then becomes that forecasters have a choice of what assumptions to follow, i.e. if the algorithm will assume hydrostatic balance or not. A second issue is that the governing equations used to develop the algorithms are *assumptions* which means that there will be inherent error with every forecast.

Ensemble forecasting uses multiple NWP integrations all of which are perturbed slightly from the control case, either in initial conditions or model configuration. This system allows forecasters to characterize uncertainty of a forecast. The drawback to

ensemble forecasting is that the computational expense combined with more demand for lower grid-spacing and more accurate forecasts can become infeasible.

It is important to note that ensemble forecasts, albeit can provide an attached uncertainty, are still made of the same assumptions and governed by the same laws as deterministic NWP. Machine learning and artificial intelligence become the next logical step. Similar to the approach taken by Baur and Rex in the 1950's, machines can identify regimes on a global and decadal level in an analysis period orders of magnitude less than manual labor.

Additionally, the computational cost to run clustering algorithms is relatively low. There are two main types of machine learning, supervised and unsupervised. This research focuses on unsupervised machine learning. A mathematical tool employed in unsupervised machine learning processes is called a clustering algorithm. Many clustering algorithms have been tested with differing mathematical procedures but all have the same end goal of delineating patterns in data. The method employed in this research is the K-means clustering algorithm which is explained in full in the background and methodology sections.

The objective of this research is to utilize unsupervised machine learning to identify recurrent atmospheric regimes by clustering the 850 hPa horizontal flow fields for different forecast times across the globe. Then forecast uncertainty values are applied to the varying atmospheric regime clusters which will develop an application to inform decision makers on the confidence of a forecast given the current atmospheric regime. This chapter introduces the motivation and objective of this research. Chapter II provides a brief review on synoptic scale weather, early clustering techniques, K-means clustering,

NWP, reforecast data, and prior research. Chapter III details the methodology including data selection, data normalization, cluster analysis procedures, and uncertainty calculations. Chapter IV presents the results of multiple clustering trials and analyzes interesting patterns that emerged in the data. There is also a case study conducted in this section. Lastly, chapter V discusses potential impacts of this research, limitations, and future work ideas.

II. Background

2.1 Synoptic Scale Weather

Synoptic-scale weather has a spatial scale of 1000-6000 km and occurs on a temporal scale of 1 day to 1 week (Lackman 2015). There are also other defining scales such as velocities in both vertical and horizontal, planetary boundary layer depth, and horizontal and vertical pressure change. Typical synoptic-scale events include extratropical high and low pressure systems (Figure 1), jet streams and associated waves, and fronts (Lackman 2015). Characterizing atmospheric phenomena through synoptic dimensions allows for large-scale predictor patterns to be discovered. Synoptic-scale characterization of the atmospheric state is useful as a planning tool, alongside smaller-scale, smaller grid-spacing products. For example, trained meteorologists can ascertain that if a maturing low-pressure system is properly aligned with upper-level forcings for ascent along with moisture feed from the Gulf of Mexico at 850 hPa and proper vertical shear, then severe weather is possible. Similarly, an omega blocking pattern is indicative

of cooler temperatures over the western and eastern United States, and warmer temperatures over the central portions of the country.

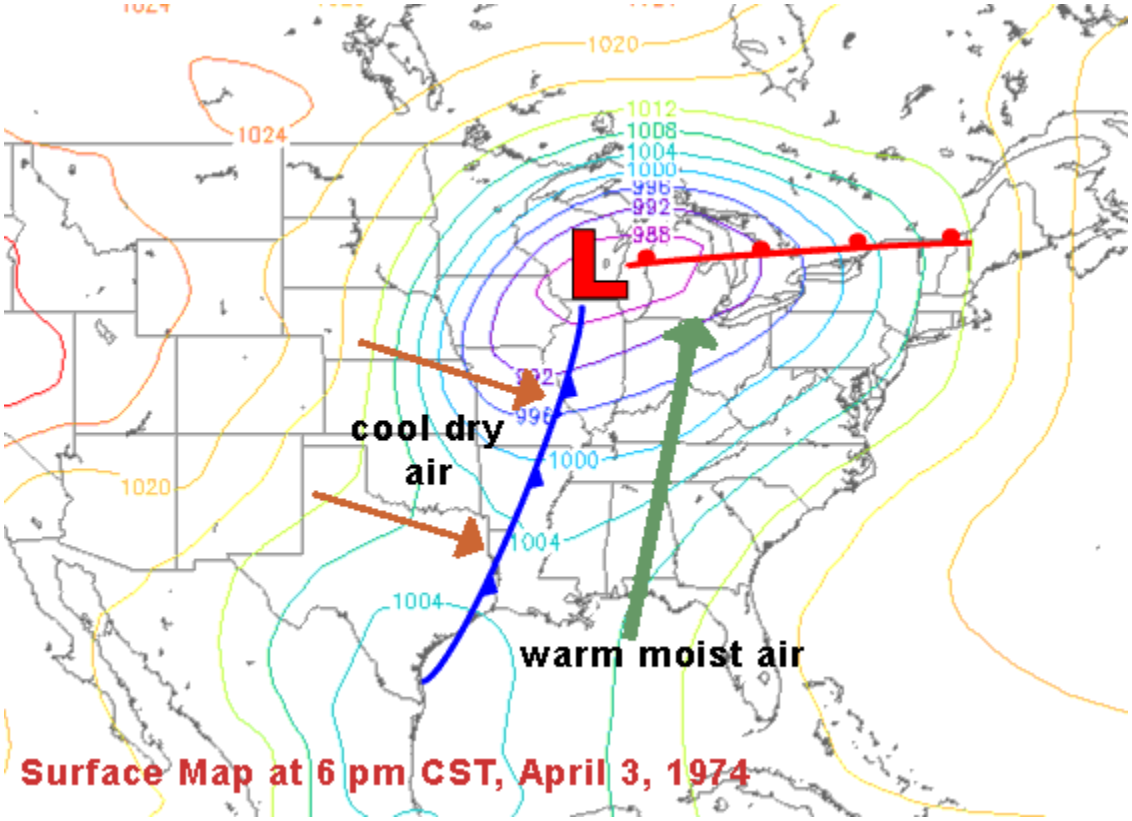


Figure 1. Synoptic-scale weather example. Image from the National Weather Service.

2.2 Atmospheric Regimes

An atmospheric regime is a recurrent synoptic scale spatial atmospheric structure (Vrac and Yiou 2010). Adenstedt (1970) states that atmospheric flow patterns can be interpreted differently based on the timescale in which they are viewed. For example, although every day atmospheric fluctuations can seem like random noise, when viewed from a yearly timescale, seasonal variations are very apparent. Feldstein (2000) states that the temporal scale for atmospheric regimes is typically 8-10 days. Mid-20th century synopticians, such as Baur (1951) and Rex (1950), noticed a recurring atmospheric flow

called an omega blocking pattern (Franzke et al.2008). The recognition of synoptic scale recurring patterns is important because it can lead to long-range predictions which is why, according to Franzke et al.(2008), there is a continuing search for such regimes.

Initially, regimes were spotted by eye. However, that is labor-intensive and does not yield quantifiable results. Lorenz (1963) agrees with this sentiment by ascertaining that hydrodynamical systems (the atmosphere) don't always follow the same patterns, even when observed for a long time. This leaves forecasters to "willy nilly" predict cyclones and anticyclones.

An attempt to classify weather regimes in stochastic meteorological models is made by R.K. Adenstedt in 1970. Adenstedt is the creator of the Markov regime process model, which through hand-calculated algorithms, determines atmospheric regimes. Adenstedt's study ends only with a series of new questions which lead to how one can better estimate large scale patterns through statistical functions. However, this does not alleviate the workload or improve the accuracy to high levels.

2.3 Clustering

According to Vrac and Yiou (2010), an objective approach to classifying regimes is possible through the use of mathematical clustering techniques. Cluster analysis is a statistical classification technique for discovering whether the individuals of a population fall into different groups by making quantitative comparison of multiple characteristics. In atmospheric science, each cluster that is produced has similar atmospheric characteristics within itself but different characteristics to other clusters. Clustering is a member of the machine learning family. There are two main types of machine learning,

(i) supervised and (ii) unsupervised. The difference being that supervised involves utilizing data where patterns are known, labeled, and then used to train the algorithm; whereas unsupervised only utilizes unlabeled data (Jain 2010).

There are two types of clustering, (i) hierarchical and (ii) partitional. Hierarchical clustering recursively finds nested clusters through either (a) agglomerative or (b) divisive. In agglomerative mode, every data point becomes its own cluster and then similar pairs are merged successively forming a hierarchy. Divisive mode places all the data into one cluster and through recursive steps divides each cluster into smaller clusters. Partitional clustering algorithms find clusters simultaneously as opposed to hierarchical clustering (Jain 2010).

2.4 K-means

The main clustering approach employed in this research is K-means clustering. This clustering technique is based on the sum-of-squares (SSQ). The first example of SSQ being utilized is in the work of Dalenius and Gurney in 1950 and 1951 (Bock 2008). SSQ is the sum of the square of variation. Variation is the measure of similarity or dissimilarity between the spread of each value to the mean. Dalenius does not use the K-means algorithm to minimize the sum-of-squares and, instead, uses a different approach. However, in 1965, Forgy became the first to use the K-means algorithm for minimization. After this point, K-means became the standard when it comes to clustering techniques (Bock 2008).

K-means finds a partition that minimizes the squared error between the empirical mean of the cluster and the points in the cluster, and then does this for all clusters. K-

means is a greedy algorithm, which means that the choice made at every step is one that picks the solution with the most obvious and available benefit. K-means starts with an initial partition of clusters and reduces the squared error between clusters by assigning patterns. The number of clusters must be fixed because the squared error always decreases with an increase in the number of clusters (Jain 2010). The following main steps for the K-means algorithm are from Jain (2010): 1) Select an initial partition with clusters; repeat steps 2 and 3 until cluster membership stabilizes, 2) generate a new partition by assigning each pattern to its closest cluster center, and 3) compute new cluster centers. This process can also be visualized in Figure 2.

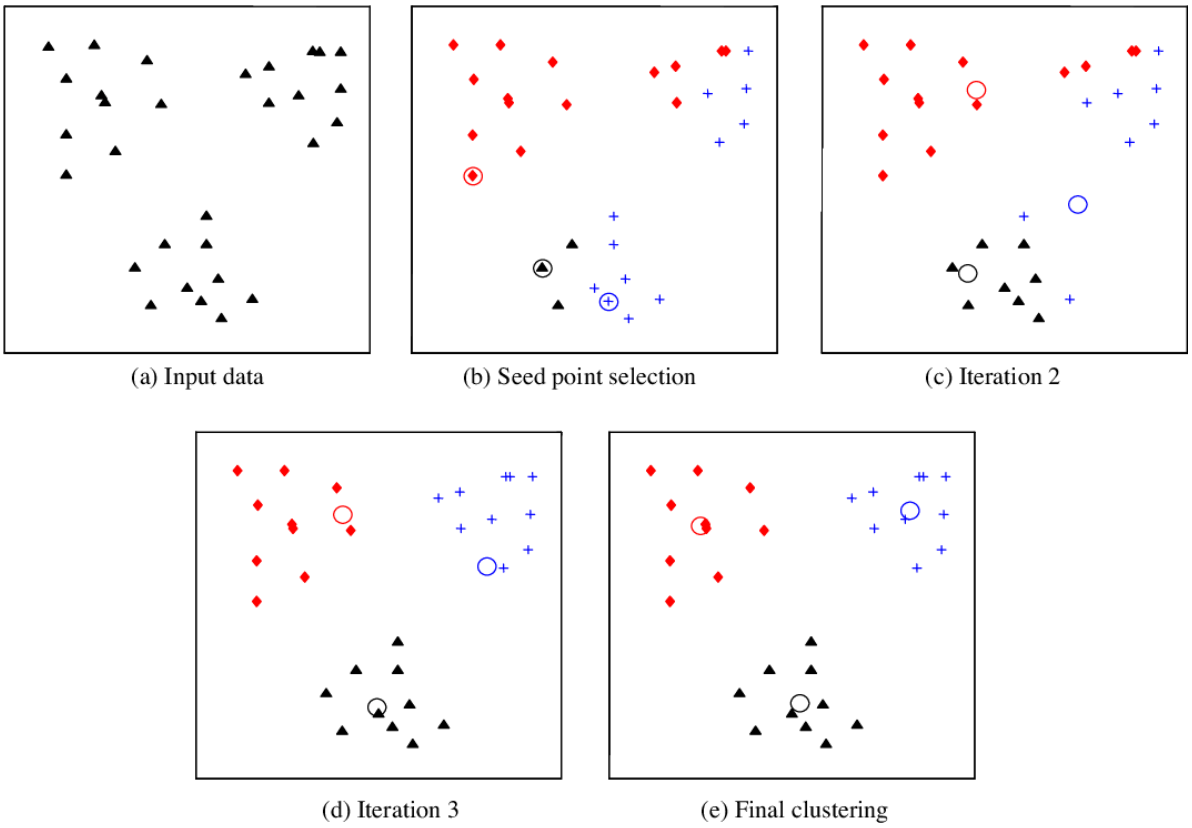


Figure 2. Visualization of the K-means clustering algorithm (Jain 2010).

One possible setback of clustering is that algorithms can find clusters in data that has no natural clustering (Figure 3). Therefore, it is important to look at cluster validity in order to evaluate clusters objectively. There are three different types of cluster validity indices: internal, relative, and external (Jain 2010). Internal criteria is a comparison of the structure of the clustering algorithm and the data solely using the data. Relative criteria is the comparison of multiple structures created by different algorithms, and then choosing the one that makes the most sense. Lastly, external indices are a comparison of the cluster structure to pre-existing class labels or ground truth.

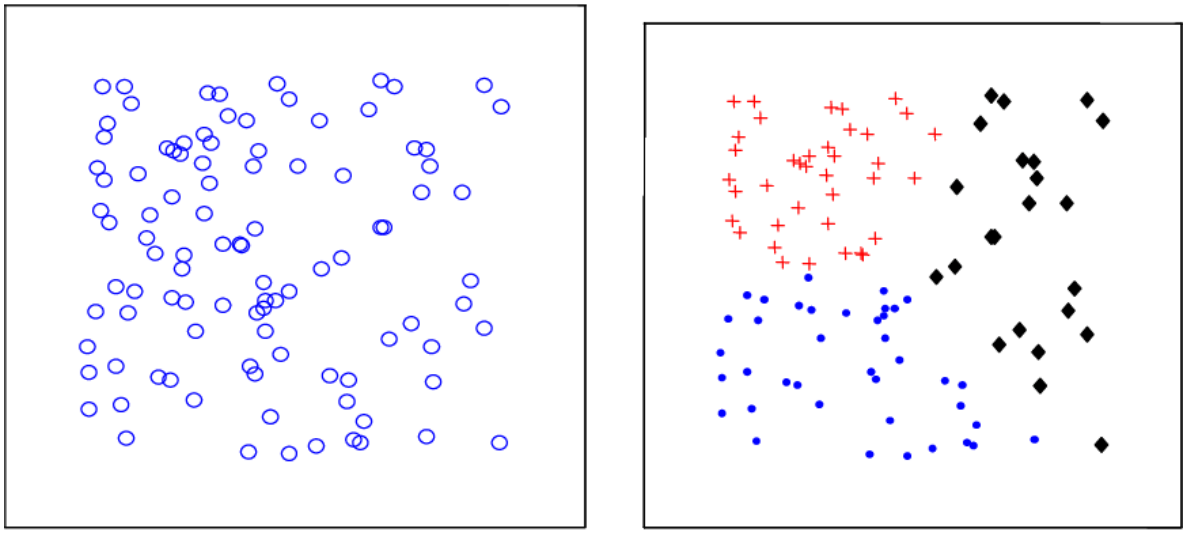


Figure 3. Example of a dataset with no natural clustering (a) and when broken into 3 clusters (b). This example illustrates that picking a cluster number that makes sense is important (Jain 2010).

2.5 Numerical Weather Prediction

NWP is the numerical implementation of the governing equations of fluid dynamics which are set to initial conditions and are fundamental to almost all dynamical weather prediction schemes. NWP models have made vast improvements since their

implementation in the mid-20th century. However, there are sources of possible error. Warner (2011) describes how initial conditions can be a source of error due to possibly a lack of or poor observations caused by improper sitting, calibration, or representativeness. Additionally, surface boundary-condition uncertainty can result from errors of the land-surface properties calculations. Also, numerical algorithms and parameterizations based off of assumptions of physical processes can be a source of error (Warner 2011). This means that the deterministic output can be incorrect due to inaccuracy of initial conditions, parameterizations or simplifying assumptions (Zarnani and Musilek 2013). There will always be some inherent uncertainty with forecasts. Thus a forecast can be more valuable if it is associated with a certainty.

The advent of the ensemble forecasts in the 1990's allowed forecasters to quantify some of the uncertainty. A forecast created by an ensemble is one in which the model is run multiple times with slight variations on the initial conditions (Scher and Messori 2018). According to Warner (2011) there are four reasons why an ensemble forecast is typically more accurate than forecasts of an individual ensemble member: (i) the mean of the ensemble is more accurate than the forecast from an individual member, (ii) the variance across the ensemble members can be a depiction of forecast uncertainty, (iii) the probability distribution function can give information about anomalous events, and (iv) quantitative probabilistic output can be used in decision-support software systems. Expounding on the second point above, the variance amongst ensemble members as it pertains to forecast uncertainty is fundamental to this thesis research. Two disadvantages of ensemble forecasting is that they are very computationally expensive to run, leading to

a several hour time-lag between initialization and delivery to the end-users (Scher and Messori 2018).

2.6 Re-analysis/Reforecast Data

When researching atmospheric regimes, a historical dataset of forecasts is necessary. Operational NWP models are regularly updated making it difficult to acquire long periods of homogeneous forecasts. Reforecasts are created by utilizing observations in the past and places them on a regularly spaced grid which aids in comparison to NWP models. These observations are then combined with an NWP model which develops a gridded set of model dependent variables which are consistent with the model and with the observations (Warner 2011). This ensures a long-term historical database of homogeneous forecasts (Scher and Messori 2018). Reforecast data is similar to reanalysis with the difference being that forecasts are created at regular intervals using the reanalysis data. Reforecast data is used in calculating biases in the model, predictability studies, clustering, and deep-neural networks (Warner 2011).

2.7 Similar/prior research

A study by Zarnani and Musilek (2013) evaluate various clustering algorithms to determine uncertainty, or the prediction interval (PI), of NWP forecasts. NWP does not automatically have corresponding uncertainties attached with their forecasts, so the authors' goal is to research and implement a possible solution. The confidence level is the expected probability that the observation will fall inside the PI range. One beneficial aspect to hypothesizing uncertainty within atmospheric science is that it is known that the

forecast uncertainty varies with the weather situation. For example, in situations where the atmospheric regime is more unstable, it is expected that there will be more uncertainty in those forecasts (Zarnani and Musilek 2013).

Zarnani and Musilek (2013) describe a PhD dissertation by Pinson in 2006 that classified the atmospheric state based on different variables manually, and empirically calculated the prediction intervals. They determine that although the results are found reliable, this approach is not efficient and will have some error due to the subjectivity of the manual clustering. Their statistical PI computation procedures are independent of the weather situation, time and location. This means that the resultant PI's are invariant on the atmospheric regime (Zarnani and Musilek 2013). However, because the goal of the research is to draw connections between certain atmospheric regime and associated forecast error, additional steps must be taken to draw conclusive comparisons. Atmospheric regimes are clustered using different machine learning methods, their errors are attached, and are compared. It is found, when compared to baseline methods, there is an average of 8% improvement in PI forecast skill when using clustering techniques (Zarnani and Musilek 2013).

Ferranti et al. (2014) studies the connection between weather regimes with ensemble spread and forecast skill. Similar to the study by Zarnani and Musilek, it is described in their research that certain types of atmospheric flow are stable and predictable, whereas others are unstable and unpredictable. Their model of choice is the European Centre for Medium-Range Weather Forecasts (ECWFM) model which uses 50 independent ensemble members. It is possible to tell predictability when analyzing ensemble spread of the members after the fact, but their goal is to find a way to know a-

priori. One of the challenges of finding flow-dependent predictability is that the flow configurations being studied must occur frequently enough so that the ensemble spread data can be gathered. Ferranti et al.,(2014) notes that the analysis of weather regime frequency provides a synthetic description of the atmospheric state. They use 4 clusters that are determined by the K-means algorithm of the 500 hPa daily anomalies from 1980-2008. These clusters can be seen in Figure 4. At the conclusion of their analysis they find that certain regimes had more associated error than others, see Figure 5. This research is important as it pinpoints atmospheric regimes which can be troublesome for NWP models and forecasts as it outlines how to assign prediction skill to a forecast.

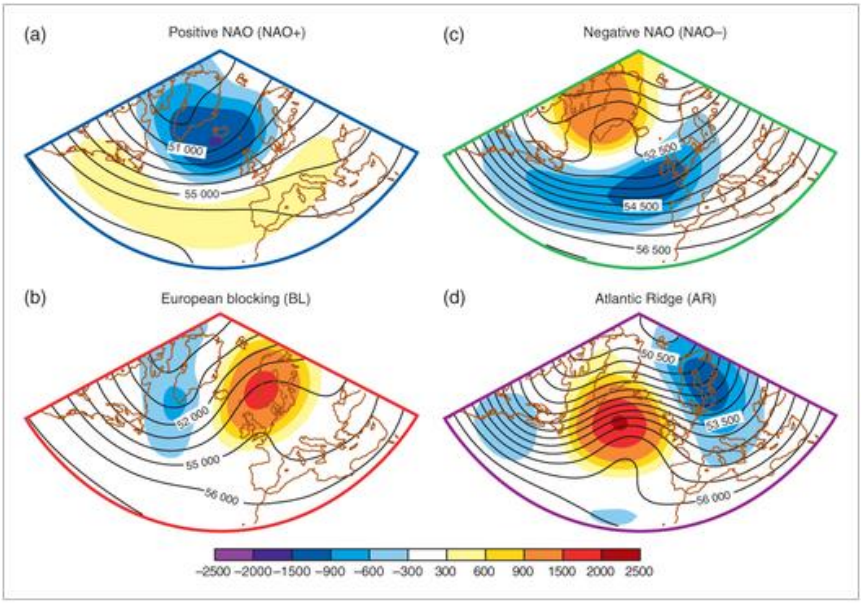


Figure 4. Depiction of the 4 clusters determined by the K-means analysis of Ferranti et al. (2014).

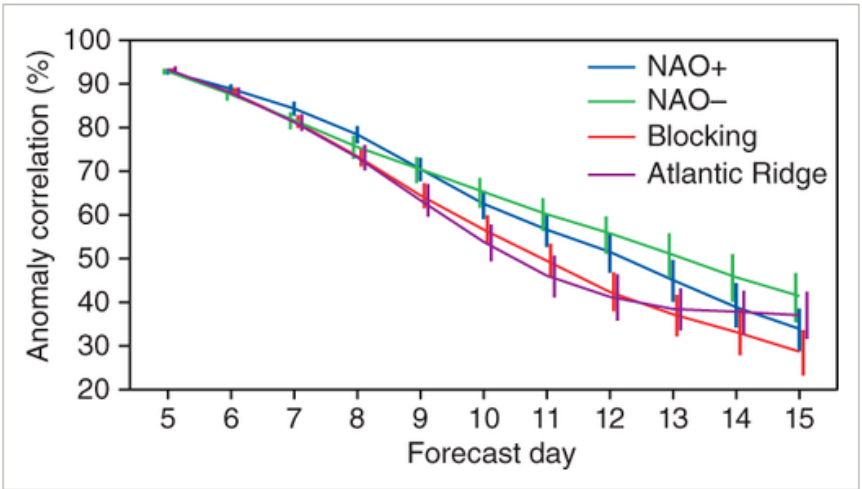


Figure 5. Associated errors with different regimes (Ferranti et al. 2014).

Vrac and Yiou (2010) study varying clustering techniques including K-means. The first type of subjective regime classification comes from the Lamb weather types in 1972. Due to its subjectivity, there is an inherent error involved with each classification. Thus, objective mathematical analysis using machine learning is the preferred method. The focus of their study is the correlation between rainfall and atmospheric regimes on the French Southern Mediterranean region. The precipitation data is derived from the European Climate Assessment and Dataset (ECA&D) from 1959 through 2004 for seven different observation stations. Their reanalysis data is from the National Centers for Environmental Prediction-National Center for Atmospheric Research (NCEP/NCAR). Five different clustering techniques are utilized, but the focus for this thesis will remain on K-means. The correlation between atmospheric regimes and precipitation is determined through canonical correlation analysis, which for the purposes of this thesis is mentioned only for context. The results (Figure 6) depict a spread between ≈5 mm to

≈17 of precipitation. Clusters 5 and 7 are clearly separated from the rest of the clusters. The remaining clusters are more similar to each other. This information is useful because, if validated, one can determine that the current atmospheric regime resembles cluster 5 and can say the percentage for precipitation is greater than zero.

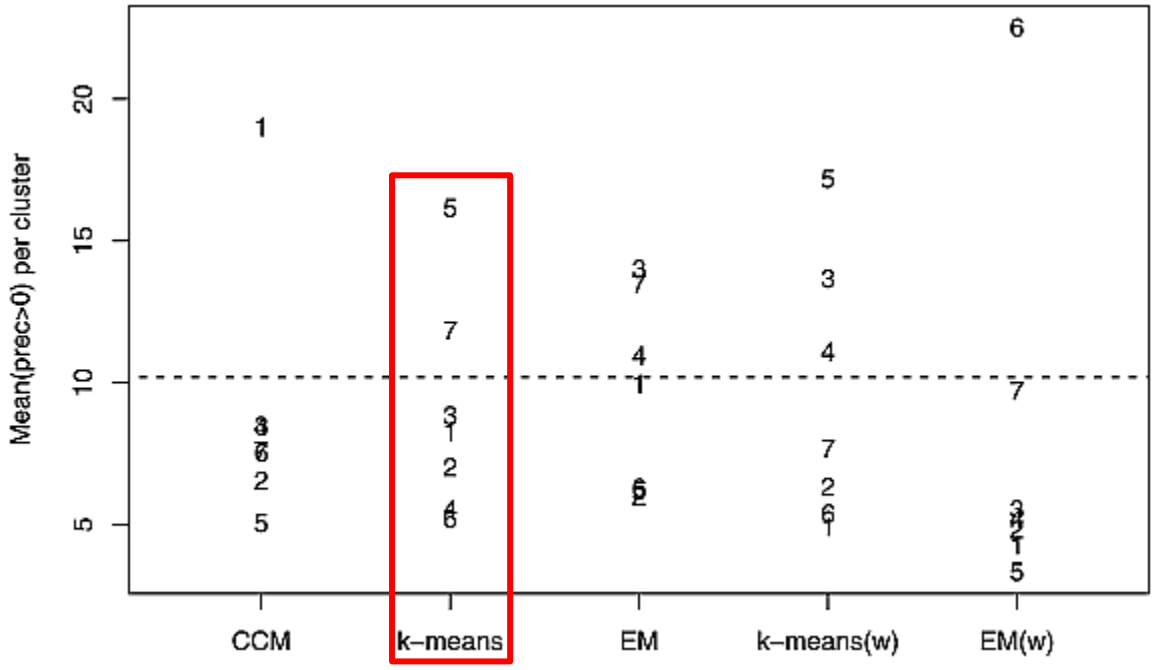


Figure 6. Tabular representation of possible correlation between rainfall amounts and cluster type (Vrac and Yiou 2010).

An additional way to view the results is depicted in Figure 7. This is the spatial distribution of daily rainfall occurrences as anomalies relative to the overall mean for each station. The conclusions of their research find that the clustering techniques used in their study show correlations between precipitation and atmospheric regimes. However, other atmospheric variables aside from the two chosen for the reanalysis can be applied

and do not have to be compared to precipitation but to other variables. Even with the additional work that will have to be done, they state that their framework can be utilized for different portions of the world.

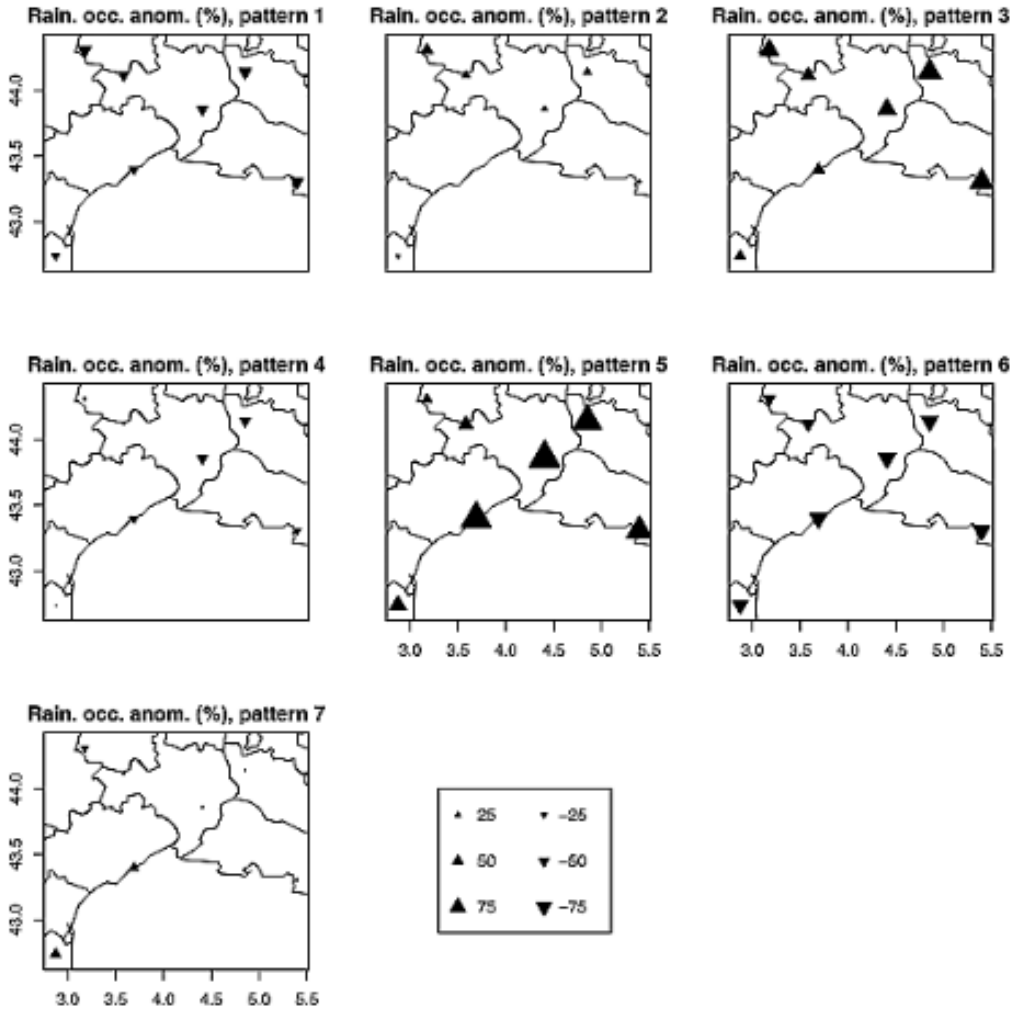


Figure 7. Visualization of possible correlation between rainfall amounts and cluster type (Vrac and Yiou 2010).

Huth (2000) conducts a study on disaggregating climatic trends by classification of circulation patterns. The motivation for this study stems from the search for signs of

global climate change. Those particular studies focused on mostly cloudiness, surface temperature and sunshine duration. However, Huth studies trends in different air-mass types. Instead of focusing on atmospheric variables, the variables are stratified based on circulation type. The study covers a period from 1949 to 1980. Nine climate elements are studied including maximum and minimum temperature, temperature range, relative humidity, cloudiness, zonal and meridional wind, sunshine duration, and precipitation occurrence. 500 hPa geopotential height fields are analyzed to identify circulation types. T-mode principal component analysis (t-PCA) is used alongside K-means analysis. T-PCA is useful at finding the underlying structure of the data but struggles to discern structures within large datasets. K-means is useful due to its ability at creating well-separated clusters. When both are combined, according to Huth, the negatives of both approaches cancel out. The K-means handles the large datasets and creates well-defined clusters on which t-PCA relies. The data is broken up in 5 subsets of 270 days which are randomly selected from the whole of the data. Within the subsets 7 circulation patterns are identified from the t-PCA. The mean of the patterns is taken and then used as initial centroids for the K-means procedure. After this process, 5 classifications of the total dataset are produced with each day classified by its membership in five types. Lastly, the days with similar types are grouped together. More procedure follows in Huth (2000), however the procedure for discerning climate trends is beyond the focus of this thesis research. From this methodology, Huth comes to 7 new findings due to the light shed by K-means and t-PCA on the mutual relationships between the trends of climate elements and causes.

Michelangeli et al.(1994) seek to find differences in clustering of recurrent weather regimes. Additionally, they develop a classifiability index to determine the number of clusters that should be chosen for their study. They define weather regimes as the states of the atmosphere with the highest probability of occurrence. This can be determined through cluster analysis. The data consists of 700 hpa northern hemisphere geopotential height maps from the National Oceanic and Atmospheric Administration (NOAA) Climate Analysis Center. Every other day from every winter from 1949 to 1992 is analyzed. Additionally, in order to quantify impacts of weather regimes near the surface, 850 hpa temperature maps are utilized.

For the cluster analysis, the algorithm used is the dynamic cluster method from Diday and Simon (1976). The algorithm finds a partition of the data that is put into clusters in order to minimize the sum of variances within the clusters. However, the authors pose three questions; (i) Given the number of clusters, what is the best partition? (ii) What is the correct amount of clusters to choose? (iii) Are the results reproducible? These questions are asked because of the inherent mechanics of the clustering method. The algorithm will always converge to a final partition. If the distribution of the data given to the algorithm is uniform, then the final partition should depend on the initial centroids. If the distribution of the data is not uniform, then the results should rely less on the initial centroids and should be more reproducible. Their way of alleviating this issue is by creating a classifiability index which tests the dependence of the results on the initial centroids. This is achieved by running the dynamical clusters algorithm 50 times with different initial seeds, and then testing, through averages, the classifiability of the clusters. The results of their tests lead to the Atlantic Sector having k=4 and the Pacific

Sector with k=3. The results for the Atlantic and Pacific Sectors can be seen in Figure 8. Furthermore, in order to verify the reproducibility, the data for each sector is split in half and the respective halves are compared. In all cases the comparisons are similar. The authors' state that their preferred approach is to classify weather regimes with cluster analysis as recurrent circulation patterns. It is mentioned that more cluster amounts need to be evaluated in order to find the most relevant number. Also, because partitioning methods are initialized with random centroids, two calculations with the same sample may not end with the same partition. However, they overcome this problem with their classifiability index.

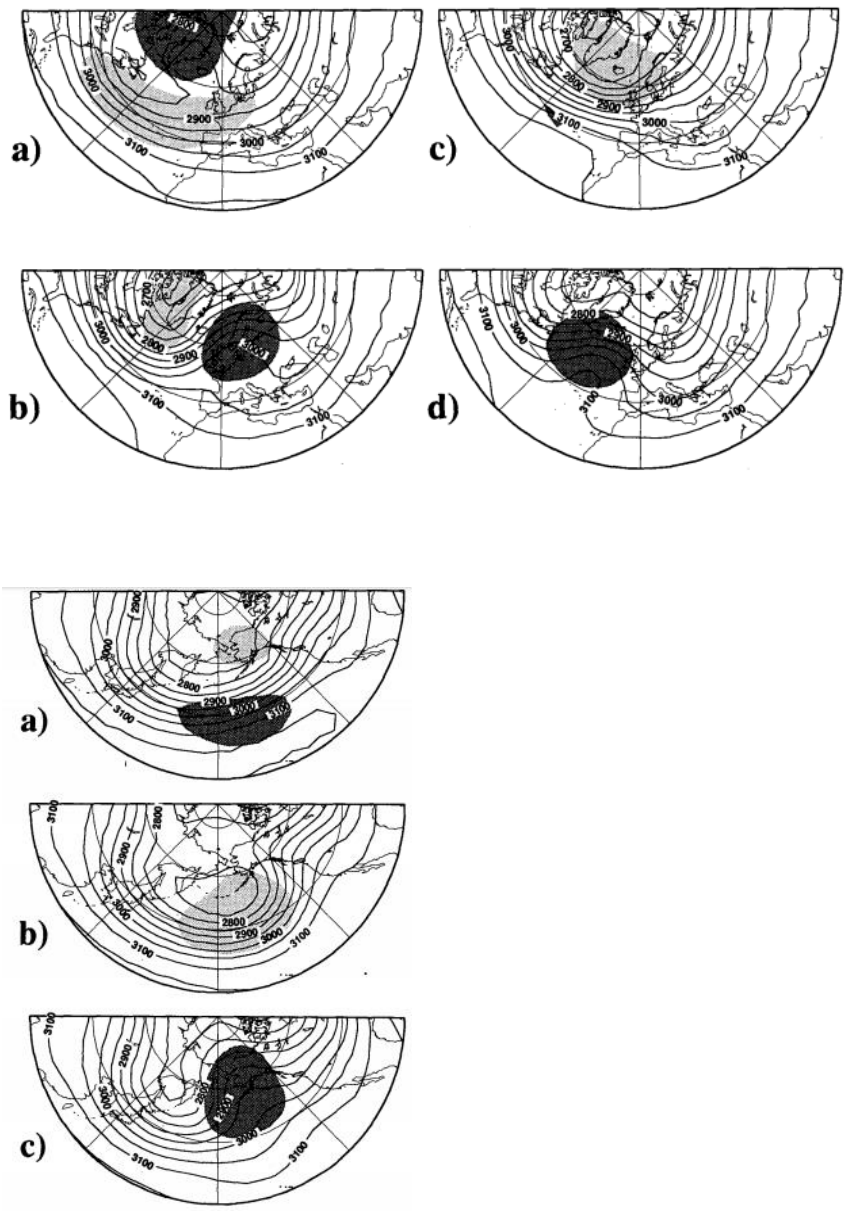


Figure 8. Graphical representation of the clusters obtained by Diday and Simon (1976).

Scher and Messori (2018) focus on predicting weather forecast uncertainty with machine learning. While their approach utilizes deep learning, as opposed to clustering, their methodology and research philosophy with respect to forecast uncertainty are

particularly useful to this thesis research. The authors state that forecasts will always have some degree of uncertainty associated with them due to issues highlighted in the above NWP section. They reiterate that weather forecasts are valuable only if the uncertainty is known. Their work is an attempt to create a measure that determines the uncertainty so that it may be associated with a forecast. They base their method only on prior forecasts independent of a NWP model. Their hypothesis is that forecast uncertainty is dependent on the atmospheric state from which the NWP model is initialized along with a random component. Thus, through analyzing past forecast spread and error, they can discern which atmospheric situations lead to higher or lower forecast certainty.

The data used in their study is from the Global Ensemble Forecasting System (GEFS) reforecast dataset. They use geopotential height at 500 hPa and zonal and meridional wind at 300 hPa for the period from 1985 to 2016. Scher and Messori (2018) compute the ensemble standard deviation of the 500 hPa geopotential height for 3-6 days after initialization, and use this as the GEFS spread. Their justification for using 500 hPa geopotential height is that it is a good proxy for large- scale weather. Also, the root mean square error of the 500 hPa geopotential height is computed for the same forecast days and is used as the forecast error.

Scher and Messori (2018) find that the predicted uncertainty correlated to ensemble spread is ≈ 0.33 and compared to error is ≈ 0.27 for day 3. When the lead time is increased to day 6, the correlation to spread and error is ≈ 0.28 and ≈ 0.19 respectively. This shows that forecast spread is low when predicted uncertainty is low, and high when the predicted uncertainty is high. In order to evaluate the effectiveness of their approach, they correlate their uncertainty predictions with the ensemble spread of

the forecast. This aids in assessing how their uncertainty predictions differentiate days with high and low forecast error. Their predictions are binned according to their error and then the mean error for each bin is calculated and plotted. This can be visualized in Figure 9. The binned forecast error increases with predicted uncertainty, which is an indicator that their method is successfully predicting uncertainty.

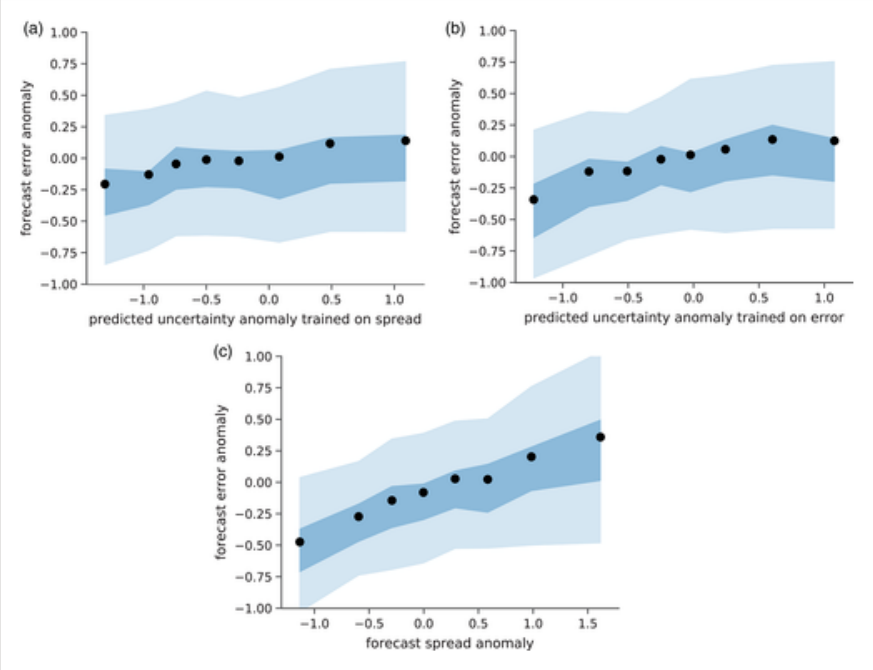


Figure 9.Relationship between predicted uncertainty and actual forecast error for (a) neural network trained on spread, and neural network trained on error (b). Forecast spread versus forecast error is depicted in (c). The data were binned along the *x*-axis in bins containing 300 samples. The points show the mean forecast error of one bin. The shading shows the 20–80 and 40–60 percentile range of forecast errors for each bin. All panels refer to forecast day 3 (Scher and Messori 2018).

III. Data and Methods

This research utilizes the Global Ensemble Forecast System Reforecast (GEFSR), version 2 (Hamill et al 2013) available through ESRL/PSD. GEFSR is an 11 member ensemble initiated at 00Z each day from 1 December 1984 until present. The horizontal grid-spacing is 1x1 degree latitude/longitude through 192 forecast hours. For this study, the analysis utilized is 24, 72, and 120-hour forecasts. Larger grid-spacing output is available from 198 through 384 hours. Output for every day from 1 January 2002 to 31 December 2016 is utilized totaling 5479 days. The data is standardized at each point in space and time utilizing the z-score method, utilizing the mean and standard deviation for five points in each direction (x,y, and time). This ensures that there seasonal cycles are not an influence in the data.

To test the method, only meridional wind at 850 hPa (V850) is utilized. Only GEFSR data from the mid-latitudes is utilized, here described as 20 to 70 latitude (in both the southern and northern hemisphere). This is because the more reproducible synoptic-patterns occur in the mid-latitudes, i.e. mid-latitude cyclone. For each point from 29 to 51 degrees latitude, a 19 by 19 degree latitude/longitude box was created, creating 5479 standardized images of V850 at each point. An independent minibatch K-means cluster analysis is conducted for each set of 5479 images at each point using the scikit learn library for the Python programming language. Through trial and error, it is determined that the optimal number of clusters necessary to represent the passage of a synoptic-scale wave is 6. When more clusters are chosen, the difference between said clusters diminishes, thus reducing their purpose. When less clusters are chosen, important phases in wave passage are lost.

Once the clusters are determined for each of the 5479 days in the data set, the clusters at each point are reordered to position the strongest negative anomaly (from the north) in cluster 0 transitioning to the strongest positive anomaly (from the south) in cluster 3 with the transition clusters in between, showing the cycle of a mid-latitude baroclinic wave. This is shown in the clusters for the southeast United States in Figure 10. The cluster analysis is conducted with the first ensemble member of the analysis data (zero hour forecast). Next, the cluster is determined for all three forecast hours (24, 72, 120) for all eleven ensemble members. The total number for each of the six clusters from the eleven forecast members is determined. The cluster with the greatest representation in the eleven ensemble member is chosen as the forecast cluster. For example, if 9 of the 11 ensemble members classified the regime as a 4, then 4 would be the overall forecast cluster choice.

With the cluster from the analysis and the forecast cluster, a determination if the correct cluster is predicted by the ensemble forecast system is can be made. A determination can then be made into the ability of the GEFSR system to correctly forecast these defined patterns over all days, in specific seasons and the variability of the system to predict specific patterns over the globe in the mid-latitudes. This is referred to henceforth as cluster likelihood of matching the forecast (CLM). CLM will show a likelihood of correctly predicting a certain cluster a certain number of hours out.

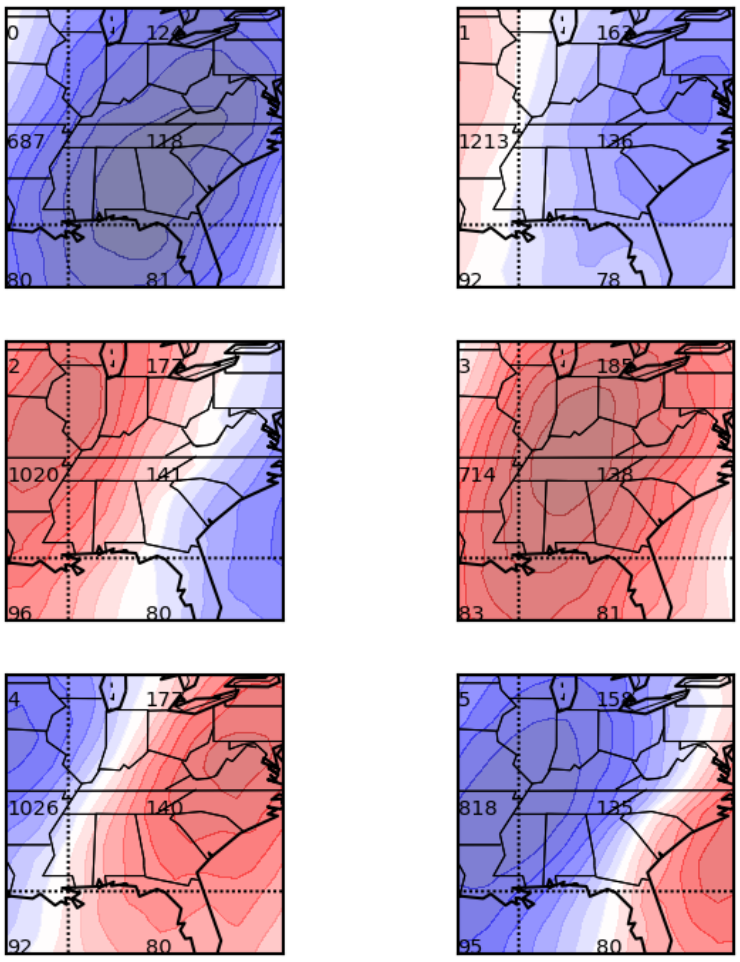


Figure 10. Idealized clusters for v-component winds. Cluster 0 is negative, cluster 3 is positive, clusters 1 and 4 represent frontal passage, and clusters 2 and 5 represent post/pre-frontal scenarios.

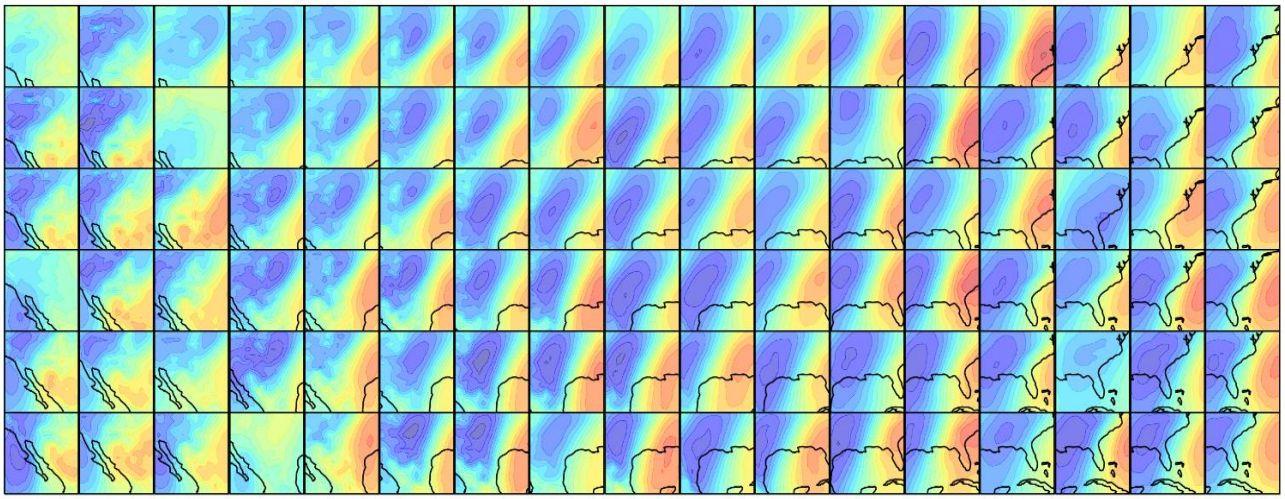


Figure 11. Visual example of the clustering process for each 19x19 grid box over CONUS.

IV. Analysis

The following is an analysis of the clusters and CLM of the forecast cluster at a given forecast hour. All examples are the V850 wind and are broken into the clusters for all 15 years of data. Additionally, clusters are analyzed for seasons which are 90 day time periods starting on Julian day 0, 90, 180, and 270. This is in order to determine if there are seasonal impacts on CLM. Forecasts are shown at +24 and +120 forecast hour (FH). Following the individual cluster analysis, there is a comparison of results as well as an attempt to explain certain findings.

It is important to mention that 850 hPa is under ground for some areas of the world, and that the results may be dubious in and around these locations. Additionally, when referring to directions positive is from equator and negative is from the poles. Furthermore, the northern hemisphere is the choice for when referring to seasons, e.g.

winter begins on Julian day zero for both hemispheres. Lastly, cold air advection and warm air advection are implied as congruently occurring with positive and negative V850.

4.1 V850 All Times Analysis

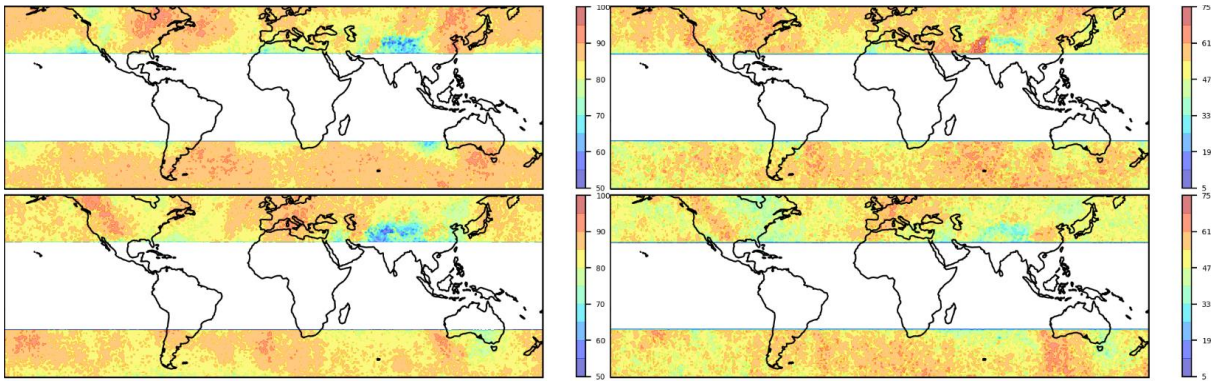


Figure 12. Probability of matching the forecast at 24 hours across all times for clusters 0 (top left) and 3 (bottom left) and 120 hrs for clusters 0 (top right) and 3 (bottom right).

V850 All Times Analysis of Cluster 0 at + 24 and + 120 FH

In figure 12, there is a high cluster likelihood of match in cluster 0 of 85% or higher for most areas in both the southern and northern hemispheres. This means that for a given forecast hour, 85% of the time cluster 0 occurred, it represented the greatest number of occurrences by the ensemble members. Areas of highest CLM exist over the Midwest and eastern United States including the coastal waters of the Atlantic Ocean. Other areas of high CLM are in eastern Russia and China as well as Iraq and far western Iran. Areas of lowest CLM occur west of the Rocky Mountains and North Africa. For the southern hemisphere, there is high CLM throughout with the highest areas of CLM being

off the southeast coast of South America as well as Australia. The lowest CLM lies off the southwest coasts of Australia, Africa, and South America. For +120 FH, the areas of CLM retain the same placement of high and low CLM but reduce as a whole in CLM.

V850 All Times Analysis of Cluster 3 at + 24 and + 120 FH

CLM for cluster 3 is overall less than in case 0 (Figure 12). Additionally, the areas of high versus low predictabilities are a mirror image in cluster 3 when compared to cluster 0. For example, where west of the Rocky Mountains had low CLM in cluster 0 now has high CLM. +120 FH still shows a decreasing trend overall, but still has the same placement of high and low areas of CLM as in +24 FH.

Comparison of Clusters 0 and 3 for All Times

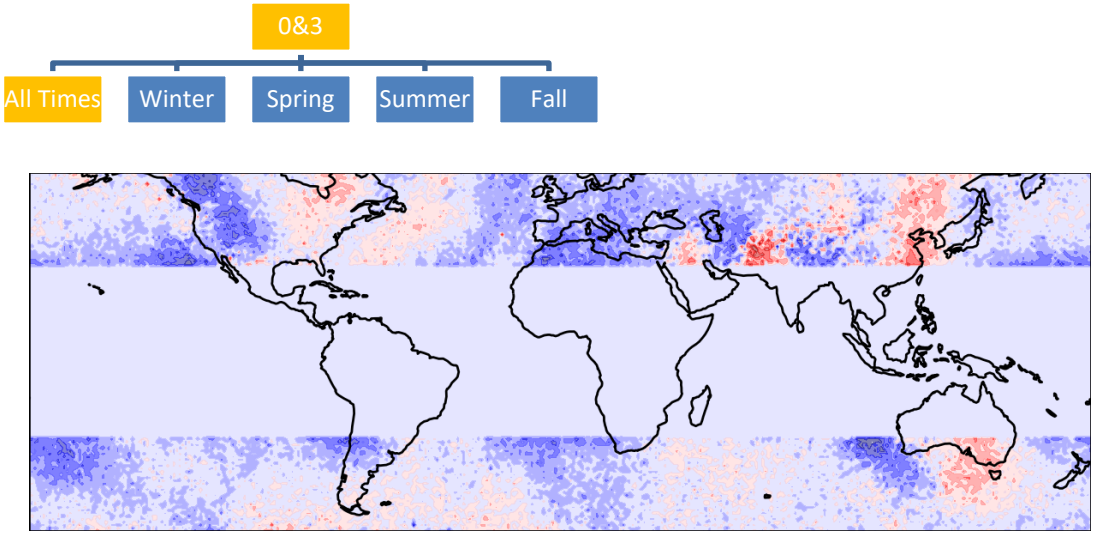


Figure 13. Difference between values from clusters 0 and 3 indicating areas of similar CLM (lighter colors).

In Figure 13, the areas of similar CLM for the northern hemisphere are located in the Pacific, West Atlantic, and central and eastern Asia. The most similarity between both cases lies in the Pacific. The areas of greatest difference lie along the United States west

coast, North Africa, and Iran. For the southern hemisphere, the areas of similarity are in the South Pacific, and off the southwest coast of Africa. The areas of greatest dissimilarity are off the western and southeast coast of Australia.

Both clusters exhibit high CLM. Cluster 0 has slightly higher values overall. A possible reason for why negative V850 winds have a higher CLM overall than positive is that positive winds indicate warmer equatorward air being driven north, and thus, an influx of instability and less predictable weather patterns.

Interestingly, the areas of maximum and minimum CLM mirror each other. In the United States, for example, the Rocky Mountains act like a natural division in the CLM. A justification for this reasoning can be found in the images. In areas where there is less terrain, the predictabilities do not change as drastically. For example, the southern hemisphere exhibits this characteristic. Another explanation for the mirrored predictabilities can be fundamentally shifted longwave patterns. A clue that this may be the case is that in cluster 0 there is a pattern of positive, negative, positive, and so on. One can imagine this being shifted to the right by ½ wavelength to get the negative, positive, negative pattern in cluster 3. This wavelike pattern is also apparent in the southern hemisphere.

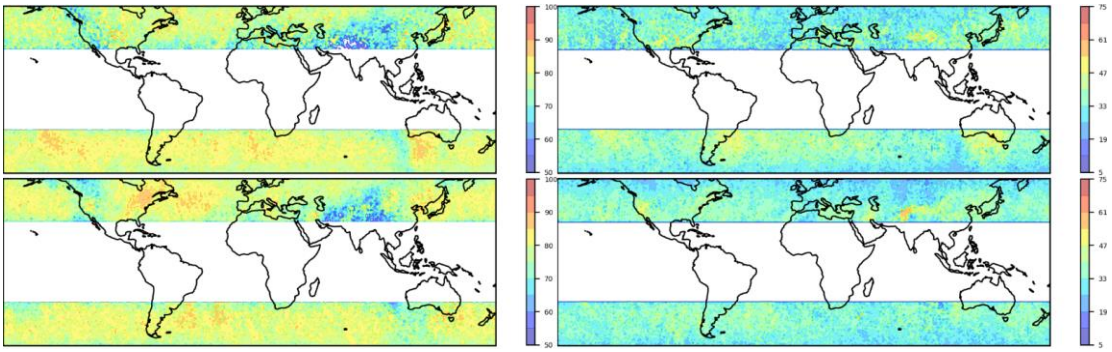


Figure 14. Probability of matching the forecast at 24 hours across all times for clusters 2 (top left) and 5 (bottom left) and 120 hrs for clusters 2 (top right) and 5 (bottom right).

V850 All Times Analysis of Cluster 2 at + 24 and + 120 FH

The areas of highest CLM lie over the Pacific, central United States, the Atlantic Europe and western Russia (Figure 14). The lowest CLM occurs over the west coast of the United States, and southwest/central Asia. The southern hemisphere CLM is higher than the northern hemisphere as well as more homogeneous. There are several areas of maxima in the southern hemisphere which include the South Pacific, South Atlantic, and off the southwest coast of Australia. There is a decrease in probability of matching the forecast at 120 hours however the areas of maximum stay the same.

V850 All Times Analysis of Cluster 5 at + 24 and + 120 FH

Overall, cluster 5 exhibits more areas with higher CLM than cluster 2 (Figure 14). In the northern hemisphere, the maximum occurs over the eastern half of the United States. Other areas of higher CLM include the Atlantic Ocean as well as eastern Asia/Russia and the Pacific Ocean. The minima occur over the western half of the United States as well as Europe, eastern Europe, and southwest Asia. For the southern hemisphere, the CLM is homogenous as usual but the maxima do not peak as high as in the northern hemisphere.

The maximum areas of CLM in the southern hemisphere include off the southeast coast of South America as well as the southeast coast of Australia. The minimum areas of CLM include the South Pacific and off the southwest coast of Australia.

Comparison of Clusters 2 and 5 for All Times

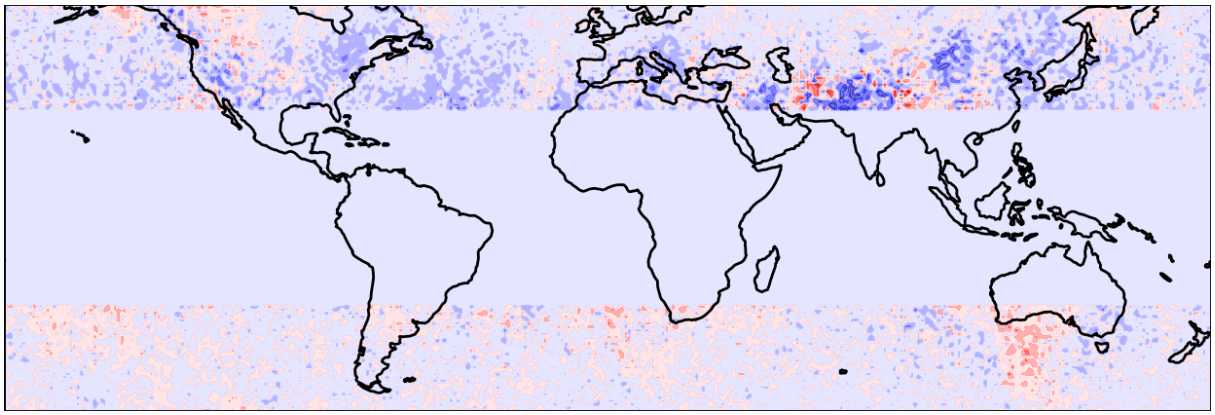


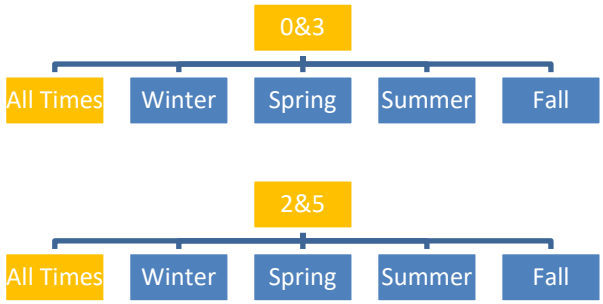
Figure 15. Difference between values from clusters 2 and 5 indicating areas of similar CLM (lighter colors).

In Figure 15, the areas with the most similarity in the northern hemisphere are the more northern sections with spotty similarity elsewhere. The United States and Asia have the greatest amounts of dissimilarity. Much of the southern hemisphere is similar, with the most dissimilarity occurring near Australia.

For the northern hemisphere, cluster 2 is much less predictable than cluster 5. The biggest increases of CLM between clusters 2 and 5 are in western Russia, the eastern United States and both the Pacific and Atlantic Oceans. The justification for this is that cluster 2 is positive anomalies to the west and negative anomalies to the east which can

be interpreted as a frontal like system moving into each 19x19 grid box. Whereas, cluster 5 can be interpreted as the development of high pressure behind an exiting front. In general, the timing of frontal development and entrance is a more difficult task than frontal exodus. When viewing cluster 5 one can notice that the eastern half of the United States forecasts frontal exit rather well. This concept makes sense due to the number of in-situ and radiosonde observations that it passed through prior to reaching the east coast. One can also see the continuance of higher CLM into the Atlantic Ocean, again because of the amount of real-time data gained on the particular frontal system. Interestingly, for cluster 2, the central portion of the United States has a higher CLM for frontal entrance into the 19x19 grid box. An explanation for this is that NWP models are initialized better for frontal development in the central portion of the United States. For both clusters the CLM decreases through Europe and Asia, with higher values in Europe. It is assumed this is due to the observation system in Europe. Additionally, in both cases the CLM suffers in the western portion of the United States. This is presumably due to the lack of observations for incoming fronts from the Pacific as well as the frontal-terrain interactions of exiting fronts. In the southern hemisphere, the situations are reverse where the better CLM lies within cluster 2. In both instances, the CLM at 120 hours is low.

Comparison of Clusters 0/3 and 2/5 for All Times



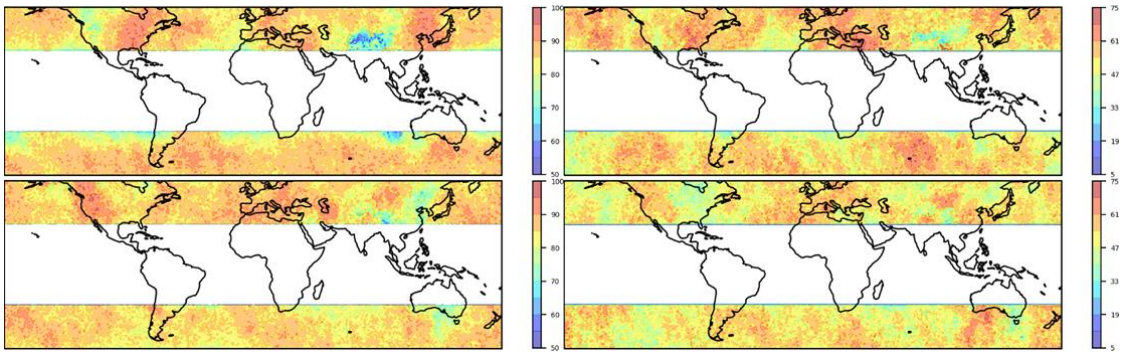


Figure 16. Probability of matching the forecast at 24 hours for winter for clusters 0 (top left) and 3 (bottom left) and 120 hrs for clusters 0 (top right) and 3 (bottom right).

Clusters 0 and 3 display higher CLM than clusters 2 and 5 at all times across the globe (Figure 16). This is because of the less dynamical situation presented to the models from clusters 0 and 3. For all cases the southern hemisphere performs with homogeneity and lack of major maxima or minima. The reason behind this finding is that the southern hemisphere has a lack of landmass interactions and acts more like an idealized "aqua planet" situation. However, the highest percentages of CLM almost always lie in the northern hemisphere because of the robust observation systems available.

4.2 V850 Winter Analysis

V850 Winter Analysis of Cluster 0 at + 24 and + 120 FH

Very high CLM exists throughout, which is the most of any season case. The Highest areas of CLM are over the central and eastern United States, the waters off the east coast of the United States, Europe, Iraq, and eastern Russia and China. The lowest area of CLM exists over the west coast of the United States. For the southern hemisphere, homogeneously high CLM throughout with the exception of a minimum off the west coast of Australia. For +120 FH, predictabilities decrease but retain the same areas of maximums and minimums.

V850 Winter Analysis of Cluster 3 at + 24 and + 120 FH

Cluster 3 shows a flip of the maximums and minimums when compared to cluster 0 (Figure 16). Best CLM now exists over the United States west coast and Europe. Higher CLM also exists over the Atlantic and Pacific Oceans. Areas of minimum CLM are the east coast of the United States and eastern Russia and China. For the southern hemisphere, areas of maximum CLM include the South Pacific, South Atlantic, and waters south of Australia with the exception of the waters off the southeast coast of Australia.

Comparison of Clusters 0/3 for Winter

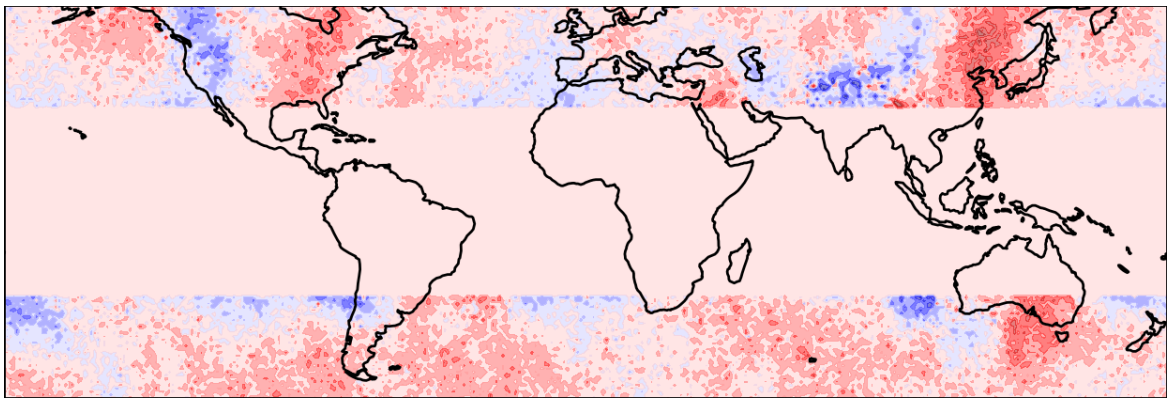
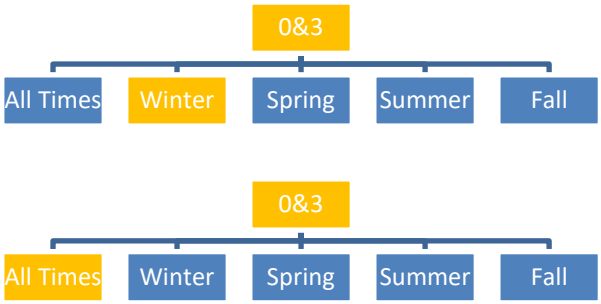


Figure 17. Difference between values from clusters 0 and 3 indicating areas of similar CLM (lighter colors).

In Figure 17, areas in the northern hemisphere with the most similarity are the Pacific, East Atlantic, and Europe. The most dissimilarity occurs over the west/northwest and eastern United States, Iraq, and East Asia. The southern hemisphere depicts similarity in the South Pacific and South Atlantic. The most dissimilarity occurs in the waters near South America, and east of Africa, especially Australia.

Both clusters exhibit high CLM. Cluster 0 has slightly higher values overall. This is due to reasons explained in the comparison of clusters 0 and 3 for all times. The same mirroring effect occurs in the winter as in all times for the same justifications as stated earlier. The same wavelike pattern appears in the southern hemisphere as well.

Comparison of Clusters 0/3 for Winter and for Clusters 0/3 All Times



Winter has higher CLM than all times for all areas and for both clusters 0 and 3. Additionally, the areas of maximums and minimums are unchanged from the all times. With repsect to cluster 0, it makes sense that the winter clusters predict the negative V850 anomaly better as that is the predominant wind direction in the northern hemisphere during that season. Another reason is that the mean is over 90 Julian days as opposed to a full year, which indicates that some seasons have high CLM and others have low CLM. For the southern hemisphere, there is very little change in the CLM, even while being in the opposite season from the northern hemisphere. There is only a very slight increase for cluster 0 and decrease for cluster 3. Additionally, cluster 3 improves in CLM for the winter analysis more than cluster 0. Improvement is expected because winter is not a transition season and that the models have fully adjusted to the climatological mean of the winter season.

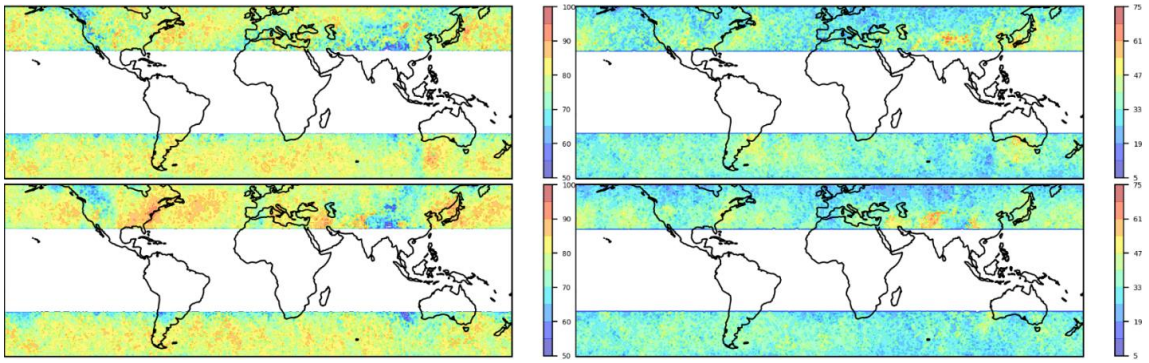


Figure 18. Probability of matching the forecast at 24 hours across winter for clusters 2 (top left) and 5 (bottom left) and 120 hrs for clusters 2 (top right) and 5 (bottom right).

V850 Winter Analysis of Cluster 2 at + 24 and + 120 FH

Overall, much less CLM than in cases 0 and 3 (Figure 18). For case 2, areas of highest CLM reside over the Pacific, central Unite States, Atlantic, and central and eastern Europe. Areas of lowest CLM exist over the west coast of the United States, western Europe, and most of southwest Asia and Asia. For the southern hemisphere, the probability is more homogenous albeit similar in CLM. Lowest area of CLM in the southern hemisphere is off the west and east coasts of Australia. For +120 FH, there is less CLM overall with the best areas being off the east coast of the United States, Turkey, the Pacific, off the coast of South America, and the southern coast of Australia.

V850 Winter Analysis of Cluster 5 at + 24 and + 120 FH

Cluster 5 also possesses much less CLM than cases 0 and 3 (Figure 18). Generally higher CLM in the northern hemisphere and slightly less in the southern hemisphere when compared to case 2. The Pacific, Atlantic, and central United States still remain areas of high CLM but with an increase and maximum over the east coast. Europe and

Asia increase in CLM throughout. For the southern hemisphere there exists slightly lower CLM, especially off the southern coast of Australia. For +120 FH, areas of highest CLM are California, east coast of the United States, Iraq, Iran, and Pacific.

Comparison of Clusters 2/5 for Winter

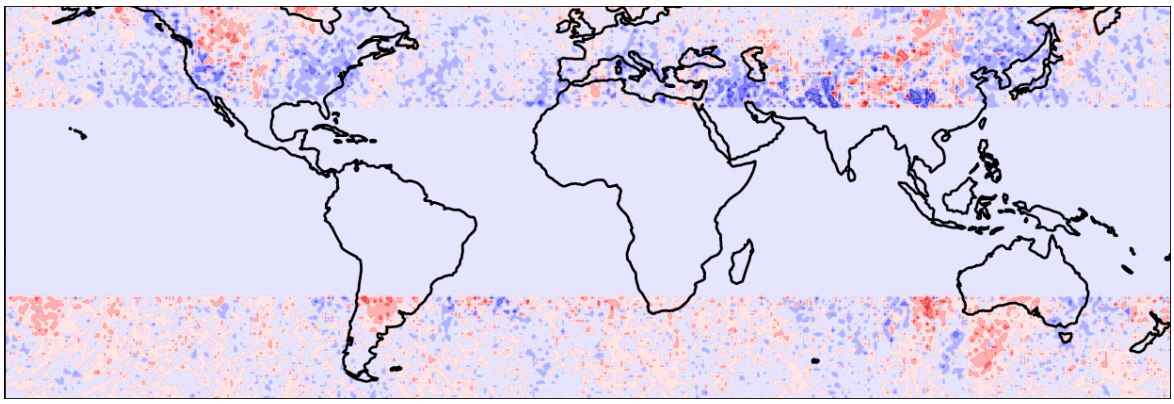


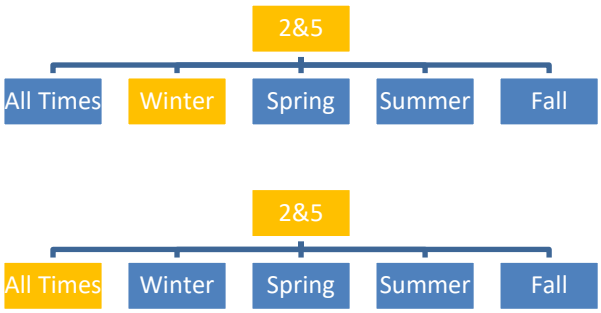
Figure 19. Difference between values from clusters 2 and 5 indicating areas of similar CLM (lighter colors).

In Figure 19, clusters 0 and 3 exhibit much more similarity than clusters 2 and 5. The areas of most similarity in the northern hemisphere are the Atlantic and Europe into northern Russia. Areas of most dissimilarity are the United States and southwest, central and east Asia/Russia. For the southern hemisphere, most areas are similar, with the southwest coast of Australia being the exception.

For the northern hemisphere, cluster 2 is much less predictable than cluster 5. The biggest increases of CLM between clusters 2 and 5 are in western Russia, the eastern

United States and both the Pacific and Atlantic Oceans. For both clusters the CLM decreases through Europe and Asia, with higher values in Europe. In the southern hemisphere, the situations are reverse where the better CLM lies within cluster 2. In both instances, the CLM at 120 hours is low.

Comparison of Clusters 2/5 for Winter and for Clusters 2/5 All Times



There is slight improvement for all areas and times for the winter as opposed to all times. The same justifications from the clusters 0 and 3 comparison of winter and all times apply. The area of biggest improvement is in cluster 5 along the east coast of the United States. Although the probability of matching the forecast improved with the seasonal filter, the improvement is not as drastic as with clusters 0 and 3.

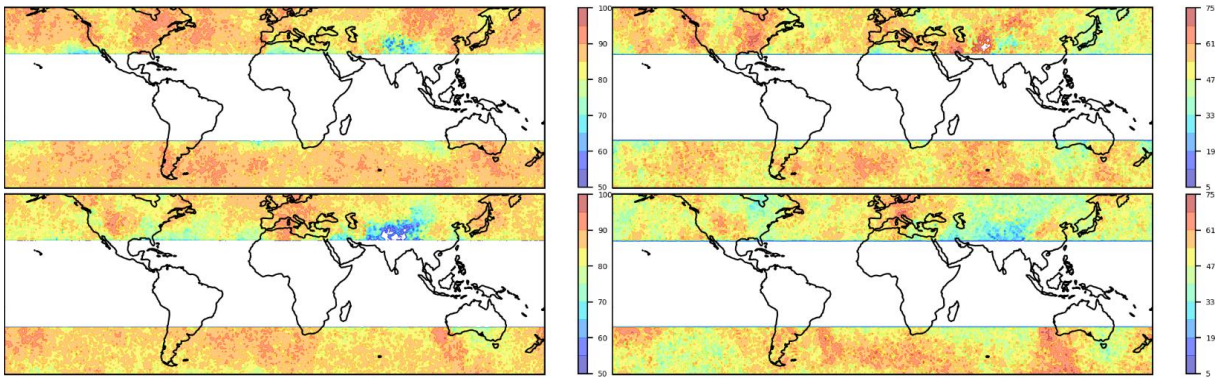


Figure 20. Probability of matching the forecast at 24 hours for spring for clusters 0 (top left) and 3 (bottom left) and 120 hrs for clusters 0 (top right) and 3 (bottom right).

4.3 V850 Spring Analysis

V850 Spring Analysis of Cluster 0 at + 24 and + 120 FH

Very high CLM exists throughout. Highest areas include Pacific, central and eastern United States, Atlantic, and central and eastern Russia and China (Figure 20). The lowest areas of CLM include the Rockies, west coast of the United States, eastern Europe, and eastern Russia and Asia. The southern hemisphere remains homogenous with high CLM throughout. The lowest area of CLM is off the western coast of Australia. For +120 FH, CLM stays higher in areas with previously high CLM but shows minimums over the Japan area of the Pacific and east Russia/China. Additionally, the highest area of CLM in the ATL off the E coast of the US turns into a minimum. For the southern hemisphere, the CLM pattern does not change much. The area of lowest CLM shifts from the west coast to southwest coast of Australia.

V850 Spring Analysis of Cluster 3 at + 24 and + 120 FH

Generally less CLM throughout than when compared to case 0 (Figure 20). The highest area of CLM over the United States shifts to the Rockies to exclude the west coast. The Pacific and Atlantic decrease significantly in CLM. western Europe increases in CLM whereas China/Russia decrease. For the southern hemisphere, the lowest area of CLM shifts from the west coast to the south/southeast coast of Australia. For +120 FH, the CLM pattern stays about the same. However, eastern Russia and Japan turn from a maximum into a minimum. Additionally, the northeast and northern United States decrease significantly in CLM. For the SH, biggest decrease is over the South Pacific and South Indian Ocean.

Comparison of Clusters 0/3 for Spring

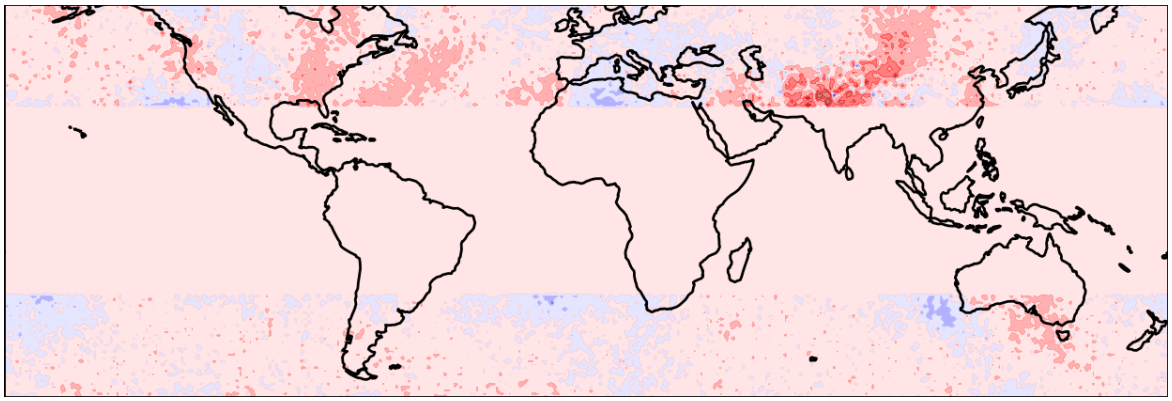


Figure 21. Difference between values from clusters 0 and 3 indicating areas of similar CLM (lighter colors).

In Figure 21, clusters 0 and 3 are similar with the greatest similarities in the southern hemisphere, as seen in Figure 21. In the northern hemisphere the Pacific is the most similar area. The most dissimilar areas are the east coast of the United States, Atlantic coastal waters, and central Asia. For the southern hemisphere all areas are very similar with the exception of Australia.

Both clusters exhibit high CLM. Cluster 0 has slightly higher values overall. This is due to reasons explained in the comparison of clusters 0 and 3 for all times. The mirroring effect which is evident in all seasons and all times for clusters 0 and 3 remains a factor, but is less pronounced than in all times and winter. This is, in part, due to the transitional nature of spring time weather. Additionally, as temperature gradients modify due to the warming of the northern hemisphere, the drastic cutoff caused by the Rocky Mountains becomes less pronounced. The southern hemisphere remains relatively the same for both clusters with only minor variations over Australia. For +120 FH, the opposite nature of CLM within clusters 0 and 3 becomes much more evident. Although the Rocky Mountain split is in place, the CLM over the west coast of the United States is still less than for the eastern half of the country. The justification for this is explained earlier with regards to observations prior to the west coast versus the east coast. Additionally, CLM over Russia and central Asia suffers much more in case 3 +120 FH versus case 0 +120 FH.

Comparison of Clusters 0/3 for Spring and for Clusters 0/3 All Times

Generally, the placement of high CLM versus low CLM remains the same to include the +120 FH for each respective cluster. The CLM or the northern hemisphere is better for case 0 in the spring versus all times. This is most likely due to the fact that NWP models have a full recent data set for negative V850 anomalies from the winter months. This also can explain why the positive V850 anomalies in cluster 3 are not predicted as well as cluster 0. The more diffuse nature of the Rocky Mountain gradient can be explained by the reasoning in the previous section. The CLM over the southeast United States is less than cluster 3 for all times. Again, this is a result of the persistence forecast of NWP models following the winter months. For the southern hemisphere, there is slightly more CLM in the spring cases for all areas.

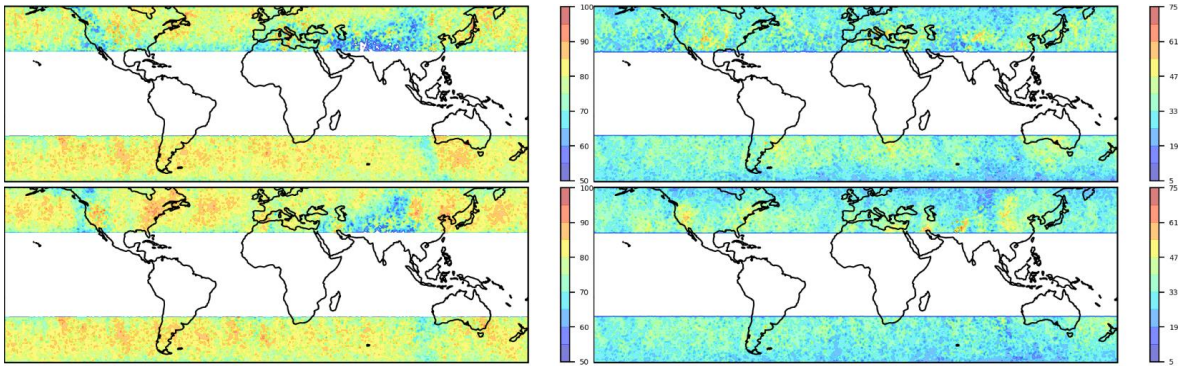


Figure 22. Probability of matching the forecast at 24 hours for spring for clusters 2 (top left) and 5 (bottom left) and 120 hrs for clusters 2 (top right) and 5 (bottom right).

V850 Spring Analysis of Cluster 2 at + 24 and + 120 FH

Much less CLM exists for cluster 2 and 5 than clusters 0 and 3 (Figure 22). For cluster 2, the Atlantic and Pacific are predicted well. Most of the United States, with the exception of the west coast, is predicted well. eastern Europe through central Russia, with the exception near Mongolia and eastern Russia and China are predicted well. A wave pattern is noticeable within the areas of maximum prediction throughout both hemispheres, with it manifesting itself better in the southern hemisphere. For the southern hemisphere, homogeneous CLM is evident with the least CLM off the western coast of Australia. For +120 FH, there is much less CLM with the highest area over the central United States.

V850 Spring Analysis of Cluster 5 at + 24 and + 120 FH

For cluster 5, there is higher CLM than in cluster 2 (Figure 22). The only major changes are that the east and west coasts of the United States and Canada are now a maxima as well as Central/East central Asia, and China and eastern Russia. More apparent than in cluster 2 is the wave like pattern of maxima. For the southern, the wave pattern is also more apparent than in case 2. What was a maximum for the southern coast of Australia now shifts to the east coast. For +120 FH, there are no major differences over all areas.

Comparison of Clusters 2/5 for Spring

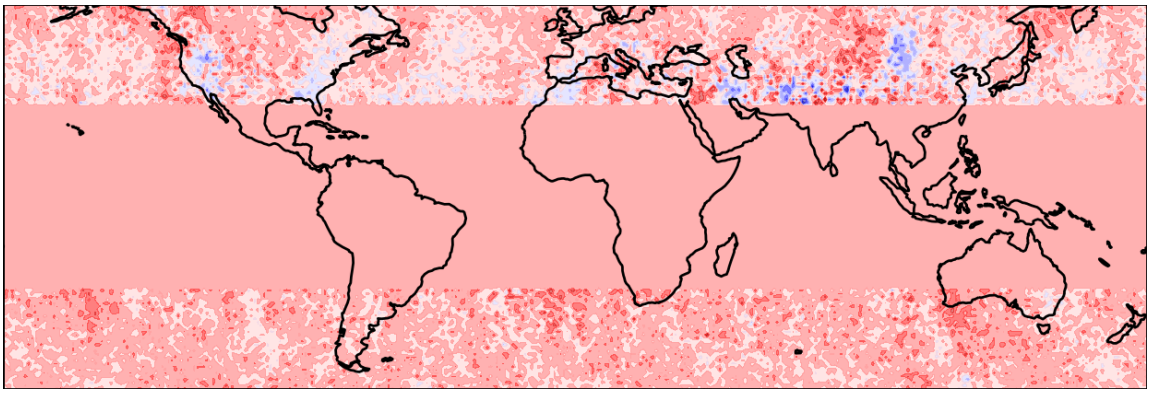
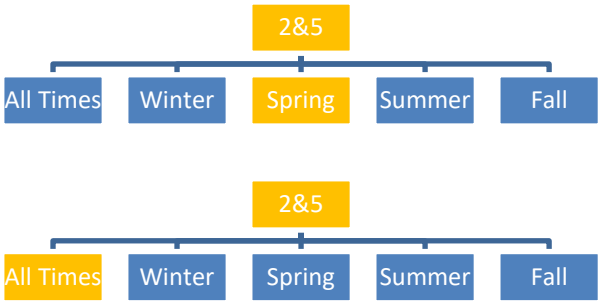


Figure 23. Difference between values from clusters 2 and 5 indicating areas of similar CLM (lighter colors).

As depicted in Figure 23, there is very little similarity between clusters 2 and 5 for the spring months. The most similarity occur within the central Atlantic. Cluster 5 exhibits more CLM than cluster 2. The reasoning for this is mentioned in the all times cluster 2 and 5 comparison. The biggest increases of CLM between clusters 2 and 5 are in western Russia, the eastern United States and both the Pacific and Atlantic Oceans. For both clusters the CLM decreases through Europe and Asia, with higher values in Europe. In the southern hemisphere, the situation is reversed where the slightly better CLM lies within cluster 2. In both instances, the CLM at +120 FH is low.

Comparison of Clusters 2/5 for Spring and for Clusters 2/5 All Times



The differences between clusters 2 and 5 for all times versus in the spring are minor. The CLM in the spring is only slightly higher for all areas. The other main difference is that a wave pattern is evident in the spring clusters as opposed to all times. A reason for this is that all times may be diluted as the 15 years of data may "wash out" minor details. For both cases, +120 FH shows very low CLM. The dynamic nature of split positive negative clusters is more difficult to predict the further one goes into the future.

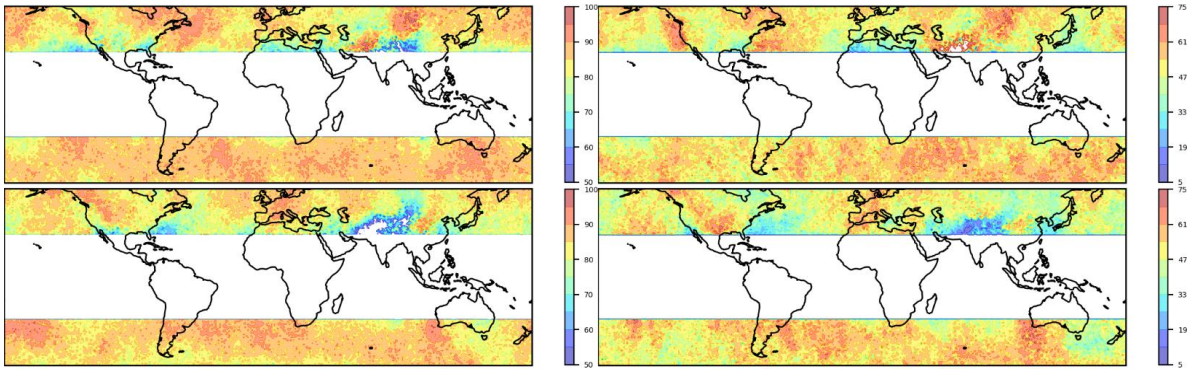


Figure 24. Probability of matching the forecast at 24 hours for summer for clusters 0 (top left) and 3 (bottom left) and 120 hrs for clusters 0 (top right) and 3 (bottom right).

4.4 V850 Summer

V850 Summer Analysis of Cluster 0 at + 24 and + 120 FH

Generally high CLM exists throughout with the highest areas over the northern United States, central/eastern Canada and the northeastern United States (Figure 24). The higher CLM also includes the waters off the eastern United States coast. Additional areas of high CLM include Iran, north-central through eastern Russia, eastern China, and the Pacific and Atlantic. For the southern hemisphere, there exists homogeneous high CLM throughout. The maximum is off the southern coast of Australia. For +120 FH, the west

coast of the United States becomes an area of high CLM as well as eastern Europe. Otherwise, no major changes.

V850 Summer Analysis of Cluster 3 at + 24 and + 120 FH

There is a general mirroring effect between clusters 0 and 3, however, boundaries between high and low CLM are not drastic (Figure 24). CLM in the United States increases in general. A maximum is located from the Rockies westward, with the exception of the west coast. The eastern United States and Canada are no longer maximums, when compared to cluster 0. What was a minimum near the United Kingdom is now a maximum. CLM increases throughout Europe as a whole and extends further into Russia. There is still decent CLM over China and eastern Russia, but less than in case 0. The Pacific decreases in CLM. For the southern hemisphere, there exists slightly less CLM throughout. A new minimum forms off the southeast coast of Australia. For +120 FH, no major changes for either of the hemispheres.

Comparison of Clusters 0/3 for Summer

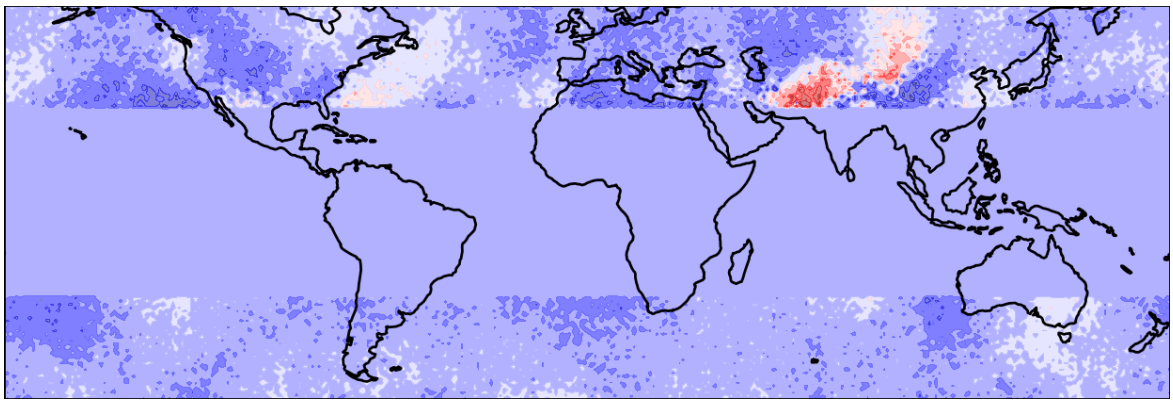


Figure 25. Difference between values from clusters 0 and 3 indicating areas of similar CLM (lighter colors).

In Figure 25, areas of most similarity are off the eastern coast of the United States and in the South Pacific. The greatest dissimilarity exists over Iran and central Asia. For the southern hemisphere, the most dissimilarity exists in the southwest Pacific, South Atlantic off the coast of Africa, and southwest Australia.

Overall, both clusters have high CLM. There is a general reversing of the predictabilities when comparing the clusters, however it is not as robust of a flip as has been observed in previous seasons. This can be due to the modified temperature gradient which reduces the uncertainty with the air-terrain interactions. The southern hemisphere remains homogenous, again, for reasons explained in the all times section.

Comparison of Clusters 0/3 for Summer and for Clusters 0/3 All Times

For cluster 0 there is considerably less CLM in the United States, especially in the southern portions, as well as eastern Russia and China for the summer months. There is a trend of less CLM along the southern periphery of the northern hemisphere. It is possible that negative V850 anomalies are more difficult to predict during months that are dominated by positive V850 anomalies. It also appears that there is a wave-like pattern for both hemispheres for the summer months. The waves are most likely present in the all times analysis as well, but washed out due to the averaging of data. However, when viewed within a smaller timescale, they are possible to detect. The wave pattern in the northern hemisphere can be seen with higher CLM over the West Pacific, less through the Pacific and southern United States, higher CLM over the Atlantic, and back to a decrease in CLM over North Africa and central Asia. The wave pattern in the southern hemisphere is less apparent but still present. For +120 FH, one feature that stands out is the higher CLM off the west coast of the United States. The predictable semi-permanent high that migrates northward during the summer months can be a reason for this increase.

For cluster 3, there are no major differences between all times and summer months. There is a slightly more apparent wavelike feature in the southern hemisphere for the summer months than all times. There are no major differences between the seasonal versus all times for cluster 3 when it comes to +120 FH.

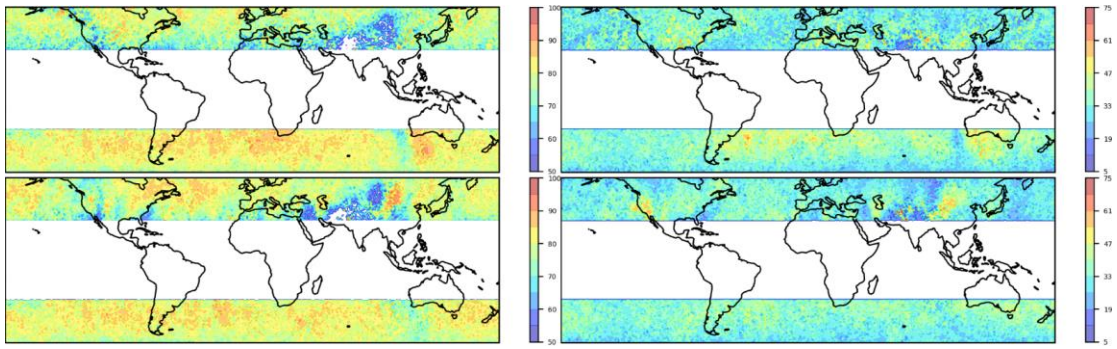


Figure 26. Probability of matching the forecast at 24 hours for summer for clusters 2 (top left) and 5 (bottom left) and 120 hrs for clusters 2 (top right) and 5 (bottom right).

V850 Summer Analysis of Cluster 2 at + 24 and + 120 FH

Higher CLM exists for the Pacific, Atlantic, and most of the United States and Canada with the exception of the Pacific coast (Figure 26). Higher CLM also exists across northern Europe and Russia as well as eastern Russia. Lower CLM exists along the southern periphery of the northern hemisphere, western and central Europe, and Asia. For the southern hemisphere, there is higher CLM throughout. The maxima for the southern hemisphere exist off the southwest coast of Australia, the South Atlantic, and the southeast Pacific. For +120 FH, the northern hemisphere drops in CLM drastically, while the southern hemisphere retains higher areas of CLM across the maxima described previously.

V850 Summer Analysis of Cluster 5 at + 24 and + 120 FH

An apparent wavelike pattern is visible within the maxima for both hemispheres (Figure 26). Otherwise, moderate CLM resides throughout with a sharp maximum over Mongolia and sharp minimum just to its west. Another sharp minimum exists over

Nevada, the central plains and just off the Carolina Coast. For the southern hemisphere there exists homogeneous moderate CLM with a minimum off the southwest coast of Australia. For +120 FH, there are no major changes from +24 FH. There are two maximums with one being over the west coast of the United States and another over Mongolia. For the SH, the maximum over southwest Australia disappears.

Comparison of Clusters 2/5 for Summer

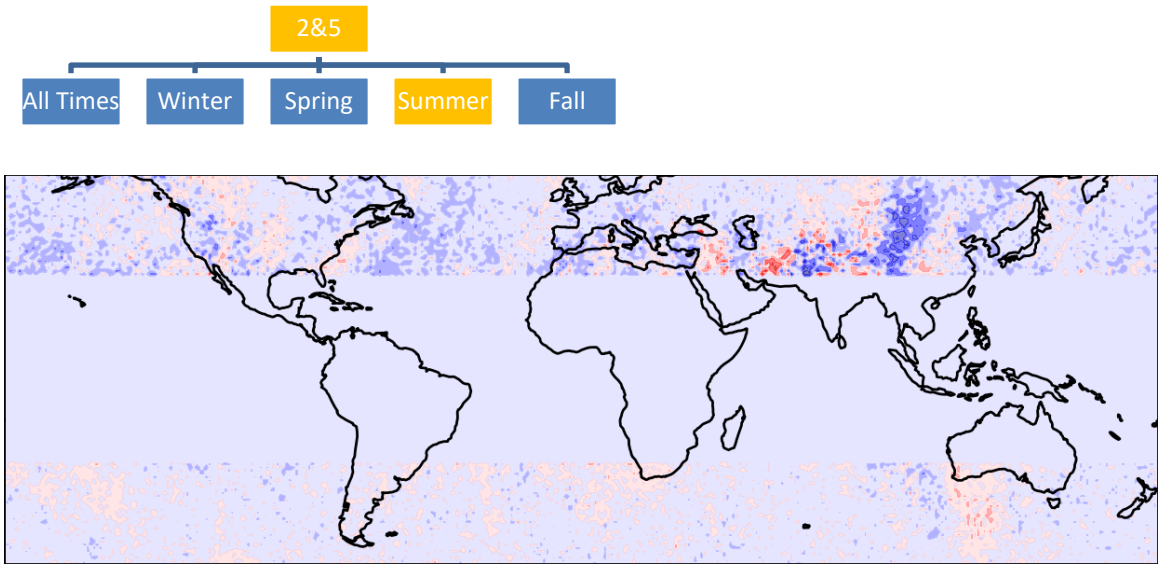
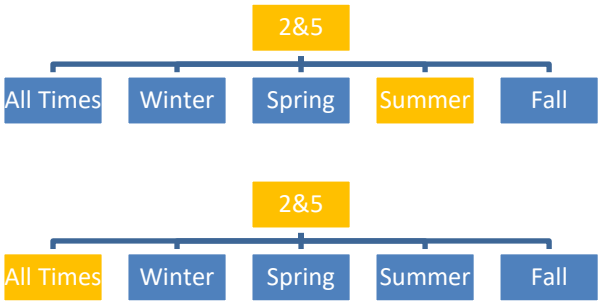


Figure 27. Difference between values from clusters 2 and 5 indicating areas of similar CLM (lighter colors).

There is more similarity between clusters than cases 0 and 3, especially in the southern hemisphere (Figure 27). The United States is similar along with the Atlantic, Pacific, and Europe/Russia. The most dissimilarity exists throughout Asia and eastern Russia. The northern hemisphere is more accurately predicted in cluster 5 whereas the southern hemisphere is more accurately predicted in cluster 2. This is due to the

hemispheres being in two opposite seasons. The cluster with the most CLM overall is cluster 2. This is for reasons explained in the all times section. Aside from CLM, the biggest difference is the apparent jagged placement of maximums and minimums in case 5. For example, the central/west central United States has a sharp minimum followed by a maximum and then another minimum. This pattern also exists less dramatically over the Atlantic. Another area of stark differential exists over eastern Asia. An explanation for this type of behavior is that because cluster 5 is split with negative anomalies on the west and positive on the east, one can expect a binary CLM pattern. When referencing clusters 0 and 3 one can notice a similarity between areas that do not have a high CLM for negative V850 anomalies in cluster 0 also do not have high predictabilities in cluster 5.

Comparison of Clusters 2/5 for Summer and for Clusters 2/5 All Times



There are no major changes in CLM between summer and all times for clusters 2 and 5. The biggest difference is how apparent the wavelike pattern becomes in the summer iteration. This is most likely due to a smaller sample size as well as how it is more difficult to predict negative anomalies in a season dominated by positive anomalies. The southern hemisphere exhibits the same tendency, albeit less apparent. Additionally,

the southern hemisphere exhibits the best wavelike patterns in cluster 2 as opposed to 5 due to the hemispheres being in opposite seasons.

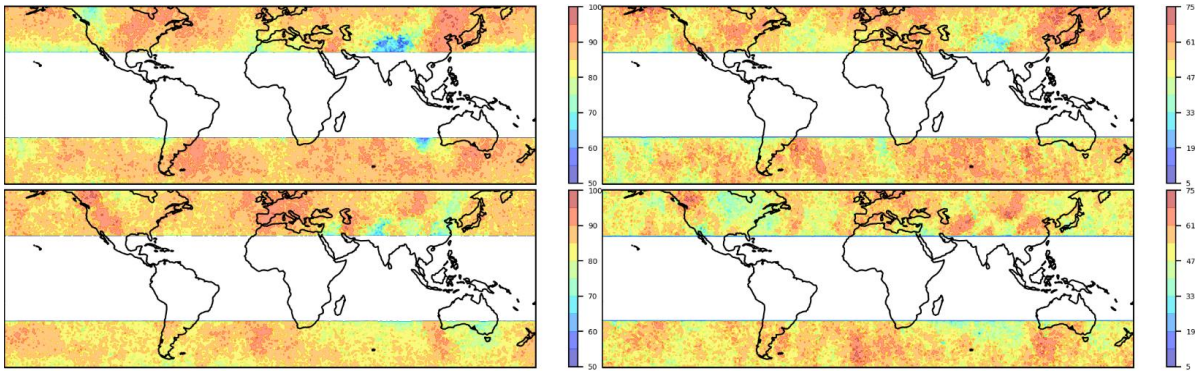


Figure 28. Probability of matching the forecast at 24 hours for fall for clusters 0 (top left) and 3 (bottom left) and 120 hrs for clusters 0 (top right) and 3 (bottom right).

4.5 V850 Fall Analysis

V850 Fall Analysis of Cluster 0 at + 24 and + 120 FH

Generally high CLM exists throughout with highest CLM over the central and eastern United States, Atlantic, and eastern China/Russia (Figure 28). Lower CLM exists over the northwest United States and Canada, western Europe, northern Africa, and central Asia. For the southern hemisphere, high CLM exists throughout with the exception of just off the west coast of Australia. For +120 FH, moderate to high CLM exists throughout with no major changes in areas of CLM for the northern hemisphere. For the southern hemisphere, the minimum off the western coast of Australia disappears and one appears off the western coast of Africa.

V850 Fall Analysis of Cluster 3 at + 24 and + 120 FH

Higher CLM is apparent in cluster 3 when compared to cluster 0 (Figure 28). Two areas are marginally opposite of case 0 which are western China/Russia and the United States. Iraq becomes a minimum. The southern hemisphere is less predictable overall with a minimum off the southeast coast of Australia. For +120 FH, the only thing that doesn't follow the pattern from + 24 FH is that E Asia increases in CLM. For the southern hemisphere, the minimum that was off the southeast coast of Australia shifts to the western Australia coast.

Comparison of Clusters 0/3 for Fall

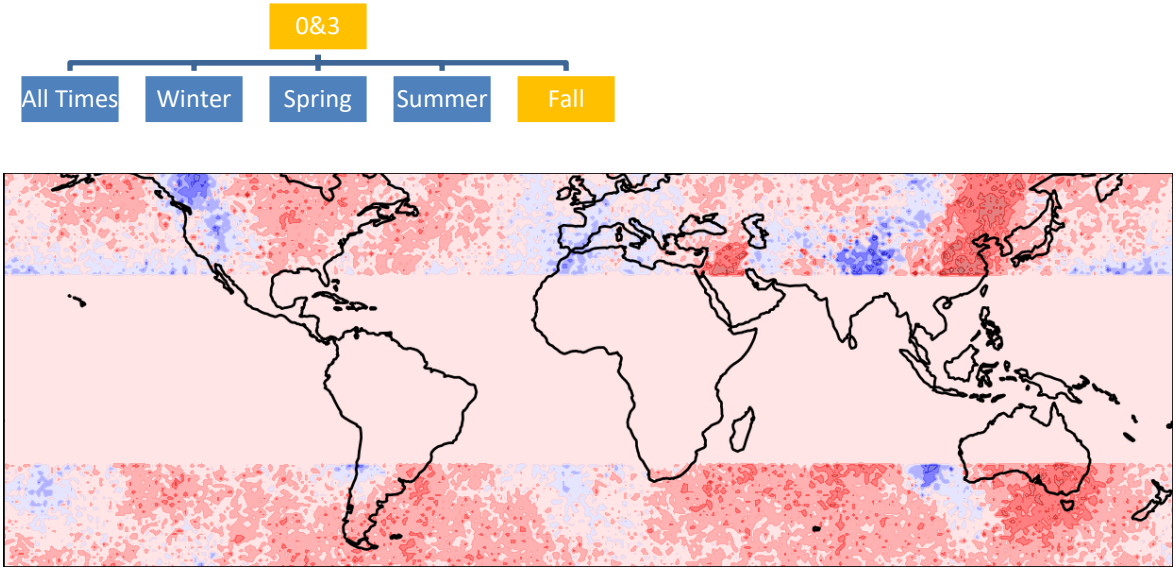


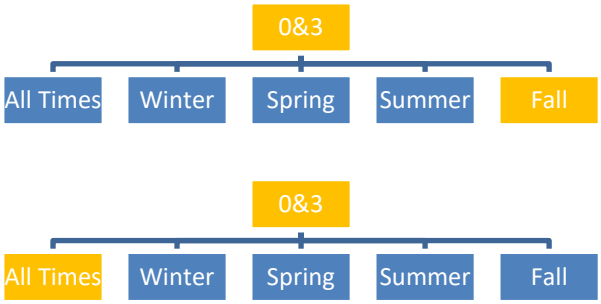
Figure 29. Difference between values from clusters 0 and 3 indicating areas of similar CLM (lighter colors).

In Figure 29, one can see that there is much more dissimilarity than similarity. The areas with the most similarity in the northern hemisphere exists over the eastern Atlantic, Europe, and portions of central Asia. The areas with the most dissimilarity are the United States, western Atlantic, northern portions of the northern hemisphere, and especially

eastern Asia/Russia. For the southern hemisphere, the areas with the most similarity are the southwest Pacific, South Atlantic off the coast of Africa, and The southwest waters off Australia. Everywhere else is dissimilar with the most dissimilarity off the southeast coast of Australia.

Cluster 0 has higher CLM overall than cluster 3. Clusters 0 and 3 are marginally opposed for areas of high and low CLM. The main areas of opposition are the US, Europe, eastern Russia/China and southern Australia. As the semi-permanent high pressure system moves to the south, the Aleutian Low moves closer to the west coast of the United States, helping in making cluster 3 more predictable. Interestingly, +120 FH for cluster 3 performs very poorly over the eastern half of the United States. This is due to the fact that fall is transitioning into a period of predominantly negative V850 anomalies, thus, predicting positive V850 anomalies becomes more difficult further into the future.

Comparison of Clusters 0/3 for Fall and for Clusters 0/3 All Times



Fall exhibits higher overall CLM than all times. The pattern of high CLM versus low CLM remains about the same. Cluster 5 is also predicted more frequently in the fall than for all times because the NWP model has acquired a strong persistence guidance for positive V850 anomalies by this point in the year. Otherwise, there are no other major differences between all times and the fall.

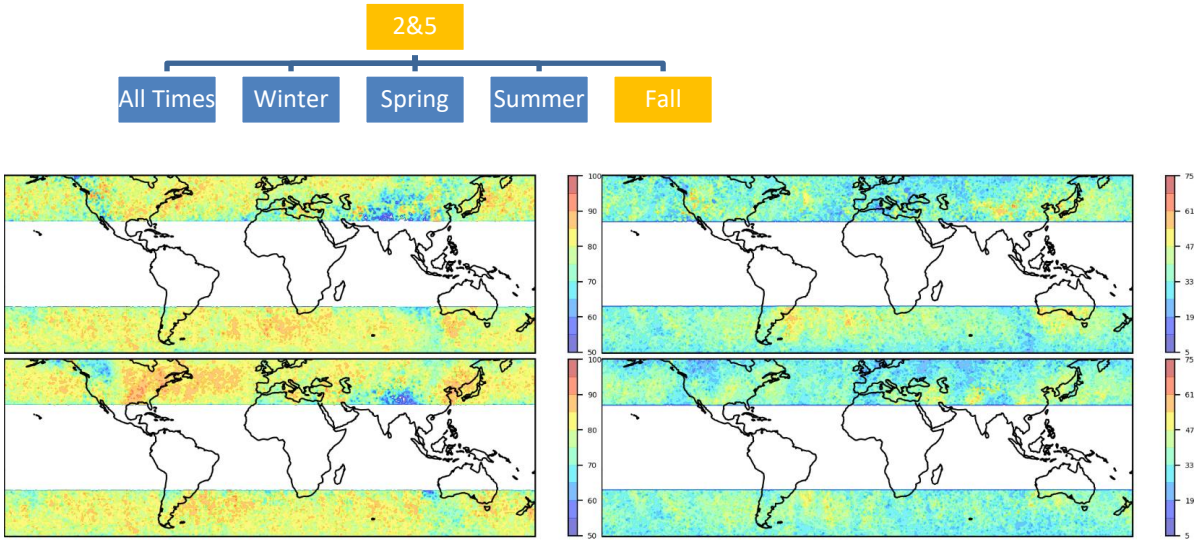


Figure 30. Probability of matching the forecast at 24 hours for fall for clusters 2 (top left) and 5 (bottom left) and 120 hrs for clusters 2 (top right) and 5 (bottom right).

V850 Fall Analysis of Cluster 2 at + 24 and + 120 FH

Moderate CLM exists throughout, although less than cases 0 and 3 (Figure 30). Highest areas of CLM are the southern United States, waters of the east coast of the United States, and waters off the east coast of Japan. A minimum exists over the west coast of the United States, Iraq, and central Asia. For the southern hemisphere, moderate CLM is in place throughout with a minimum off the west coast of Australia. For +120 FH, no major changes for the northern hemisphere or the southern hemisphere.

V850 Fall Analysis of Cluster 5 at + 24 and + 120 FH

High CLM exists throughout (Figure 30). A maximum is in place over the Midwest and eastern United States which stretches across the Atlantic. CLM in Europe is

moderate, but there is no real maximum. Another maximum is over eastern China/Russia. A minimum exists over the western United States and Canada. For the southern hemisphere, a wavelike pattern is marginally visible. There is a minimum off the southwest coast of Australia. For +120 FH, it becomes easier to see the wavelike pattern in the northern hemisphere, but no major changes to CLM areas.

Comparison of Clusters 2/5 for Fall

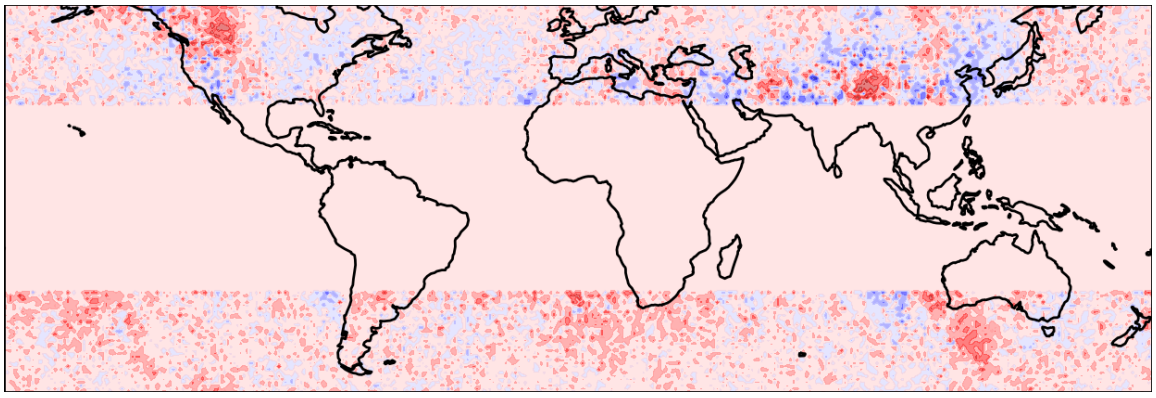
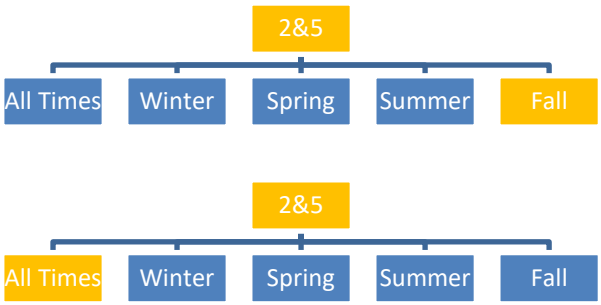


Figure 31. Difference between values from clusters 2 and 5 indicating areas of similar CLM (lighter colors).

In Figure 31, one can see that most areas are predicted similarly with respect to clusters 2 and 5 for the fall months. The areas with the most similarity in the northern hemisphere are the oceans. The most dissimilar areas are the western united states and central/eastern Asia. For the southern hemisphere, the most dissimilarity occurs off the

southwest coast of Australia. Cluster 5 has a higher CLM for the northern hemisphere, but cluster 2 has a higher CLM for the southern hemisphere. This is for reasons explained in previous comparison sections. Most major features noted in other clusters 2 and 5 analysis are present in the fall analysis.

Comparison of Clusters 2/5 for Fall and for Clusters 2/5 All Times



As with every seasons, the CLM for fall is greater than that of all times. This is the most evident for the east coast of the United States and Atlantic Ocean in cluster 5. This is most likely due to the fact that fall is transitioning to more negative V850 anomalies than positive, therefore, the cluster that most accurately represents cold air advection should be the cluster that is most predictable.

4.6 Comparison of Clusters 0/3 for All Seasons and Times

Winter and fall are the most predictable seasons for both clusters 0 and 3. The differences are extremely slight. The highest values lie within the winter months, however there is slightly more homogeneity to the fall predictabilities. There is a slight positive bias for CLM of positive anomalies for the fall months and negative anomalies for the winter months. This plays in favor for a slight edge to the fall months, because not only does it perform strongly with negative anomalies, the models also have the best database for persistence forecasting of positive anomalies.

One can notice that during the summer months, the bottom portion of the northern hemisphere analysis has wavelike areas of low CLM, and if one were to look at the winter months, the same pattern appears in the southern hemisphere. This bolsters the idea that in months where the V-component is predominantly positive, then CLM will suffer. This can be due to the development of low pressure systems and an unstable atmospheric environment from warm air advection(positive V850 anomalies). Additionally, cold air advection drives atmospheric motion, whereas the warm sectors can be more turbulent and unpredictable. Furthermore, predicting positive V850 anomalies is less accurate than negative anomalies in most cases with several exceptions: (i) along the west coast of the United States, (ii) during the summer months in both hemispheres, (iii) Australia, (iv) central and southwest Asia.

For the western United States, this is due to the seasonal fluctuations between the Aleutian Low and subtropical high. Another area of interest is the dichotomous CLM across Australia with respect to positive and negative V850 anomalies. The CLM there is driven by the dynamic interaction between the trade winds, subtropical ridge and westerlies. On average, the western side of Australia will experience better prediction of positive anomalies versus the eastern side with better prediction of negative anomalies.

Another area of interest for drastic changes in the prediction of positive V850 anomalies is central Asia. During the spring and summer months the Indian Summer Monsoon as well as the East Asia Summer Monsoon are active. One can see a reflectance of higher CLM for positive V850 anomalies in the summer and spring months along the East Asia coast indicative of the summer monsoon pattern. There is also a lack of CLM in positive V850 anomalies over much of central Asia during the summer and spring

months. This may seem counterintuitive, but as the thermal gradient relaxes north of the Tibetan Plateau, the weather patterns will become less predictable than when the Siberian High is bringing predominantly negative V850 anomalies to the region during the winter months. This is why the CLM of negative V850 anomalies increases throughout central Asia for the winter months.

Lastly, the CLM in southwest Asia (primarily Iraq) has interesting characteristics due to the summer monsoon pattern. The CLM of negative V850 anomalies is higher in every case, which is in line with the general trend globally. However, there is a period during the summer months that both the negative and positive forecasts perform equally as poorly. Climatologically, the pattern during the summer in Iraq is fairly unchanging, so a lack of CLM must be explained another way. Similarly, the southwest United States displays a similar trend. Both regions are under a summer monsoon pattern during these times. The difference between the two regions is that the southwest United States has higher CLM for positive V850 anomalies most times of the year, which is not the case for Iraq. Although the East Asian and Indian summer monsoons are predicted more accurately, it is possible that because the smaller spatial scale does not allow for a more accurate representation of the North American Monsoon and summer pattern in Iraq.

4.7 Comparison of Clusters 2/5 for All Seasons and Times

Clusters 0 and 3 have higher CLM than clusters 2 and 5 in all cases. The justification for this finding is that models will perform better in environments of lesser flux. Clusters 2 and 5, by nature, indicate a transitioning pattern, and the reduction in CLM reflects such. The southern hemisphere has higher CLM than the northern

hemisphere in general, however, the highest CLM values lie within the northern hemisphere. The higher values in the northern hemisphere is most likely due to a better coverage in observations. In all cases for the northern hemisphere, cluster 5 has higher CLM than cluster 2. When it comes to the southern hemisphere, clusters 2 and 5 perform similarly, with the exception of Australia. With respect to cluster 5 having higher CLM in all cases in the northern hemisphere, cold air advection and approaching stability is the main reason as cluster 5 represents negative V850 anomalies on the left and positive V850 anomalies on the right halves of the 19x19 grid box.

As clusters 2 and 5 with respect to all times have already been described, the focus is directed towards areas with anomalously high or low CLM. Additionally, comparisons will be drawn between all seasons. The highest predictabilities lie within cluster 5 for fall and winter, which aligns with the predominantly negative V850 anomaly nature of the cooler months. One area that does not follow any of the standard patterns in British Columbia which performs poorly for all clusters and times. Climatologically speaking, British Columbia experiences rainfall year round which is due to upslope but also, in part, due to opposing air masses. Because of this, one expects this to be reflected in the split V anomaly charts. Therefore, terrain must be a factor in the low CLM as mountains line the western coast of Canada. One can notice that the poorest area of CLM is along the mountain range. This is likely due to the terrain along the western coast of Canada.

Compared to cluster 2, cluster 5 is more predictable over the western coast of the United States at all times. As is the case with clusters 0 and 3, the predictabilities along the west coast are driven by the subtropical high and Aleutian Low dynamics. For a

majority of the United States, cluster 2 is moderately predictably for all times, although cluster 5 has higher CLM cluster 2 in all cases but summer because of negative V850 anomalies being less common during this time. The other area within the United States with comparatively high CLM is the east coast for cluster 5 in all cases with the exception of summer. The justification for this finding is provided in the clusters 2 and 5 all times analysis. The lower CLM off the Carolina Coast during the summer months is due to the Bermuda High seasonality. Additionally, the anomalies in Asia and Australia are caused by the factors mentioned previously.

An interesting feature exists along the southern periphery of the northern hemisphere for the spring and summer months and the northern periphery of the southern hemisphere in the winter months. It is noted there is significantly less CLM in these areas during these times. This is due to the polar jet advancing more northward, thus, reducing the times patterns reflect those identified in clusters 2 or 5. It is also noted that this phenomena is not as identifiable in clusters 0 and 3, but this is justifiable because the requirement for clusters 2 and 5 is one that exists predominantly along frontal boundaries which do not occur in the same intensity and or frequency as one approaches the equator.

Furthermore, wave-like phenomena appears more readily in clusters 2 and 5 than clusters 0 and 3. If this were due to the physical negative and positive anomalies, one expects to notice such a pattern in clusters 0 and 3. However, this is not the case, and another explanation is necessary but beyond the scope of this work. A check is performed to deduce if there is seasonal influence. The most notable wave patterns occur in cluster 5 in this northern hemisphere during the summer months while it is the spring months for the southern hemisphere. Additionally, the wave pattern is always more apparent in the

southern hemisphere. When observing all times analysis for clusters 2 and 5 it is apparent that the wave patterns do not always offset as there are still faint wave patterns identified.

4.8 Case Study

A case study is conducted in order to relate the findings above to real-world events. The time chosen for this case study is for April of 2011 for the southeast United States centered on Tennessee. This region and time is chosen because of the active nature of the weather patterns during this time. This variability in the weather pattern is evident in Figures 32 and 33 which represents the changes in ensemble members for one day in April and December. December is chosen as a comparison with April based on the findings in the analysis that the CLM is high in winter and low in spring. One can see that the initial ensemble members fall into the same cluster categorization for both months, but as time progresses, variations start to manifest. This is due to the initial perturbations of the ensemble members. However, the variation is much higher within April when compared to December. This finding is also evident in the cluster analysis for spring versus winter as discussed previously. Not only do the figures provide a visualization of the variability, but also provide insight as to the certain types of patterns that fall into certain clusters.

Figure 32. 11 ensemble members (across) changes over +24, +72, and +120 FH (down) for April 2011. The cluster number is assigned for each box (top left).

Figure 33. 11 ensemble members (across) changes over +24, +72, and +120 FH (down) for December 2011. The cluster number is assigned for each box (top left).

Additionally, the number of times that any of the clusters are correctly predicted is calculated for + 24, +72, and +120 FH. The results can be seen in Figure 34. As expected, the number of matches decreases with increased forecast hour. There is a marked decrease in matching cluster forecasts between ~20 and 30 April for the +24 FH. Satellite imagery and surface analysis is retrieved from the Weather Prediction Center which show an increase in frontal activity. This is a likely factor in the decrease of CLM. There is in increase in CLM following this period, which is accompanied by development of strong positive winds over the study area. When compared to each other, the trends from the different forecast hours decrease and increase at relatively the same times. The main difference between the forecast hours is the number of matches per the last sampling period of 28 days with +120 and +24 FHs having the least and most matches respectively.

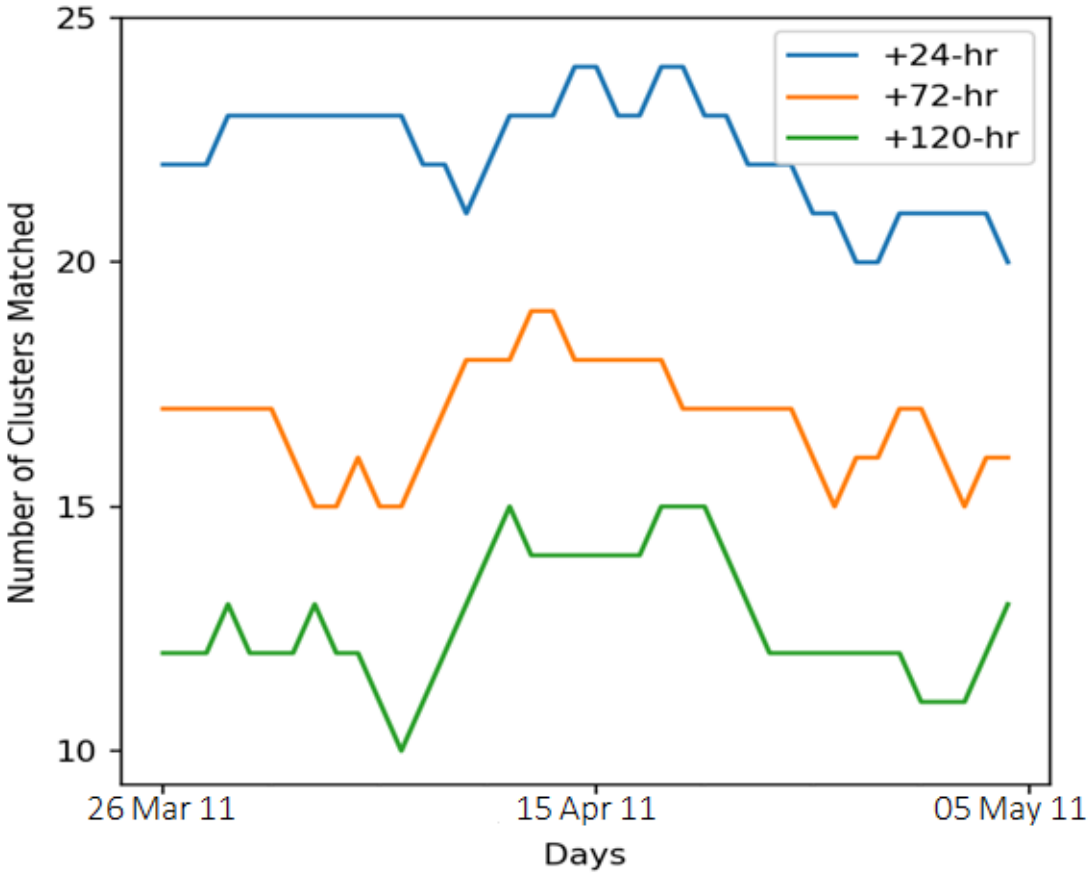
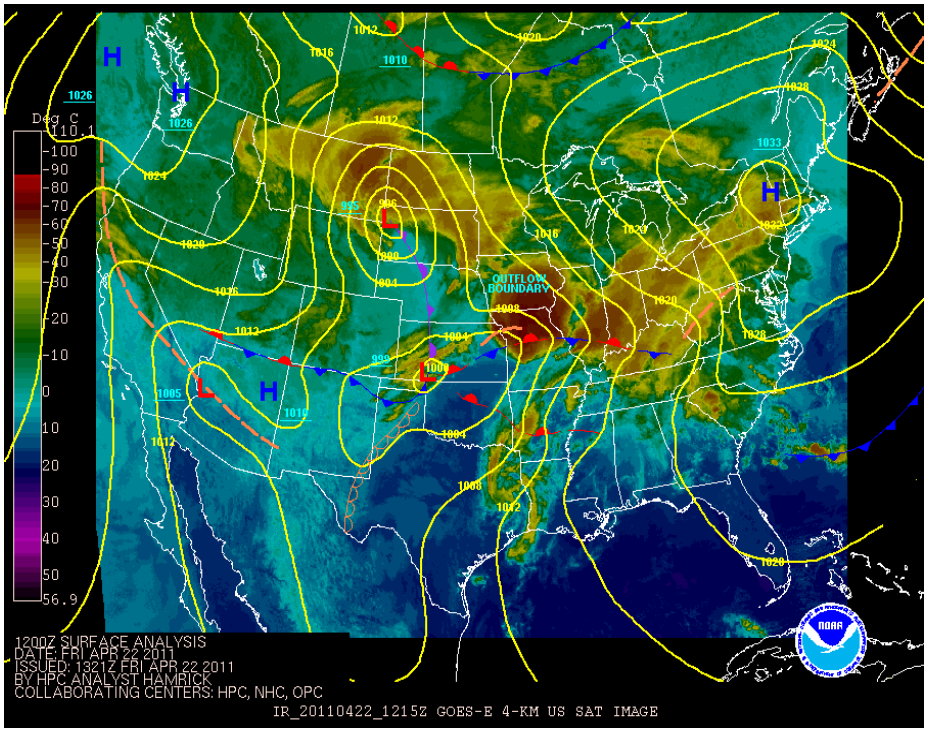
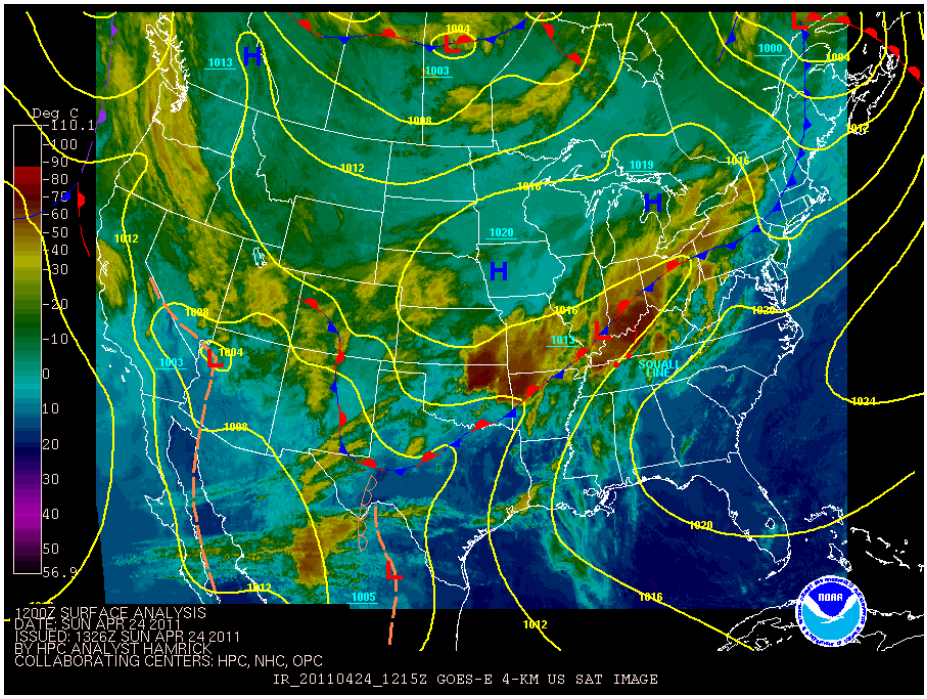


Figure 34. Number of times any of the 6 clusters are matched per their respective forecast hour for dates March 26- May 5 2011 centered on April 15th. For example, on April 15th there were 24 matches within the 28 days prior.





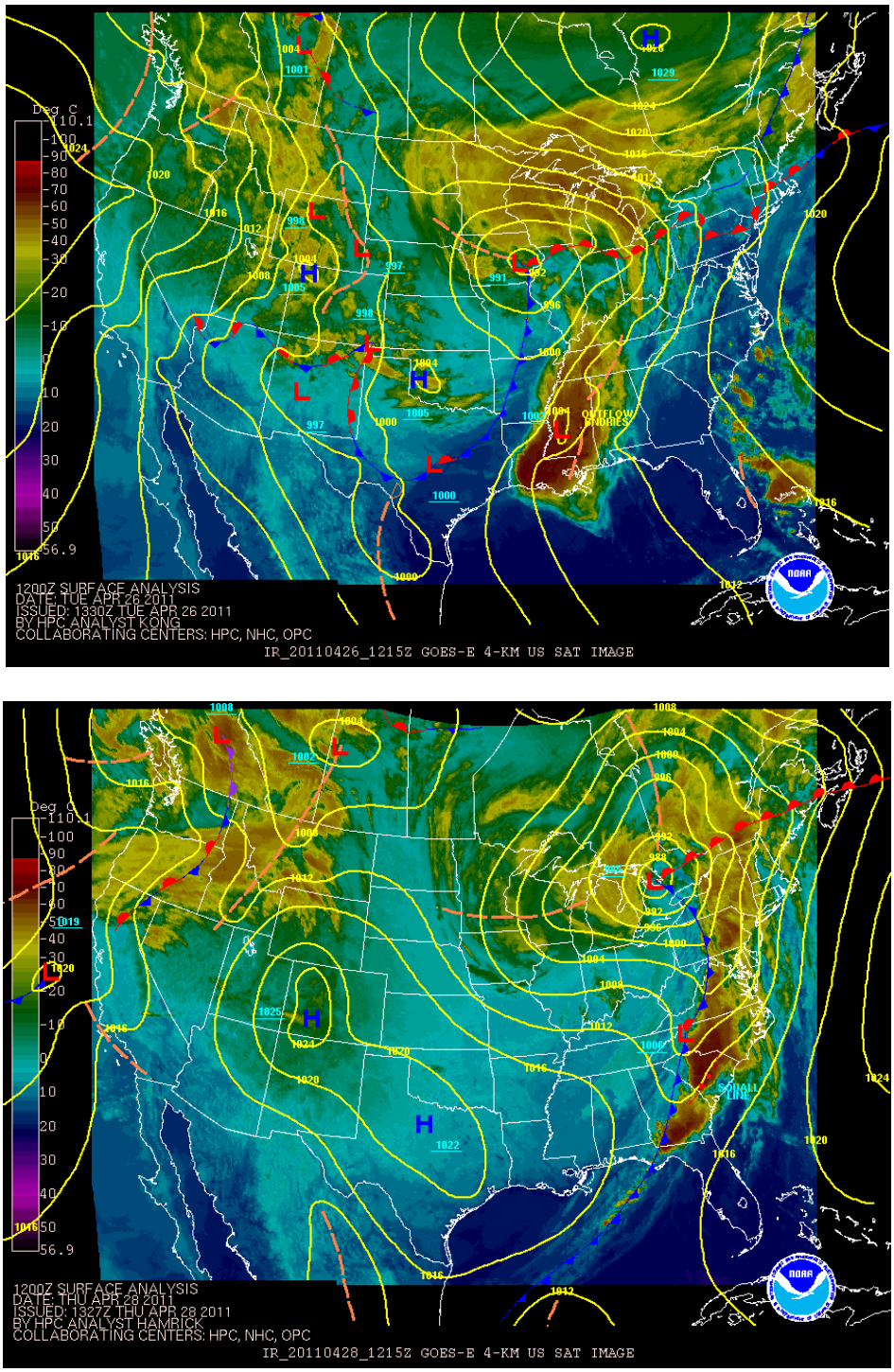


Figure 35. Series of satellite and surface analysis snapshots for 22-28 April 2011 to illustrate the atmospheric environment present during cluster match sampling.

Figure 36 shows the raw V850 for 15 Apr through 22 Apr, with the 24-hr forecast on the top, 72-hr in the middle, and 120-hr on the bottom row to demonstrate the an example of the performance of the algorithm from day-to-day. The columns are representative of the date for which the forecast is valid. The number in the top left indicates to which cluster the pattern is binned. The bottom number inside the image indicates a hit (1) or miss (0). The images are representative of only one ensemble member, however, the cluster outcomes are displayed underneath each image for every ensemble member. One can note that for the +24 FH the ensemble members resolve a fairly homogenous outcome. Over time, the spread of outcomes widens, which is comparable to the decrease in hits for +120 FH in Figure 36. Additionally, frontal progression can be visualized in this example. Starting with the first column (15 April) one can see pre-frontal signatures of positive wind on the right and negative on the left. Following throughout the series of columns, frontal passage, post-frontal, and another frontal passage is noticeable (Figure 36). Comparing column 1 to the satellite and analysis (Figure 37) it is clear that +24 FH is the best performing as +72 and +120 are too progressive with frontal movement. As the low pressure system matures and as more observations can be taken, the ensembles at all forecast hours correctly forecast the system. By day 4 (18 April), however, all ensemble members incorrectly forecast positive winds. This is most likely due to a quasi-stationary frontal system that drapes across the northern half of the study area.

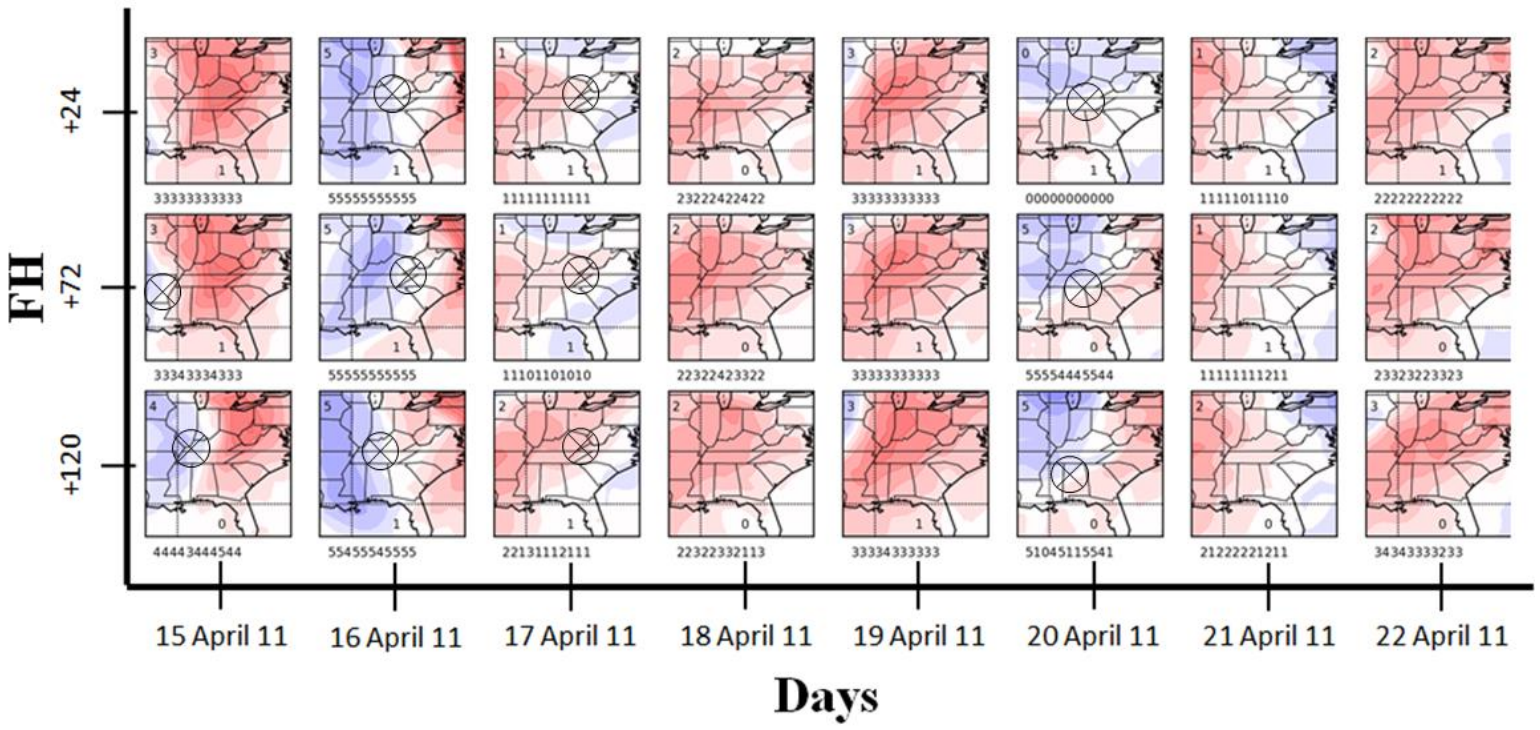
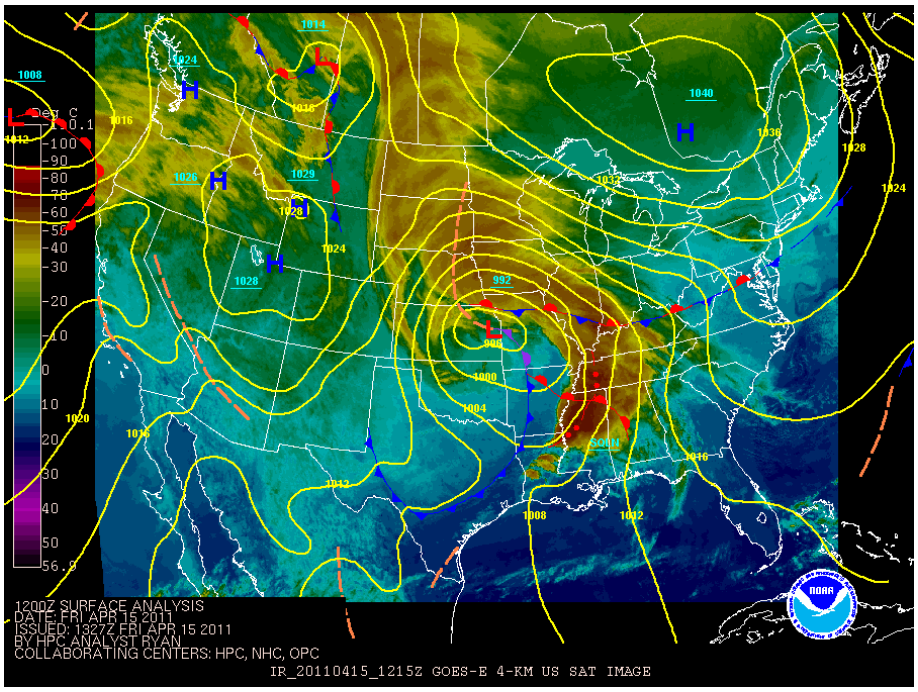
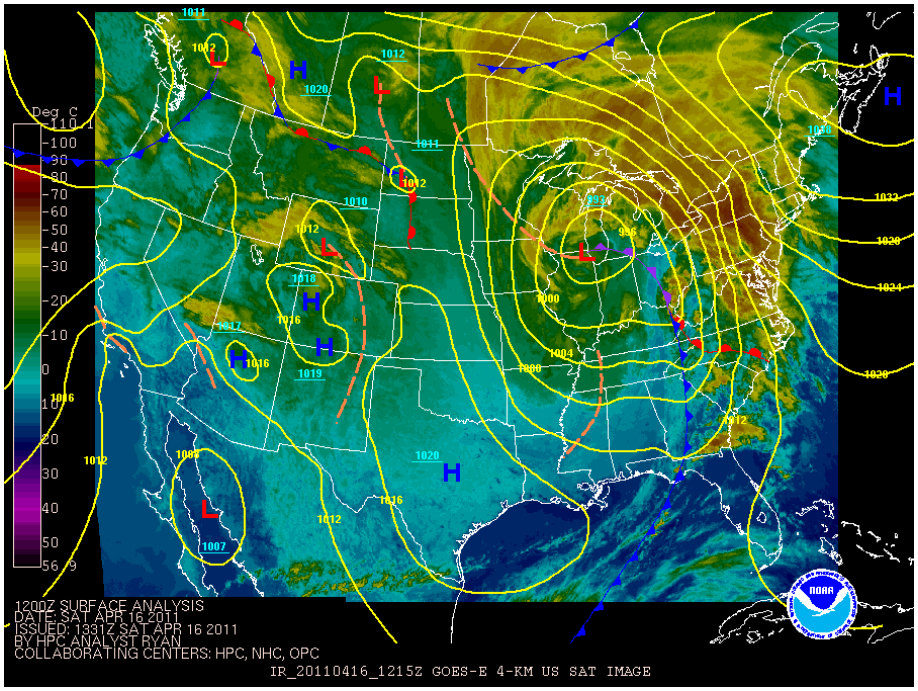


Figure 36. Raw binned clusters for one ensemble for April 15-22 2011.





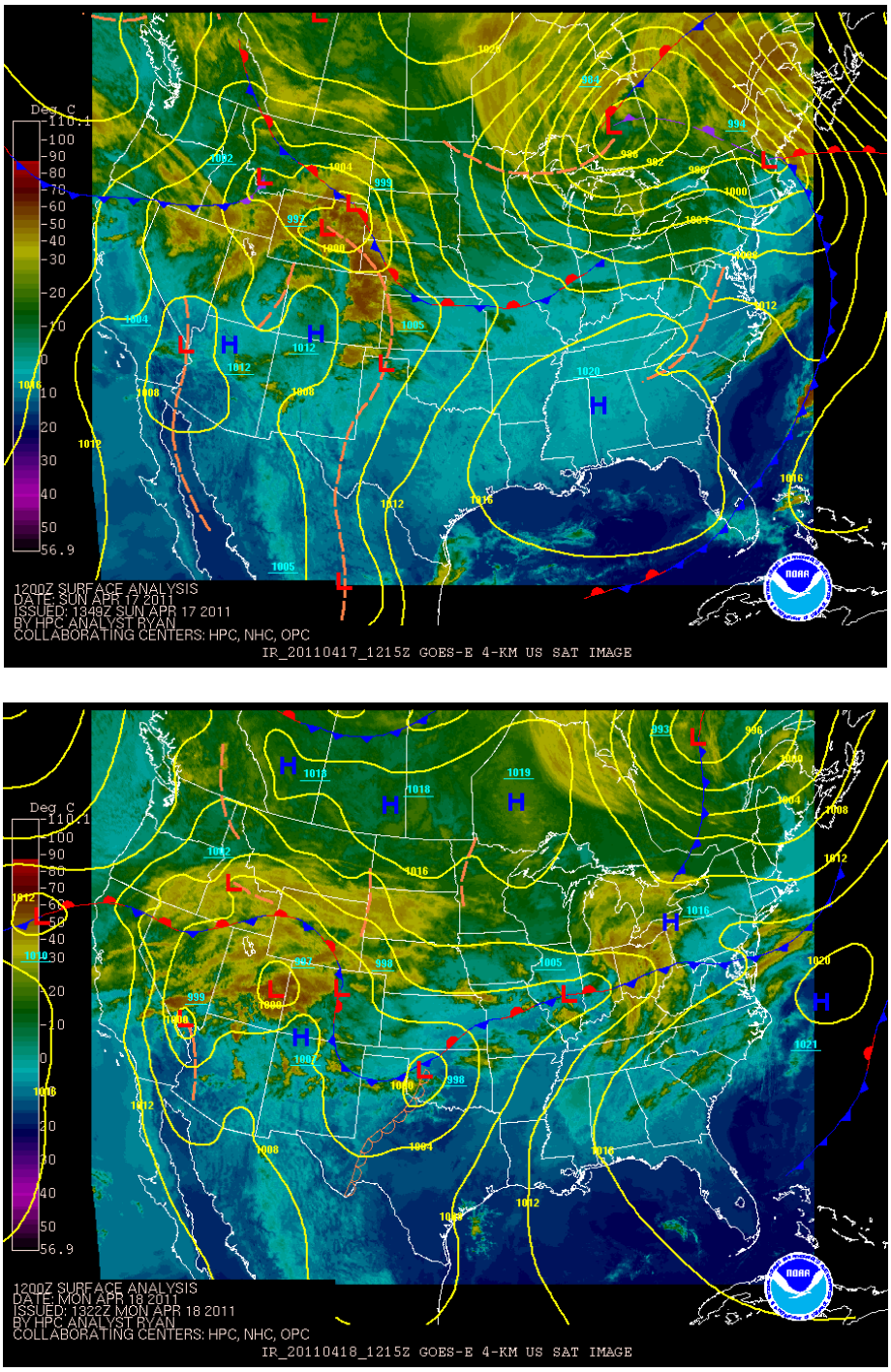


Figure 37. Series of satellite and analysis images for comparison with Figure 36.

Upon inspection of random samples from the case study, it is noted that the cluster number assigned to certain clusters can be suspect. In figure 38, one can notice that some of the clusters look alike, yet are categorized into different bins. Therefore, it is determined that a statistical analysis of the data must be performed.

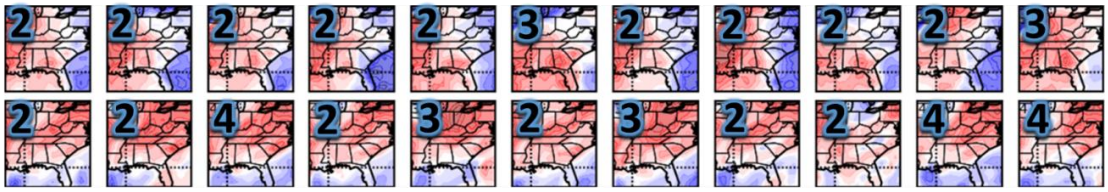


Figure 38. Results of an inspection of random samples of clusters.

This statistical analysis is for a sample area in the eastern United States and western Atlantic. In order to conduct the analysis a 6x6 contingency table is created for the 80 points within the test area. A Chi-squared test is run on these points in order to determine if the clusters are significant. The null hypothesis that the clusters are not significant is rejected with a 99% confidence.

A secondary statistical analysis is conducted. Although the Chi-squared test determines that the clusters are significant, it is also beneficial to understand how the significance changes between clusters. A Mann-Whitney U test is performed on the CLM versus each cluster. The results can be seen in Figure 39.

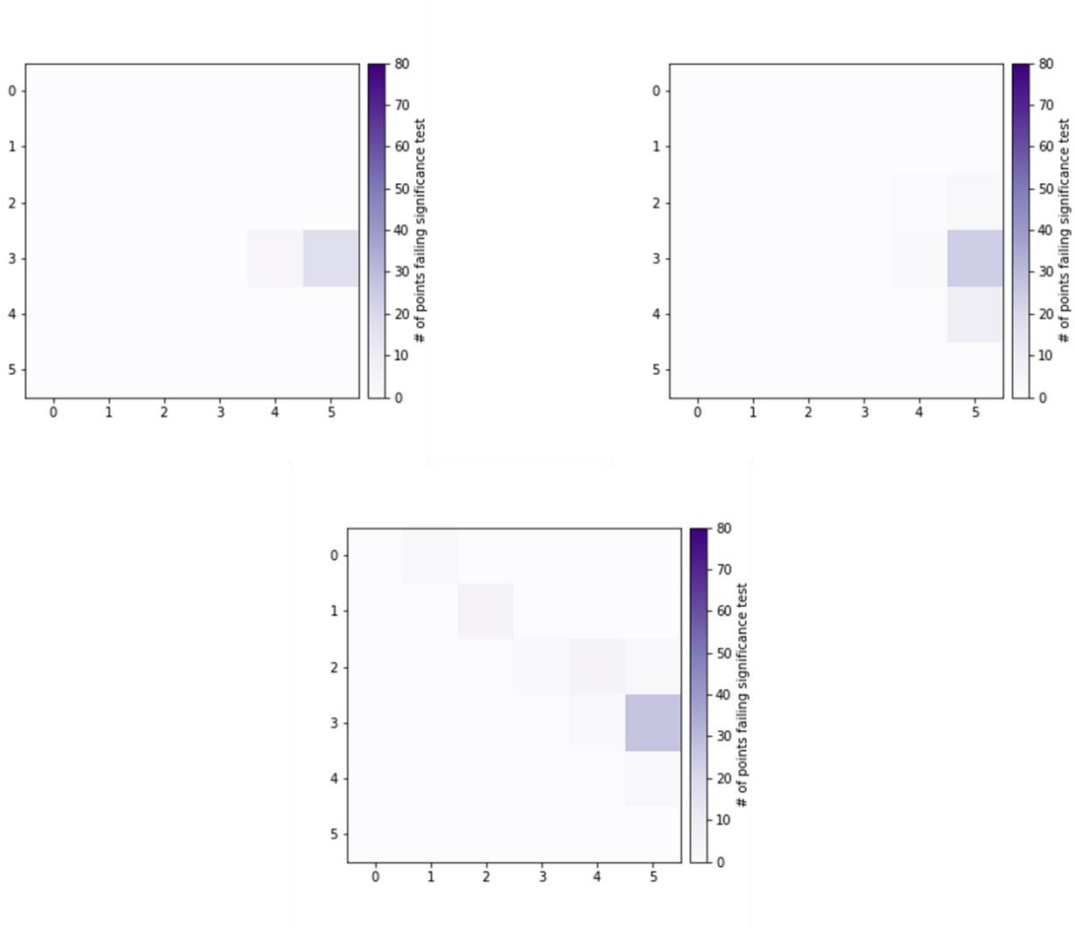


Figure 39 shows the results of the Mann-Whitney U test for 24 FH (top left), 72 FH (top right) and 120 FH (bottom).

The x and y axes indicate the cluster numbers for the 6 clusters. The color bar indicates the number of points that fail the significance test. The first observation is that the number of points which fail the test increase as FH increases. This is in line with the previous findings of CLM decrease with increased forecast hour. The most notable trends occur in the 120 FH. One can notice that most of the failures occur to the right by one. For example, when comparing clusters 0 to 1, 1 to 2, 2 to 3, 3 to 4, and 4 to 5, one can see that the most failures occur with these cases. This is because adjacent cluster to the right

and its partner cluster are the most similar to each other of any of the other pairings, and are expected to have the highest levels of mismatches. Another finding is the cluster 3 compared to 5 have the most failures. It is not immediately obvious as to why this may be the case. However, it might be due to a regional bias, since this statistical analysis is focused on only a portion of the United States.

V. Discussion and Conclusions

5.1 Summary of Results

Cluster analysis is employed on a data set consisting of V850 horizontal flow field variables in order to quantitatively depict various atmospheric regimes. Each cluster is then binned according to a uniquely made set of constraints. This binning mimics realistic atmospheric scenarios such as pre-frontal, frontal, and post-frontal patterns. Additionally, a forecast uncertainty is calculated for each bin for +24, +72and +120 FH. This enables end-users to place a quantitative uncertainty on different atmospheric regimes and FH's.

For the analysis, clusters 0 and 3 represent the purely negative and positive standardized V850 components and 2 and 5 represent the split positive/negative V850 components. For the results, clusters 0 and 3 have the highest CLM for all cases. Clusters 2 and 5 typically have 10-15% less CLM for all cases. The conclusion for this result is that forecasting non-homogeneous V850 wind fields will have less CLM as their very nature implies a transition from balance, which can be more difficult to predict. Additional reasons are discussed in section IV.

The 15 years of data is also partitioned and analyzed in seasons. The impetus for partitioning the data is to gain insight into seasonal influences. This is based off the common meteorological knowledge that transition seasons (fall/spring) have more pattern variability than summer and winter. Thus, it is thought that transition seasons will have a lower overall CLM when compared to the analysis for all 15 years. After analyzing it becomes clear that there are fluctuations in CLM within each season. CLM climbs during the fall and winter months, but then reduces overall throughout the spring and summer months. This result is slightly counter to the original expectation. This result also means that summer and winter are not equal in predictability within the algorithm described in this thesis. Further research may be needed in order to ascertain a reason for why this is the case.

Lastly, a case study is conducted for a 40 day period centered on April 15 2011 with the purpose of attaching a real-world meteorological example to an otherwise abstract topic. The results are promising in that they do show a relationship between CLM and frontal passage dynamics. The finding that these are dependent is important as this means that this data can be given to a non-naïve machine learning program which can learn the patterns and uncertainties for 15 years as opposed to just 40 days, but this will be covered in more depth in the future work section.

5.2 Limitations

One limitation of this research is that the K-means algorithm is a function of a user-defined input. One such input is the cluster initialization. The original choice of cluster initialization affects the final outcome, and when changed, will yield inconsistent

results. A work around for this issue, according to Steinley (2003), is that one should conduct thousands of random initializations until one is determined to have the smallest sum of squared error. Due to the time constraints, as well as the fact that the results are meteorologically understandable and accurate, this certain technique is not employed.

Another user-defined input that has a far greater effect on the final outcome is the choice of the number of clusters that will be found (k). According to Jain (2010) this is the most serious and critical limitation of K-means clustering. A clear objective formula for determining the correct amount of clusters is not available. The most widely used formula is the elbow method (Figure 39) in which the K-means algorithm is run for a range of k values and which for each result the sum of squared error is calculated and plotted. Then one locates the portion of the plot that looks like an elbow which is the area where the error decreases less rapidly and is considered the optimal k (Tibshirani et al.,2001). However, because of the conceptual and intuitive nature of atmospheric patterns, the values of k can be determined by running various values of k and determining which value is the best. Nonetheless, this is a source of error and is possibly a limiting factor to this research.

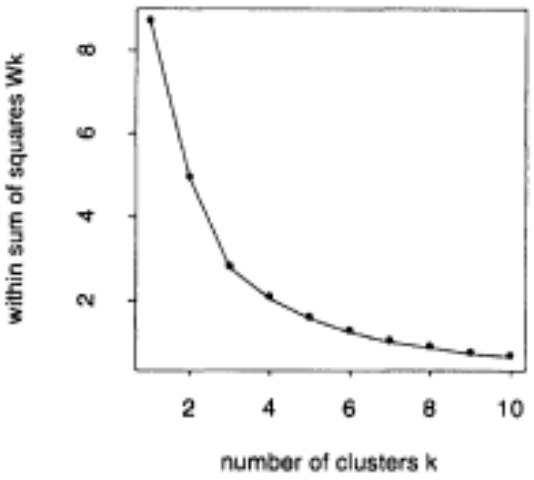


Figure 40. Example of the elbow method (Tibshirani et al.2001).

Another limitation that is experienced while using K-means for atmospheric patterns is that although the average of the 6 clusters used in this study look distinctively different, some members within the particular clusters are similar or indistinguishable from other clusters for some geographic regions, especially in the summer. An example of this was shown in the case study section. Although the clusters appeared indistinguishable, a Chi-squared significance test was applied which indicated a 99% confidence to reject the null hypothesis of the clusters not being statistically significant.

Additionally, the method to bin the clusters into specific atmospheric patterns can run into ambiguity which is further compounded by the original ambiguity from the K-means algorithm. An example from Jain (2010) can help conceptualize this issue. Referring to Figure 3 from section 2.5, one can note that the data members are not separated into 3 distinct groups. Atmospheric patterns mimic such behavior. Thus, it is expected that transitional patterns can fall into a gray area, akin to the transition zones

between clusters in the Jain (2010) example, and cannot be intuitively separated from the others except for the black and white classification from the clustering algorithm.

5.3 Future Work

One area of future work can include utilizing different ensembles. Utilizing an ensemble with more members than the GEFSR may show different results than gleaned in this research. More ensemble members will also provide more examples of clusters and could enable a more accurate choice of forecast clusters. Another way of increasing examples of clusters and potential accuracy is to utilize more years in the study.

This research focuses on V850 winds but studies on the plethora of other atmospheric variables can prove extremely beneficial. A study into the zonal wind field at 250 hPa could possibly show connections between CLM and jet strength. Drawing comparisons of 500 hPa relative humidity and CLM might shed light on mid-level moisture content and its effect on CLM.

Another option can be to use the regimes from this research and attach another variable such as moisture or a wind field at a different atmospheric level. For example, one might find that in cluster 0 the CLM is lowest on days when relative humidity is XX. It is also possible to use multiple atmospheric levels simultaneously to depict atmospheric regimes. For example, meridional wind can be analyzed at 850, 500, and 250 hPa concurrently and can be used to create a vertical wind profile which can then be analyzed by K-means clustering. It is important to note, however, that connections drawn can be purely coincidental, and must be objectively and exhaustively analyzed for meteorological merit.

An example application is a regional air quality monitoring office. This office performs both proactive and retrospective air quality monitoring in a mountainous region over a large area. For proactive monitoring, while most of the office's sensors are fixed in position, some of the key sensors are mobile but can be moved only on Monday's (due to resource limitations). The system described in this research gives the office additional information than is otherwise available from the operational ensemble NWP data. With forecasts of the mesoscale regimes over the forecast area, the confidence in the regime likelihood occurring, and the tendency that the model errs in forecasting the regime, the office can develop a more informed decision, and act more quickly on resource allocation.

The retrospective monitoring mission is very different. An example is the development of source term estimation (STE) from the given emplacement of sensors. The office has the ability to run retrospective numerical weather model runs at increasingly finer horizontal and vertical grid-spacing, but resources are very limited and cases requiring STE are always greater than the resources potentially available to run the numerical models are the finest grid-spacing. With the regime determination and confidence system developed here, along with pre-developed retrospective studies, the office can develop a set of rules for what horizontal, vertical, and temporal resolution is acceptable for a given atmospheric pattern. For example, an archive of 00 UTC initiated global NWP data is archived on-site for retrospective studies and for initiating finer grid-spacing limited area model runs in the air quality monitoring office's area of responsibility. Imagine two scenarios in which different patterns occurred over the area. Past studies indicate that each pattern is suitably quiescent enough to not require

additional numerical modeling at finer grid-spacing, but the second pattern tends to break down even at short forecast time scales. This second pattern may require higher grid-spacing modeling at the end of the retrospective analysis time period due to the lack of confidence in the short term forecast upon which the retrospective is based.

5.4 Conclusion

This research is foundational and there are many extensions and applications to which this research can be applied. Ultimately, the goal of determining the confidence of the forecast pattern or regime is an act of data reduction and cost savings. The data deluging the modern forecaster is immense and grows every year. Any effective way to reduce this data and deliver an equal or higher quality forecast product is necessary and worthwhile. Computational resources are limited. The framework researched here provides not only the forecast atmospheric pattern, but also provides a likelihood that the pattern is correctly forecast. Regime-based forecasting has long been used to provide forecasters with the knowledge of the high-impact weather associated with past atmospheric patterns. The system developed in this research adds an order of magnitude to the decision-makers toolbox.

Bibliography

Adenstedt, R. K. (1970). Weather regimes in stochastic meteorological models. *Quarterly of Applied Mathematics*, *28*(3), 343–353. doi:10.1090/qam/99787

Baur, F. (1951). Extended-Range Weather Forecasting. *Compendium of Meteorology*, 814–833. doi:10.1007/978-1-940033-70-9_66

Bock, H.-H. (2007). Clustering Methods: A History of K-means Algorithms. *Selected Contributions in Data Analysis and Classification Studies in Classification, Data Analysis, and Knowledge Organization*, 161–172. doi:10.1007/978-3-540-73560-1_15

Dalenius, T., & Gurney, M. (1951). The problem of optimum stratification. II. *Scandinavian Actuarial Journal*, *1951*(1-2), 133–148. doi:10.1080/03461238.1951.10432134

Dalenius, T. (1950). The Problem of Optimum Stratification. *Scandinavian Actuarial Journal*, *1950*(3-4), 203–213. doi:10.1080/03461238.1950.10432042

Diday, E., & Simon, J. C. (1976). Clustering Analysis. *Digital Pattern Recognition Communication and Cybernetics*, 47–94. doi:10.1007/978-3-642-96303-2_3

Feldstein, S. B. (2000). The Timescale, Power Spectra, and Climate Noise Properties of Teleconnection Patterns. *Journal of Climate*, *13*(24), 4430–4440. doi:10.1175/1520-0442(2000)013<4430:ttpsac>2.0.co;2

Ferranti, L., Corti, S., & Janousek, M. (2014). Flow-dependent verification of the ECMWF ensemble over the Euro-Atlantic sector. *Quarterly Journal of the Royal Meteorological Society*, *141*(688), 916–924. doi:10.1002/qj.2411

Franzke, C., Horenko, I., Majda, A. J., & Klein, R. (2009). Systematic Metastable Atmospheric Regime Identification in an AGCM. *Journal of the Atmospheric Sciences*, *66*(7), 1997–2012. doi:10.1175/2009jas2939.1

Hamill, T. M., Bates, G. T., Whitaker, J. S., Murray, D. R., Fiorino, M., Galarneau, T. J., et al.,(2013). NOAAs Second-Generation Global Medium-Range Ensemble Reforecast Dataset. *Bulletin of the American Meteorological Society*, *94*(10), 1553–1565. doi:10.1175/bams-d-12-00014.1

Huth, R. (2000). A circulation classification scheme applicable in GCM studies. *Theoretical and Applied Climatology*, *67*(1-2), 1–18. doi:10.1007/s007040070012

Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*, *31*(8), 651–666. doi:10.1016/j.patrec.2009.09.011

Lackmann, G. (2015). *Midlatitude synoptic meteorology: dynamics, analysis, and forecasting*. Boston, MA: American Meteorological Society.

Lorenz, E. N. (1963). Deterministic Nonperiodic Flow. *Journal of the Atmospheric Sciences*, *20*(2), 130–141. doi:10.1175/1520-0469(1963)020<0130:dnf>2.0.co;2

Michelangeli, P.-A., Vautard, R., & Legras, B. (1995). Weather Regimes: Recurrence and Quasi Stationarity. *Journal of the Atmospheric Sciences*, *52*(8), 1237–1256. doi:10.1175/1520-0469(1995)052<1237:wrraqs>2.0.co;2

Rex, D. F. (1950). Blocking Action in the Middle Troposphere and its Effect upon Regional Climate. *Tellus*, *2*(4), 275–301. doi:10.1111/j.2153-3490.1950.tb00339.x

Scher, S., & Messori, G. (2018). Predicting weather forecast uncertainty with machine learning. *Quarterly Journal of the Royal Meteorological Society*, *144*(717), 2830–2841. doi:10.1002/qj.3410

Steinley, D. (2003). Local Optima in K-means Clustering: What You Dont Know May Hurt You. *Psychological Methods*, *8*(3), 294–304. doi:10.1037/1082-989x.8.3.294

Tibshirani, R., Walther, G., & Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *63*(2), 411–423. doi:10.1111/1467-9868.00293

Vrac, M., & Yiou, P. (2010). Weather regimes designed for local precipitation modeling: Application to the Mediterranean basin. *Journal of Geophysical Research*, *115*(D12). doi:10.1029/2009jd012871

Warner, T. T. (2011). *Numerical weather and climate prediction*. Cambridge: Cambridge University Press.

Zarnani, A., Musilek, P., & Heckenbergerova, J. (2013). Clustering numerical weather forecasts to obtain statistical prediction intervals. *Meteorological Applications*, *21*(3), 605–618. doi:10.1002/met.1383

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From – To)*
22-03-2020	Master's Thesis	August 2018 – March 2020

TITLE AND SUBTITLE

Characterizing Regime-Based Flow Uncertainty

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Fioretti, John L., Captain, USAF

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S)
Air Force Institute of Technology
Graduate School of Engineering and Management (AFIT/ENP)
2950 Hobson Way, Building 640
WPAFB OH 45433-8865

8. PERFORMING ORGANIZATION REPORT NUMBER

AFIT-ENP-MS-20-M-093

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Technical Applications Center (spelled out)
1020 South Patrick Dr.
Patrick AFB FL 32935
321-854-8931 astrid.suarez.mullins@us.af.mil
ATTN: Dr. Astrid Suarez-Mullins

10. SPONSOR/MONITOR'S ACRONYM(S)

AFRL/RHIQ (example)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
DISTRUBTION STATEMENT A. APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

13. SUPPLEMENTARY NOTES
This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States.

14. ABSTRACT

The goal of this work is to develop a regime-based quantification of horizontal wind field uncertainty utilizing a global ensemble numerical weather prediction model. In this case, the Global Ensemble Forecast System Reforecast (GEFSR) data is utilized. The machine learning algorithm that is employed is the mini-batch K-means clustering algorithm. Horizontal flow fields are clustered and the forecast uncertainty in these flow fields is calculated for different vertical levels and for different forecast times for regions across the globe. This provides end-users quantified flow-based forecast uncertainty.

15. SUBJECT TERMS
Machine learning, K-means, clustering, atmospheric regime, classification, forecast uncertainty

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Lt Col Robert C. Tournay, AFIT/ENP
U	U	U	UU	19	19b. TELEPHONE NUMBER *(Include area code)* (937) 255-6565 x 4743 (Robert.tournay@afit.edu)

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39-18