Final Report

An Evaluation and Application of Multi-Model Ensembles in Operations for High Impact Weather over the Eastern U.S.

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
Collaborative Science, Technology, and Applied Research (CSTAR) Program

29 December 2017

PI: Dr. Brian A. Colle and Co-PI: Dr. Edmund Kar-Man Chang

School of Marine and Atmospheric Science

Stony Brook University

State University of New York

Telephone (631) 632-3174

FAX (631) 632-6251

Email: brian.colle@stonybrook.edu

1. Background

This goal of this project was to increase the use of ensembles in operations by (1) demonstrating multi-model ensemble performance for high impact weather during the cool season along the U.S. East coast, (2) calibrating ensemble gridded data, and (3) developing and training forecasters on new ensemble display tools to better understand ensemble predictions and the evolution of ensemble uncertainty. Partners for this project include several NWS offices and operational centers at National Centers for Environmental Prediction (the Weather Prediction Center, Environmental Prediction Center, Ocean Prediction Center, and Aviation Weather Center), and the Developmental Testbed Center.

2. Scientific Objectives and Accomplishments

i. *Fuzzy clustering for U.S. East coast winter storms*

a) Introduction to the fuzzy clustering approach

An efficient evaluation of ensemble models' performance for high impact weather events over U.S. East Coast region is very important to improve weather forecasting in this heavily populated region. For this project we developed a new ensemble clustering tool to help forecasters condense ensemble information, and evaluate ensemble models' performance by separating different scenarios and comparing them with an analysis. An operational goal is to apply this tool to ensemble forecasts, and improve the interpretation of multi-model output as well as provide guidance to potential model bias and outliers. The details of the approach are described in Zheng et al. (2017).

To quantify the variability of ensemble forecasts and recognize different scenarios among different operational models, a fuzzy cluster analysis is applied to group ensemble members with similar forecast scenarios. At the verification time, an EOF analysis is employed over ensemble MSLP anomalies over concern region. In general, as we have showed with our previous ensemble sensitivity efforts (Zheng et al. 2013), the leading EOF patterns represent the main characteristics of an ensemble forecast at the verification time. Given the first and second principal components (PCs) for each member of an ensemble set, Fuzzy clustering analysis can be performed to separate the members into groups. The detailed process is as follows:

a) To start the iterative cluster procedure, a predefined number of clusters or initial guess was randomly placed in the EOF PC1-PC2 phase space. Each ensemble member denoted by the pair of PCs is then assigned to the nearest group center.
b) New centers are computed by minimizing an objective function that represents the distance from each point to each new cluster center. Each point is examined again relative to the updated cluster centers. If no points can be reassigned because they are closer to another center, the iterations stop.
c) Each member is assigned a weight value that identifies their relative strength of membership to their cluster (Harr et al. 2008). For a point k, the weight associated with the *i*th cluster is defined as,

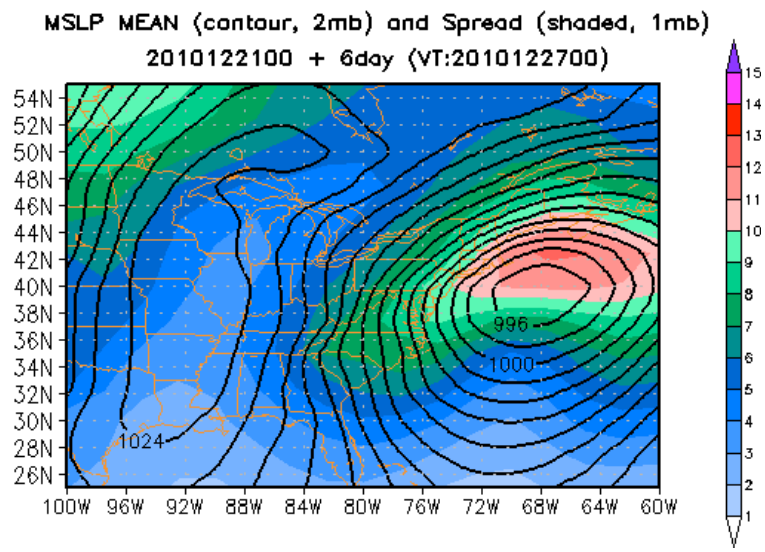
$$w_{i,k} = \frac{1}{\sum_{j=1}^{C} \left(\frac{d_{i,k}}{d_{j,k}}\right)^{2/(q-1)}}, \tag{1}$$

such that $d_{i,k}$ is the distance between point k and the centroid of cluster i, and $d_{j,k}$ is the distance between point k and the other cluster centers j. The fuzziness coefficient q determines the level of cluster fuzziness. A large q results in smaller memberships $w_{i,k}$ and hence, fuzzier clusters. In the limit q=1, the memberships converge to 0 or 1, which implies a crisp partitioning. In the absence of experimentation or domain knowledge, q is commonly set to 2. A total of C clusters can be calculated using equation (1) and the above procedure.

b) Case study using fuzzy clustering method

This section provides a forecast example to illustrate how to interpret the Fuzzy clustering result. The ensemble data is the 6-day ensemble forecast initialized at 0000 UTC Dec 21 2010 from ECMWF 50 members, CMC 20 members and NCEP 20 members. Mean sea level pressure (MSLP) is used for performing EOF analysis. One thing worth noting is that the analysis data (NCEP 6-hourly data) is also included as an extra member. Hence, the total ensemble member number is 90+1=91 members. EOF analysis is performed based on MSLP anomalies among these 91 members.

Figure 1 illustrates ensemble spread/mean of ensemble MSLP and the first two EOF patterns. For each EOF pattern, there is a set of corresponding principal components (PC) including 91 members. Ensemble mean suggests a cyclone off shore of coastal region. EOF1 pattern has a dipole around the ensemble mean cyclone, suggesting an east-northeastward shift of this cyclone (negative EOF1 pattern will correspond to an opposite shift). EOF2 pattern shows a monopole around cyclone center, suggesting a deeper cyclone (negative EOF2 will suggest a weaker cyclone). Therefore, EOF1 and EOF2 represent cyclone location and intensity uncertainty among the 91 members.



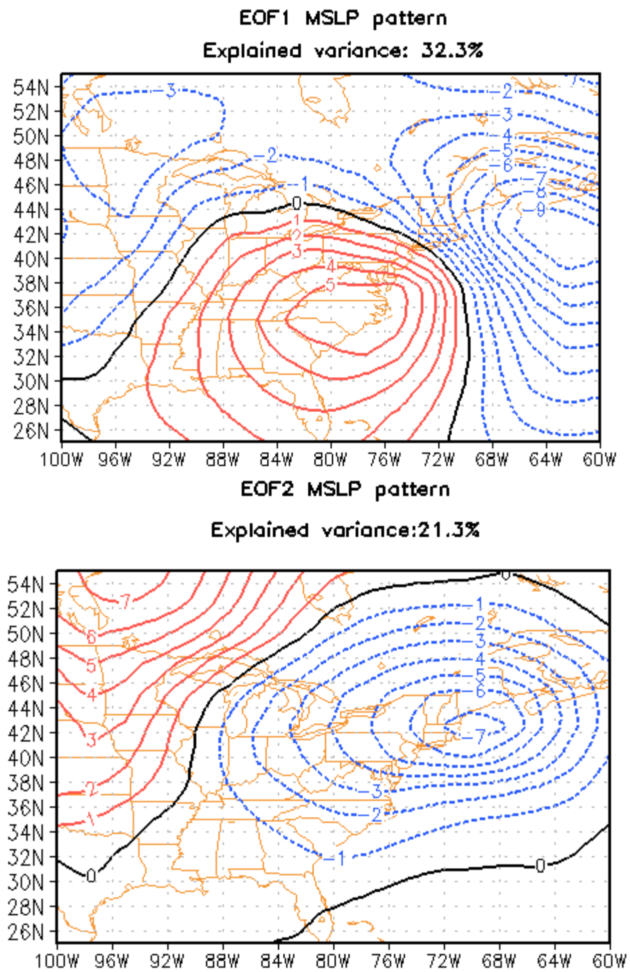


Figure 1 a) (top) Ensemble mean MSLP (contours, unit:[hPa]) and spread (shading, unit: [hPa]); b) (middle) EOF 1 MSLP pattern, unit: [hPa]; c) (bottom) EOF 2 MSLP pattern, unit: [hPa]. VT: 0000 UTC 27 December2010; IT: 0000 UTC 21 December 2010.

PC1 and PC2 corresponding to EOF1 and EOF2 pattern form the coordinate of the individual members on the phase space (Fig. 2). In this example, 90 ensemble members and 1 analysis are grouped into 4 clusters, which are represented by different markers. The cluster with analysis point is called group analysis or "Group ANA". The remaining three clusters are named by Group 1-3. Different marker colors (green, red, and blue) represent members from different forecast models (NCEP, CMC and ECMWF).

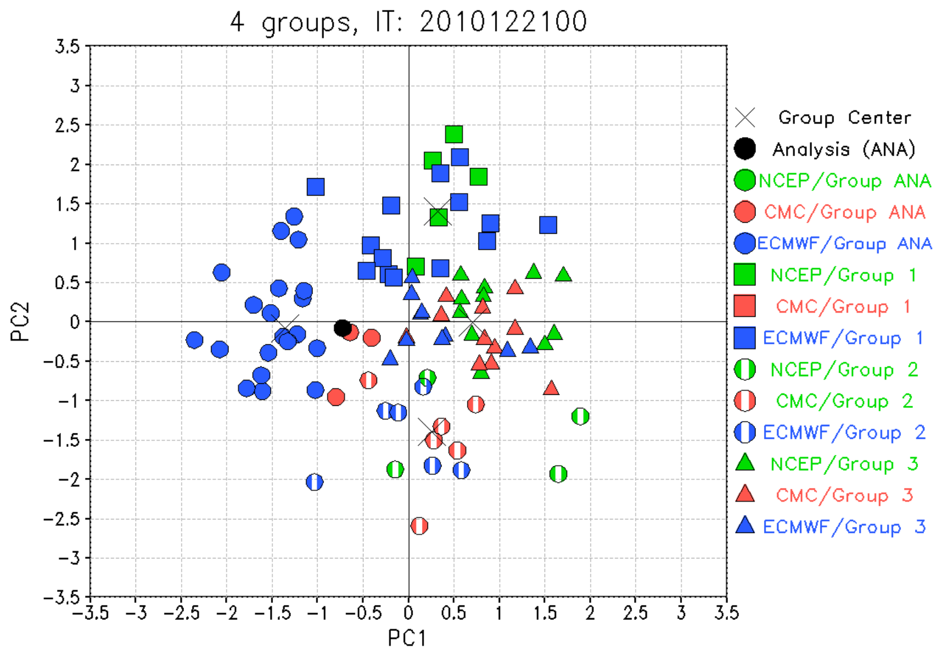
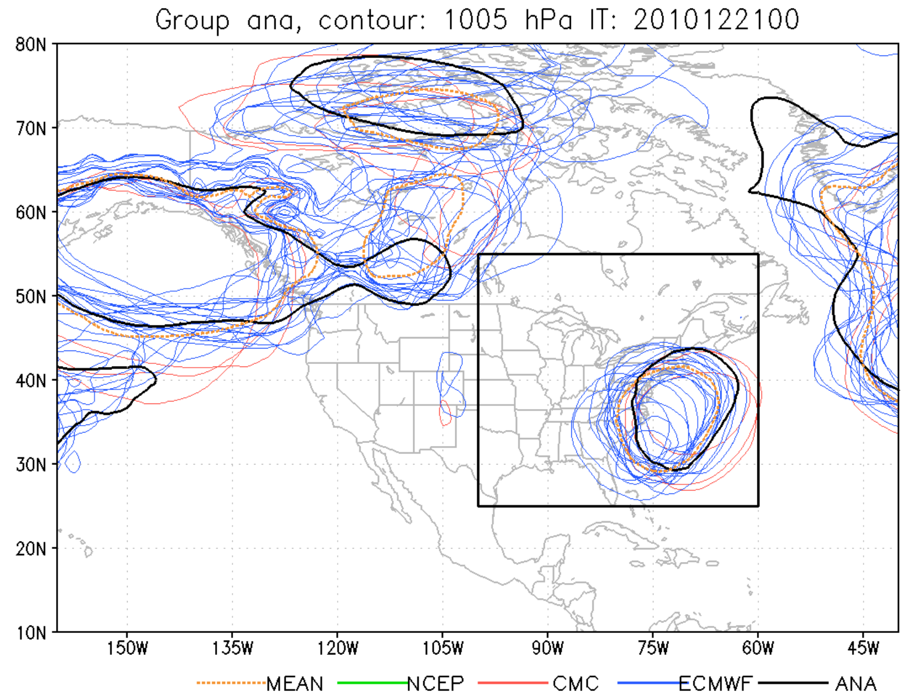
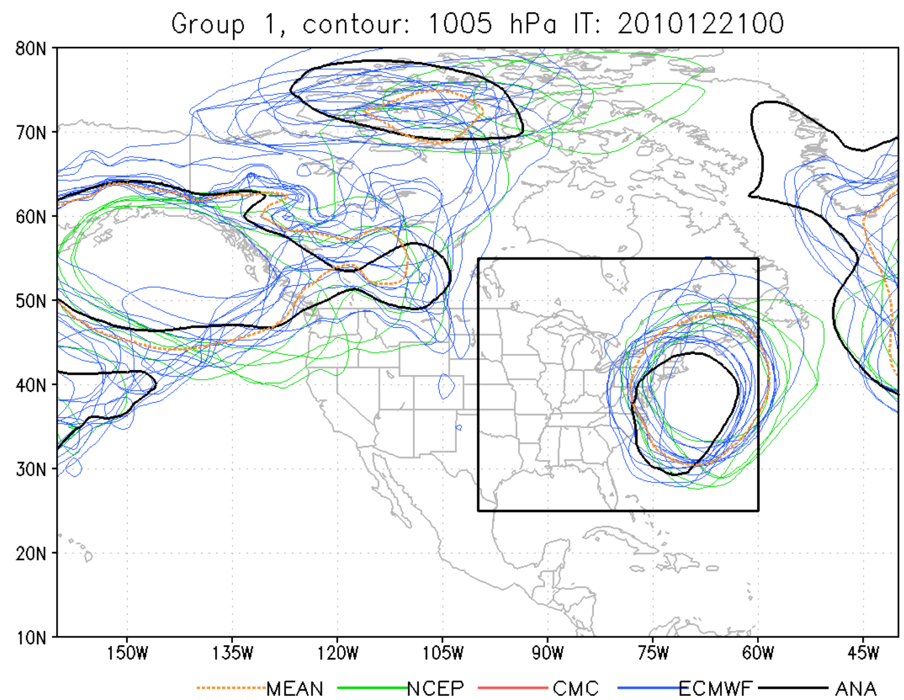


Figure 2. The first and second principal components for the four-cluster solution based on principal components from 90 ensemble members (colors) and 1 analysis member (black dot) initialized on 0000 UTC 21 December 2010 and valid at 0000 UTC 27 December 2010. Markers (filled circle, filled square, filled circle with a vertical bar, and filled triangle) represent clusters anal, 1, 2, and 3; black multiplication signs define the group centers for each cluster.

For this particular example, the analysis point is around (-0.7, 0), which belongs to negative PC1 pattern, or more onshore solution cluster. There are 20 ECMWF (blue dot) members, and 3 CMC (red dot) in this analysis group. None of the NCEP members are in this group. Therefore, for this 6-day forecast, 20/50 ECMWF and 3/20 CMC members predict a close-to-analysis scenario, while NCEP members miss this scenario. This helps the forecaster quickly assess the different ensemble scenarios. Looking at the NCEP members (green color for all marker), most of them tend to be over the right side (positive PC1) of the plot, suggesting most NCEP members forecast an offshore cyclone.

With the separation of 4 clusters as well as the corresponding ensemble members, we generated spaghetti plots for each cluster by picking up one MSLP contour line (Figure 3, 1005 hPa isobar). Among the four groups, analysis group (Group ana) mean is the one most close to analysis (top panel in Fig. 3). Most of the members (20) in this cluster are from ECMWF forecasts (thin blue contours), and the rest 3 members are from CMC (thin red contours). Group 1 mean (orange dashed contour) shows a deeper and more northeastward cyclone when comparing with analysis, and 14 ECMWF members and 5 CMC members form this group. Group 2 mean shows a more eastward cyclone; 6 ECMWF members, 6 CMC members and 4 NCEP members form this Group, hence all three models contribute to this group. Group 3 mean shows a deeper and more eastward-northeastward cyclone; 10 ECMWF members, 11 CMC members and 11 NCEP contribute to this group.


Group ana, contour: 1005 hPa IT: 2010122100
80N
70N
60N
50N
40N
30N
20N
10N
150W
135W
120W
105W
90W
75W
60W
45W
MEAN
NCEP
CMC
ECMWF
ANA


Group 1, contour: 1005 hPa IT: 2010122100
80N
70N
60N
50N
40N
30N
20N
10N
150W
135W
120W
105W
90W
75W
60W
45W
MEAN
NCEP
CMC
ECMWF
ANA

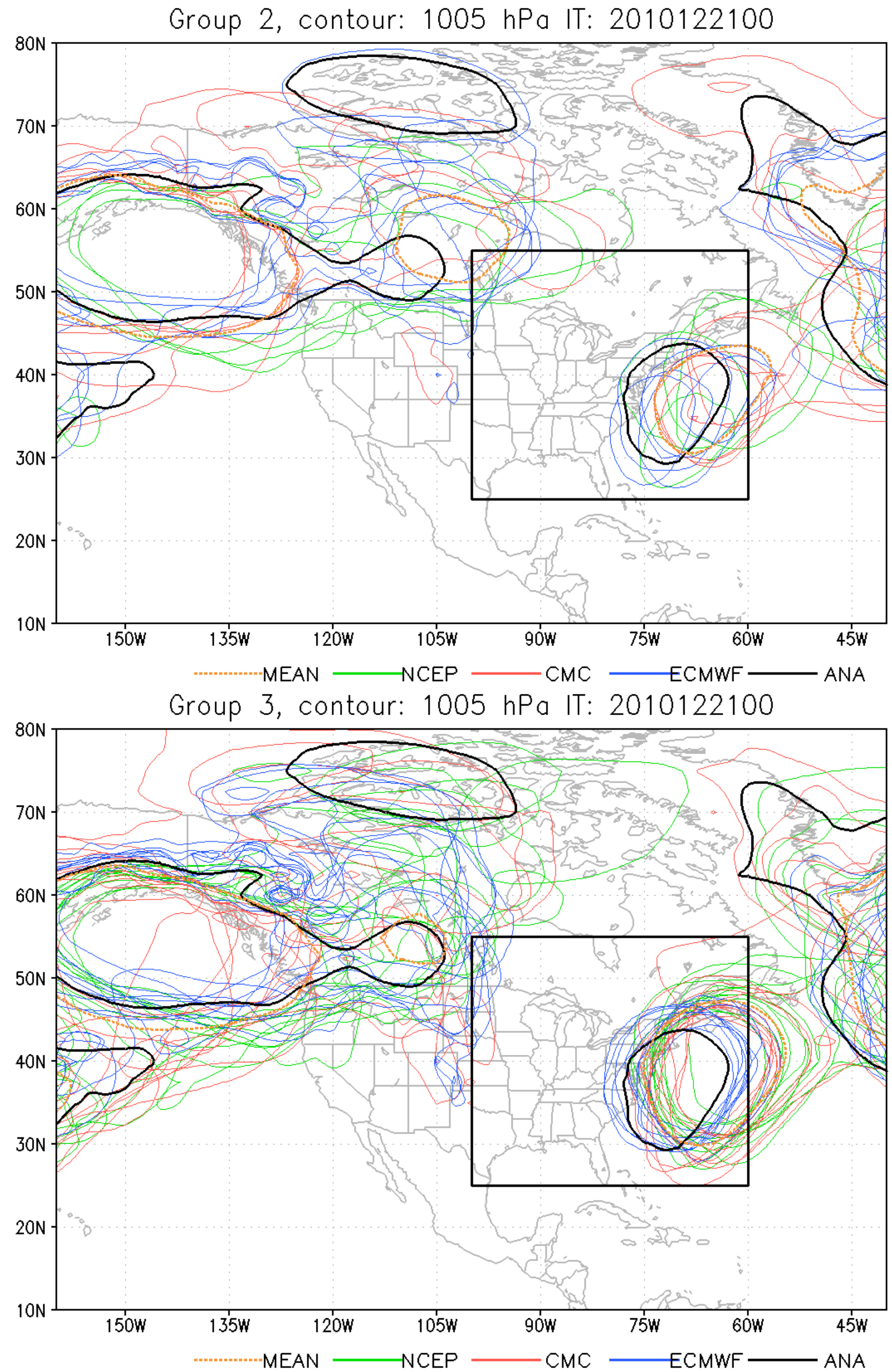


Figure 3. The spaghetti plot of 1005 hPa MSLP from ensemble members in group ANA, group 1, group 2 and group3. Dashed orange contour shows cluster mean contour; green, red and blue contours represent m`embers from NCEP, CMC and ECMWF. The black contours represent the analysis for 1005 hPa. The ensemble forecast is initialized on 0000 UTC 21 December 2010 and valid at 0000 UTC 27 December 2010.

c) Applying fuzzy clustering analysis to assess model performance

An EOF analysis is employed over ensemble MSLP anomalies over verification region at verification time (e.g., day 6) to characterize dominant forecast uncertainty patterns. Given the first two leading EOF principal components (PCs) for each member of an ensemble run, Fuzzy clustering analysis can be performed to separate the members into different groups (in this study, we use 5 groups). By doing this, similar forecast scenarios are often represented by the same groups.

We are using two verification regions in our analysis (Fig. 4). Given a 6-day ensemble forecast run, EOF PCs will be first calculated. Based on PCs, 5 groups will be partitioned using Fuzzy clustering analysis. The 90 multi-model members plus the ensemble meanare used when performing Fuzzy clustering grouping. Mean of each group will show the average scenario among that group, while spaghetti plots can show more details in the group. The analysis field is projected on the first two leading EOF patterns to get a pair of projection coefficients. The group including ensemble mean point is Group Ensemble (Group EM). The group with the center closest to analysis point is assigned to be analysis group. These two groups are not necessarily the same group in one case.

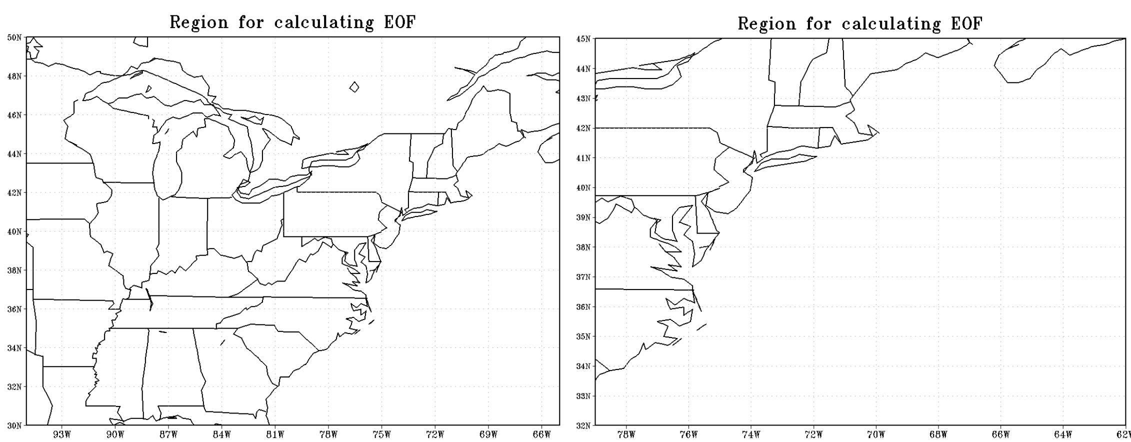


Figure 4. Two verification regions. Region 1 (left): North Central & Eastern U.S. Coordinates: 95W-65W, 30N-50N; Region 2 (right) Mid-Atlantic Coast & Western Atlantic. Coordinates: 79W-62W, 32N-45N

The case we are using to illustrate fuzzy clustering procedure is the February 2013 nor'easter impacting Northeastern U.S. and parts of Canada, which brought heavy snow and hurricane-force winds on 8-10 February 2013. The run we are using has initial time on 1200 UTC 3 February 2013 and valid at 1200 UTC 9 February 2013. Figure 2 shows the scatter plots using EOF PC1 and PC2 as well as the partitions of 5 groups. PC1 and PC2 corresponding to EOF1 and EOF2 pattern form the coordinate of the individual members on the phase space (See figure 5). In this example, 90 ensemble members and the ensemble mean are grouped into 5 clusters, which are represented by different markers. The cluster with ensemble mean (origin) is called group ensemble or "Group EM". The rest four clusters are named Group 1-4. The cluster with analysis point is group analysis, which can be any group among five groups (in this particular example, it is Group 3). Different marker colors (green, red, and blue) represent members from different forecast models (NCEP, CMC and ECMWF).

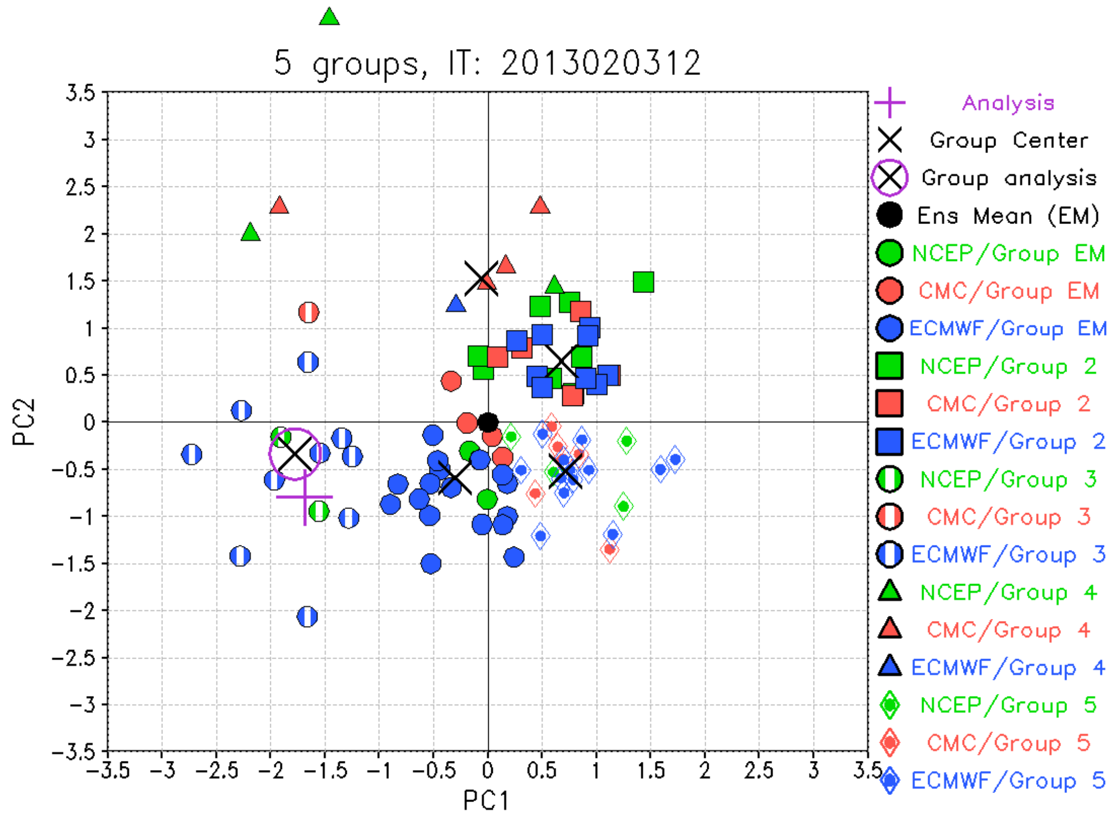


Figure 5. The first and second principal components for the five-cluster solution based on principal components from 90 ensemble members (colors) and the ensemble means (black dot) initialized on 1200 UTC 3 February 2013 and valid at 1200 UTC 9 February 2013. Markers (filled circle, filled square, filled circle with a vertical bar, filled triangle and open) represent clusters Ensemble (EM), 1, 2, 3 and 4; black multiplication signs define the group centers for each cluster. Purple plus sign shows the analysis position. Black multiplication sign with a purple circle represents the group with analysis.

For this particular example, the analysis point is around (-1.7, 0), which is closest to Group 3 and represents negative PC1 pattern, or a strong storm solution cluster (EOF patterns are not shown here). There are 12 ECMWF (blue dot with a vertical bar) members, 2 CMC (red dot with a vertical bar) and 1 NCEP (green dot with vertical bar) in this analysis group. Therefore, for this 6-day forecast, 12/50 ECMWF, 2/20 CMC and 1/20 members predict a close-to-analysis scenario. If you look at NCEP and CMC members (green and red color for all marker), most of them tend to be over the top-right side (positive PC1 and PC2) of the plot, suggesting most NCEP and CMC members forecast a weak and offshore cyclone. In this case, NCEP members have the most outliers; there is one NCEP member is even out of plotting chart (the green triangle on the top left of figure 5).

A summary of 5 groups using 1000 hPa MSLP contour line and 5460 m Z500 contour line is shown in Figure 6. All the group means missed the correct position of analyzed cyclone. However, Group 3 captured the comparable amplitude of the analyzed cyclone as seen in the MSLP plot. In the upper level (Right panel in Figure 6), the green dashed line (Group 3) is closest to the purple line (analysis).

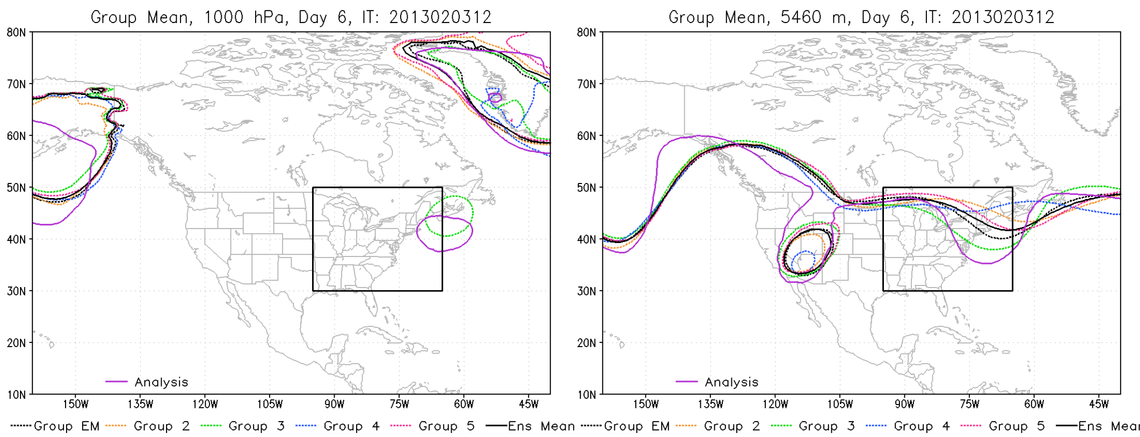


Figure 6. Summary of 5 group means using 1000 hPa line (left) and Z500 5460 m line (right). Purple solid line is the analysis; black, orange, green, blue and magenta dashed lines represent Group EM, 2, 3, 4 and 5, respectively. Black solid line shows the ensemble mean using all 90 members.

More details can be seen in spaghetti plots for each group. Figure 7 shows the spaghetti plots of Group 3. From Figure 5, most of the members (10) in this cluster are from ECMWF forecasts (thin blue contours), and the rest 3 members are from CMC (orange contours) and NCEP (green contour). Although Group 3 mean (purple dashed contour) shows a more northeastward cyclone when comparing with analysis, the amplitude is comparable between these two.

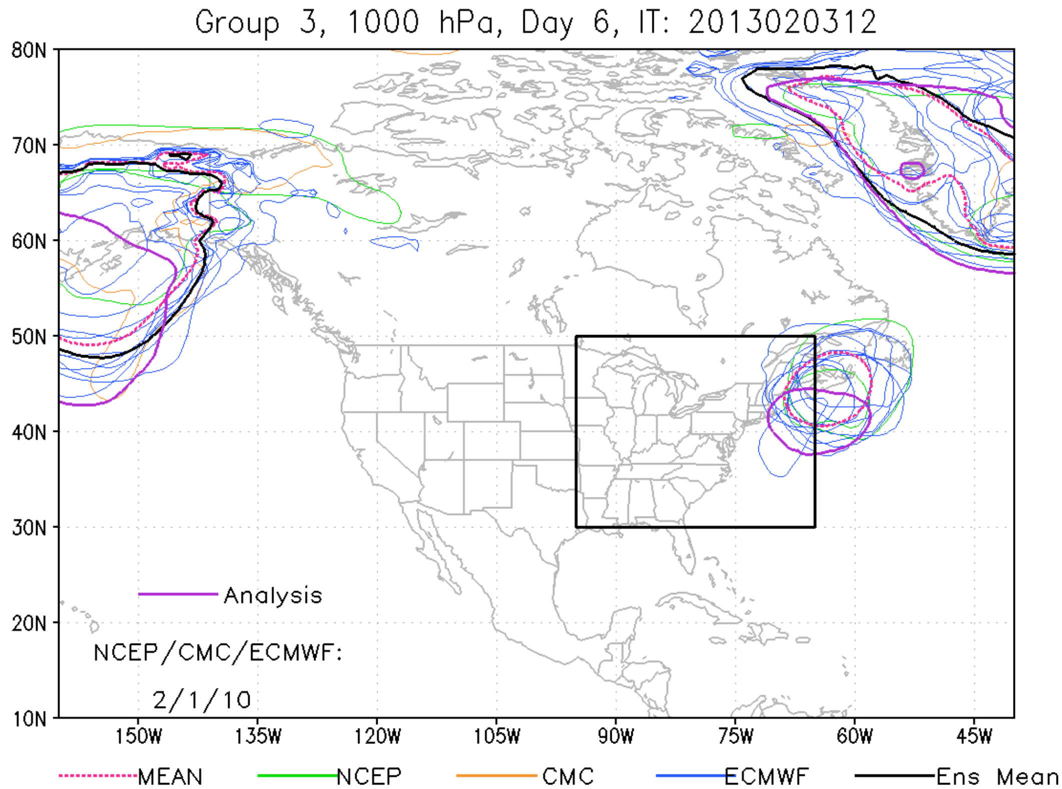


Figure 7. Spaghetti plot of MSLP from ensemble members in Group 3. Dashed magenta contour shows cluster mean contour; green, orange and blue contours represent m`embers from NCEP, CMC and ECMWF. The black solid line shows ensemble mean of 1000 hPa MSLP contour. The purple contour represents the analysis for 1000 hPa. The ensemble forecast is initialized on 1200 UTC 3 February 2013 and valid at 1200 UTC 9 February 2013.

Two verification regions are selected to calculate statistics of different models' performance in capturing the scenario closest to analysis for 145 HIW storm events over East Coast for a 6-day lead time. Figure 8 shows the percentage of ensemble members that are in the same cluster with analysis for each forecast model, a combination of NCEP and CMC models (NCEP+CMC), and a combination of three models (NCEP+CMC+ECMWF). For 6-day forecast the mean percentage of ensemble members out of all three models is 21.1% (21.2% for region 2). On average, 22.9% (23.7% for region 2) ECMWF (50 total members) are in the same group with analysis, which is the best among the three ensembles. An average of 19.9% (20.1% for region 2) NCEP models (total 20) and 18.0% (15.9% for region 2) CMC models (total 20) are in the same cluster with analysis.19.0% (18.0% for region 2) of a combined NCEP and CMC (total 40) members are in the analysis groups. Both regions' statistics suggest ECMWF has the best skill in terms of capturing the analysis scenarios, while the CMC model has the worst skill among three models. Regarding to missing model, NCEP has 14 cases that totally miss the analysis scenario; CMC has 16 cases missing the analysis cluster. In contrast, for all 145 cases, ECMWF model did not miss any of the analysis scenarios, again suggesting its stable performance in medium range forecast. One thing worth noting is that a combination of NCEP and CMC models only have 2 cases missing the analysis group, implicating a combination of these two models can significantly decrease the possibility of missing analysis scenario.

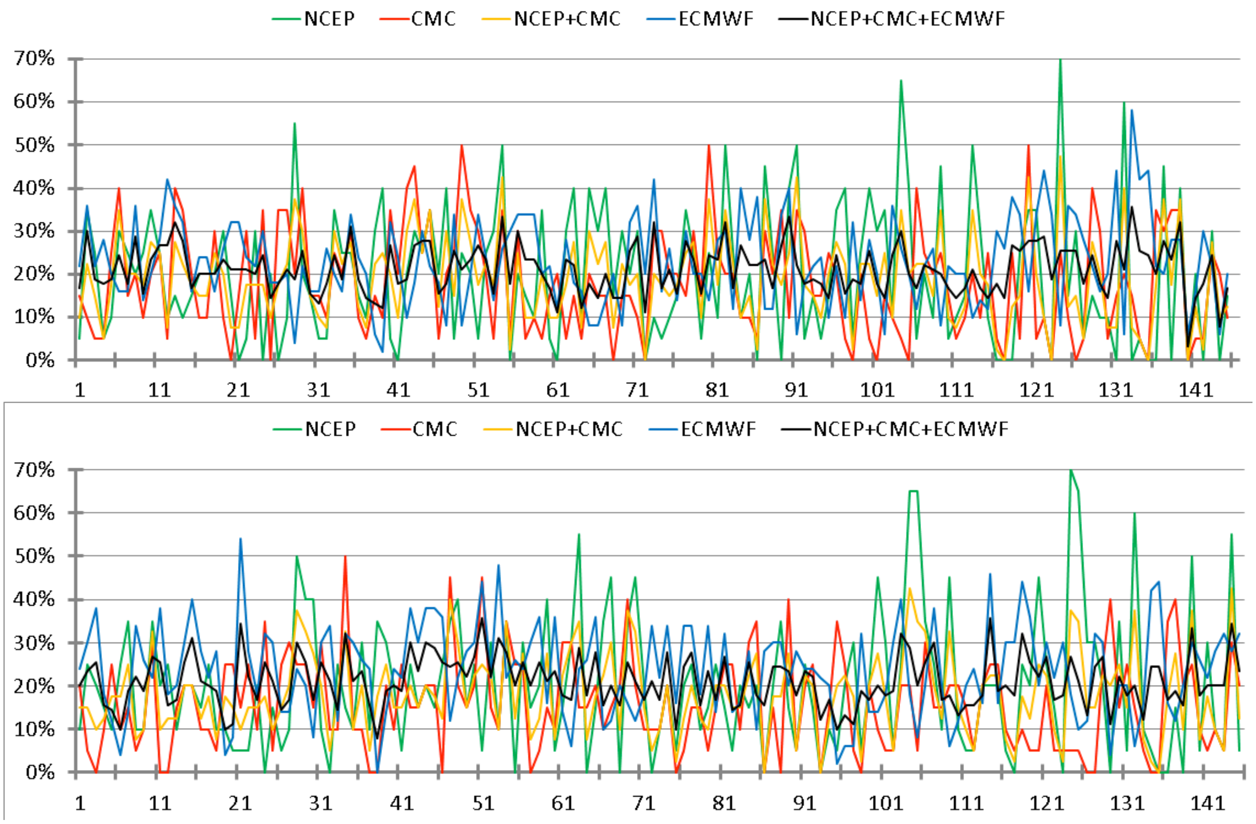


Figure 8 Percentage of ensemble members in each model in the same cluster with analysis for 145 HIW storm cases based on 6-day forecasts for region 1 (top) and (bottom). Green, red and blue line represent NCEP, CMC and ECMWF members, respectively; orange line represents NCEP+CMC combination members; black line represenst the percentage of members out of total 90 members (NCEP+CMC+ECMWF).

Table 1 Models' average percentage in capturing analysis groups for region 1 and 2.

	NCEP	CMC	NCEP+CN	ECMWI	NCEP+CMC+EC
Region	19.9%	18.0%	19.0%	22.9%	21.1%
Region	20.1%	15.9%	18.0%	23.7%	21.2%

As previously mentioned, the analysis is projected onto the EOF1 and EOF2 patterns and the corresponding projection coefficients are assigned to the closest group among 5 groups on the PC1-PC2 phase space. By employing this method to each historical cyclone case, we can find out the analysis group representing analysis scenario at verification time for each ensemble run. We repeated this calculation for all 126(116) HIW cyclone cases over regions 1 and 2. One thing worth noting is that the analysis could be out of the ensemble envelope in some cases, which will be investigated separately in the next section. Henceforth, there are 117(107) HIW cyclone cases in the following statistics excluding outside-of-envelope cases.

Figure 9 shows the percentage of ensemble members relative to their total ensemble members in the analysis group for NCEP, CMC and ECMWF over region 2 at a lead time of 3 days, 6 days, and 9 days respectively. For the 3-day forecast (Fig. 9a), on average over all cases there are 21.7% (20.7%) members out of 90 ensemble members in the analysis group for region 1 (2). Among the three models, ECMWF members show the highest percentage (24.9% of a total of 50 members) over region1 and significantly higher than that in NCEP, CMC and NCEP+CMC; the average percentage is 22.7% over region2, which is significantly higher than CMC and NCEP+CMC. This demonstrates that for the 3-day forecast, ECMWF members have the highest chance to include analysis scenarios. At the same time, although the error bars overlap with each other, the average percentages (18%/19.6% for regions 1/2) in NCEP (the total member is 20) members are higher than that in CMC members (16.4%/16.7% for region 1/2, the total member is also 20). Therefore, the CMC ensemble members have the lowest chance to include analysis scenarios.

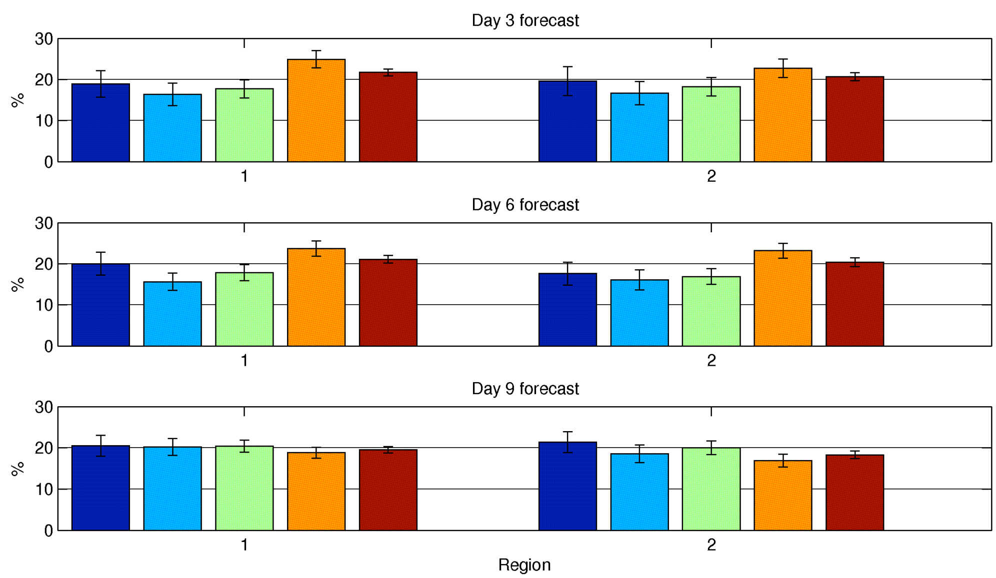


Figure 9. A summary of averaged percentage for 3-day, 6-day and 9-day forecast over both regions. From left to right: NCEP, CMC, NCEP+CMC, ECMWF, and analysis group.

For the 6-day forecast (Fig. 9b), the mean percentage of ensemble members out of the combined 90 members is 21.1% (20.4% for region 2). On average, 23.7% (23.2% for region 2) ECMWF models are in the analysis group, which is the best one among the three models. An average of 20.0% (17.6% for region 2) NCEP models and 15.6% (16.1% for region 2) CMC models are in the analysis group. 17.8% (16.9% for region 2) of the combined NCEP and CMC members are in the analysis groups. Both regions' statistics suggest that ECMWF has the best skill in terms of capturing the analysis scenarios, while the CMC model has the worst skill among three models.

For the 9-day forecast (Fig. 9c) the mean percentage of ensemble members out of all three models is 19.5% (18.3% for region 2). On average, 18.8% (16.9% for region 2) ECMWF members are in the analysis group, which becomes the worst one among the three models. An average of 20.5% (21.4% for region 2) NCEP models and 20.2% (18.6% for region 2) CMC models are in the analysis group. 20.4% (20.0% for region 2) of the combined NCEP and CMC members are in the analysis groups. Both regions' statistics suggest ECMWF has the most limited skill in terms of capturing the analysis scenarios for day 9 forecast, while the NCEP model becomes the best one among three models.

There are cases in which analysis is out of the multi-model ensemble envelope. Figure 10 is an example of comparing case with the analysis out of the forecast envelope. The outlier cases are defined by outside-of-envelope cases conditional on that the distance between the analysis and the closest member is larger than the average distances between the vertices and the closest members for all cases. There are 4 outlier cases for 3-day forecast, 9 for 6-day forecast, and 19 for 9-day forecast. For the statistics in the previous section, we have excluded 9 outliers based on 6-day forecasts considering that the outlier cases do not have the analysis group and those cases should not be included in the statistics.

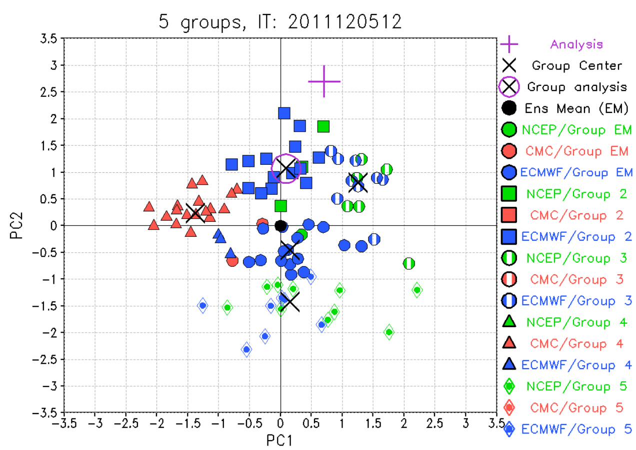


Figure. 10. Out-of-envelope case example from 3-day forecast.

In some cases, the analysis can be outside of one or two individual models among three models. In other words, one or two models can completely miss the analysis scenarios by contributing zero in the analysis group. For example, in day-3 forecast for region 2, there are 17(15.9%), 20(18.7%) and 2(1.9%) cases for NCEP, CMC and ECMWF models missing the analysis scenario out of a total of 107 cases. For day 3 forecast, NCEP has 24/17(20.5%/15.9%) cases, which totally miss the analysis scenario; CMC has 16/20(13.7%/18.7%) cases missing the analysis cluster among the 117/107 cases over region 1/2. On the other hand, ECMWF model only misses 2/2(1.7%/1.9%) of them. For day 6 forecast, NCEP has 12/12(10.3%/11.2%) cases that totally miss the analysis scenario; CMC has 12/15(10.3%/14.0%) cases missing the analysis cluster among the 117/107 cases over region 1/2. By contrast, ECMWF model only misses 1/0(0.9%/0.0%) of them. This missing cases statistics again suggest ECMWF models stable performance in medium range forecast. One thing worth noting is that a combination of NCEP and CMC models only have 1/3(0.8%/2.8%) and 2/4(1.7%/3.7%) cases missing the analysis group for 3- and 6-day forecasts, implicating a combination of these two models can significantly decrease the probability of missing analysis scenario.

There is case-to-case variability of each individual models performance in capturing the analysis scenarios. We show that models, especially NCEP models, can miss analysis scenarios in some cases. This leads us to explore another importance performance of different EPS: the error-spread relation. We are aiming to explore whether the ensemble models could provide sufficient forecast variability in simulating cyclone-related winter HIW cases over East Coast. The metrics we are utilizing to do cluster analysis are EOF PCs, with the standard deviation normalized to be 1 for each case. To be consistent, we choose this metric to study the spread-error relation. It is worth noting that the calculations include the aforementioned outside-of-envelope cases, which are denoted by large error-spread ratios.

Figure 11 shows the error-spread ratios for both regions and both EOF PCs based on 126(116) cyclone cases in three EPSs. One represents the perfect relation. The model is defined as "underdispersed" if the ratio is greater than 1; otherwise, it is defined as "over-dispersed" if the ratio is less than 1. For the 3-day forecast, the ECMWF model is the closest to 1 over both regions, suggesting that this model has perfect error-spread relation in 3-day forecast. Both NCEP and CMC models are under-dispersed, of which the former is under-dispersed more over both regions, suggesting the NCEP model may not have enough ensemble dispersion. On the other hand, the multi-model (NCEP+CMC+ECMWF) has the ratio less than 1, suggesting a combination of three models is over-dispersed for day 3 forecast. For the 6-day forecast, the ECMWF model is also a bit under-dispersed; however, it is still the closest one to perfect value among three models. NCEP and CMC models have comparable level of under-dispersion, with the latter being more underdispersed for PC1 over region 2. The multi-model is also under-dispersed in day 6, but it is closer to one than the rest, suggesting the benefit of combining different EPSs to provide more forecast variability for medium range forecast. For the 9-day forecast, all three models are severely underdispersed, of which the ECMWF model is even under-dispersed more in PC1 over region 2 than the rest. The multi-model is always closer to perfect value, again showing the increased forecast variability when combing models.

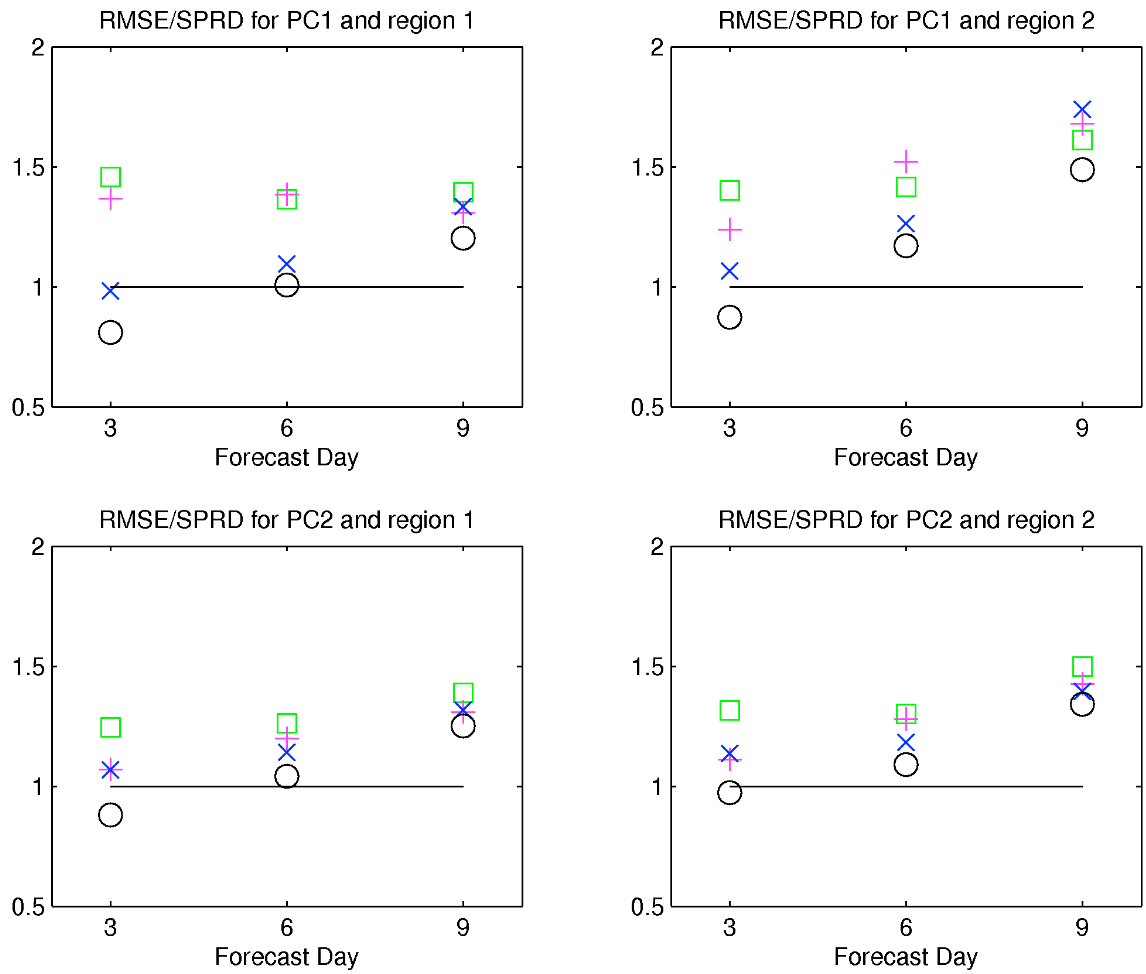


Figure 11. The ratio of RMSE to spread of PC1 and PC2 over region 1 and region 2. 1 represents the perfect error-spread relation. The green square, magenta plus, blue cross, and black circle represent NCEP, CMC, ECMWF and multimodel mean, respectively. The black line is the reference line 1.

Forecasters tend to look more at ensemble mean as guidance to the forecast scenarios especially in medium-range forecast. However, it is not clear whether the ensemble means have more chance to include the real scenario. In this study, we provide a new way to explore this problem. Since we can always find out the group including analysis in our historical cyclone cases, it is easy to check whether this group is the same with the ensemble mean group or Group EM, which can represent the ensemble mean scenario. Figure 12 shows the ratio of cases when the analysis group is the same with Group EM to the total cases equally divided by five groups. If the ratio is great than 1, the Group EM has more chance to include the analysis scenarios, otherwise, the Group EM doesn't include the analysis more often than any other group. For the day 3 forecast, the ratios for both regions are greater than 1, suggesting the ensemble mean group include the analysis scenario more often than the rest groups, which particularly hold true over region 1 with the ratio greater than 1.5. However, for the day 6 and day 9 forecasts, the ratios are either close to 1 (region 1) or even less than 1, indicating the analysis scenario is not more likely to be in the ensemble mean group than other groups. For day 9, the ensemble mean group even has less chance to include analysis. This result suggests that focusing too much on the ensemble mean could be misleading in many cases for medium to longer range forecasts. Based on the previous results, the ECMWF model seems to be the best one among three in terms of capturing the analysis scenarios. When counting the cases in which more than 40% of ECMWF members contributing to the analysis group (not shown), a majority of the ensemble members are from the ECMWF ensemble model for both day 3 and day 6 forecast. However, this is not the case for day 9 forecast. Only less than one quarter cases have the dominant members from ECMWF model.

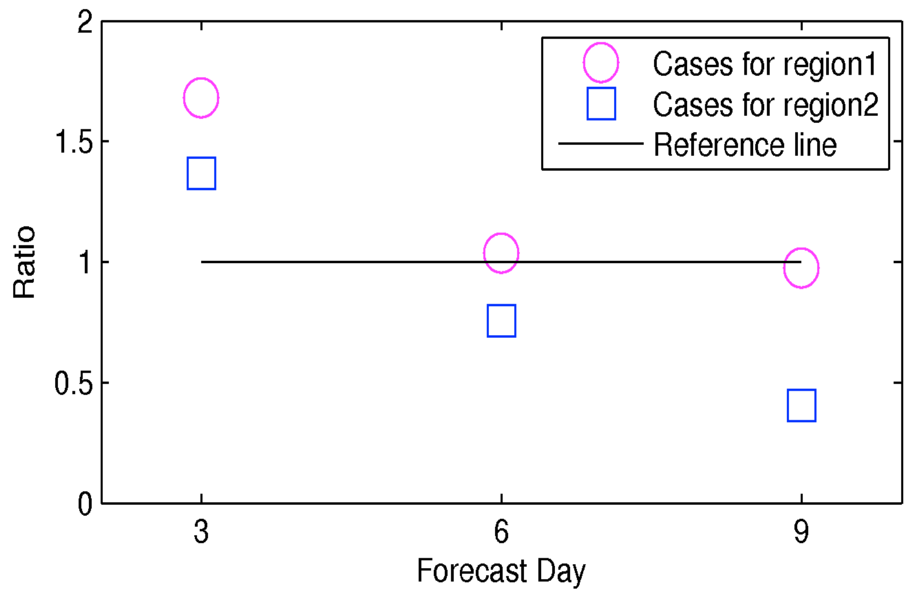


Figure 12. The ratio of the case numbers with the analysis scenario similar to the ensemble mean group to the 20% the total cases over both regions.

To further look at this problem, we have counted the cases when the analysis falls into the same quadrant including ECMWF means. For day3 and day6 forecasts, 34% (37%) and 30% (32%) cases over region1 (region 2), the analysis falls into the ECMWF means' quadrants. For day 9 forecast, these values drop to 21% (24%). On the other hand, in only 21% (24%) and 15% (16%) cases, the analysis falls into the same quadrant with NCEP+CMC means for day 3 and day 6 forecast. This value increases to 28% (32%) in day 9 forecast.

To sum up, the analysis does fall into the ECMWF means' quadrants more often than the NCEP+CMC means for day 3 and day 6 forecast. However, it falls into the latter's quadrants more often for day 9 forecast. One thing worth noting is that even in medium range forecasts, there are still around two thirds cases in which the analysis is outside the ECMWF means quadrants. For example, in the winter of 2014-2015, the ECMWF models tend to have problem in forecasting the correct position of cyclone in several HIW storm cases over East Coast, implicating that forecasters need to keep in mind that it could be misleading to rely on the

ECMWF models ensemble output too much when expecting a severe storm impacting East Coast.

ii. Validation of extratropical cyclones in operational ensembles

CSTAR graduate student (Nathan Korfe) validated extratropical cyclones within the global ensembles. The approach and results are described in detail in Korfe and Colle (2017). The Hodges surface cyclone-tracking scheme is used to construct cyclone tracks using 6-hourly MSLP from the Observing System Research and Predictability Experiment (THORPEX) Interactive Grand Global Ensemble (TIGGE), which includes the 50-member European Centre for Medium-Range Weather Forecasts (ECMWF), the 20-member Global Ensemble Forecast System (GEFS), and the 20-member Canadian Meteorological Centre (CMC). The ECMWF ERA-Interim Re-Analysis MSLP data has also been tracked to verify cyclone properties for the October to March cool seasons from 2007-2015. It is important to note that each forecast and analysis cyclone is created with a unique storm identification number (ID) that identifies all the points in the cyclone lifetime.

Some of the questions being addressed were the following:

1. How well do the operational ensembles (NCEP, CMC, ECMWF) predict cool season extratropical cyclones with lead times from days 1-6?
2. What are the intensity biases in the ensembles? Where are the largest intensity errors located within domain?
3. What flow patterns are associated with the high error cyclone cases in the ensemble?

Figure 13 shows a sample of the validation results by illustrating the cyclone displacement and intensity mean absolute errors for the CMC, NCEP, NCEP, and EC ensembles as a function of forecast lead time. Error bars are calculated using a bootstrap resampling approach (over 1000 random iterations of the data). For cyclone displacement, the CMC mean ensemble has the largest track errors, while the NCEP and EC ensembles are similar from hours 0 to 72. After hour 72, EC has the smallest displacement errors. The results are similar for cyclone central pressure errors. It is interesting that the NCEP control member of the ensemble has similar errors to the mean of all members, which suggests that there are cases in which ensemble mean has difficulties with the forecast as well. Figure 14 illustrates that the relatively deep cyclones are underpredicted after day-2, especially in the CMC and NCEP ensembles.

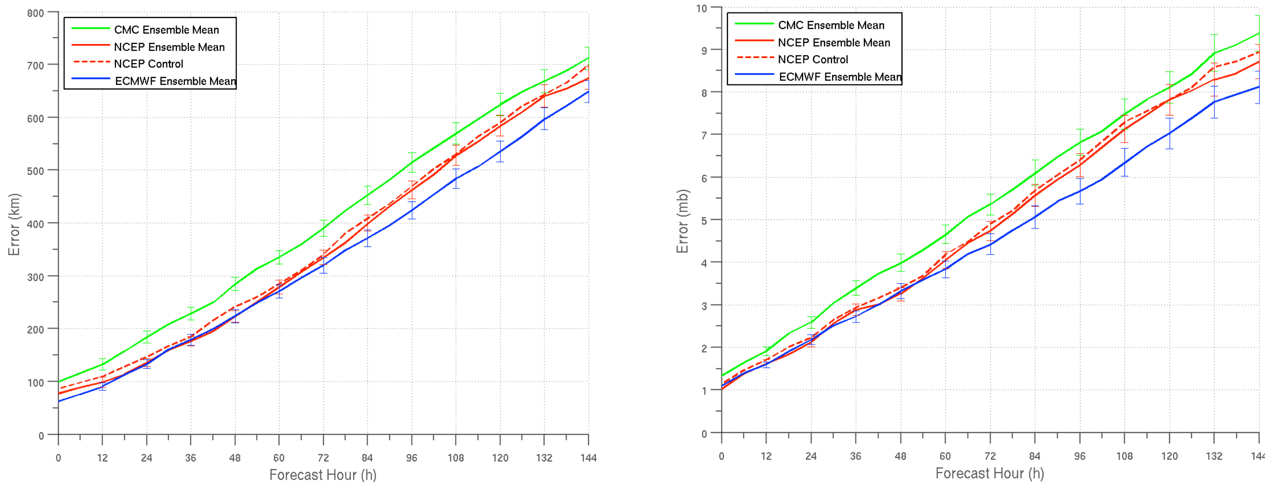


Figure 13. Cyclone displacement (left) and central pressure (right) mean absolute errors for the day 4-6 ensemble mean forecasts for the East Coast and western Atlantic for the ECMWF, NCEP, and CMC for lead times from hour 0 to 144.

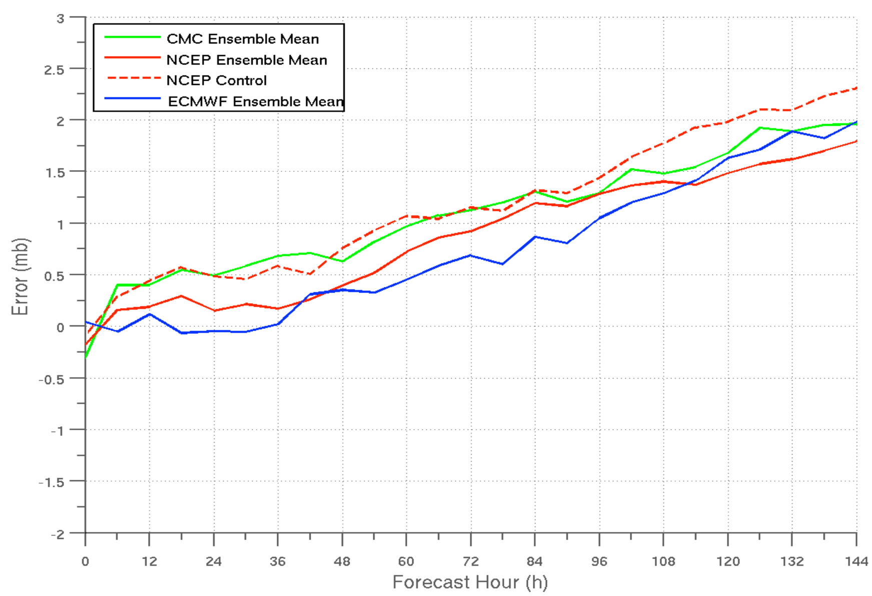


Figure 14. Cyclone central pressure mean errors for the day 4-6 ensemble mean forecasts for relatively deep observed and simulated cyclones (> 1.5 standard deviation) for the East Coast and western Atlantic for the ECMWF, NCEP, and CMC.

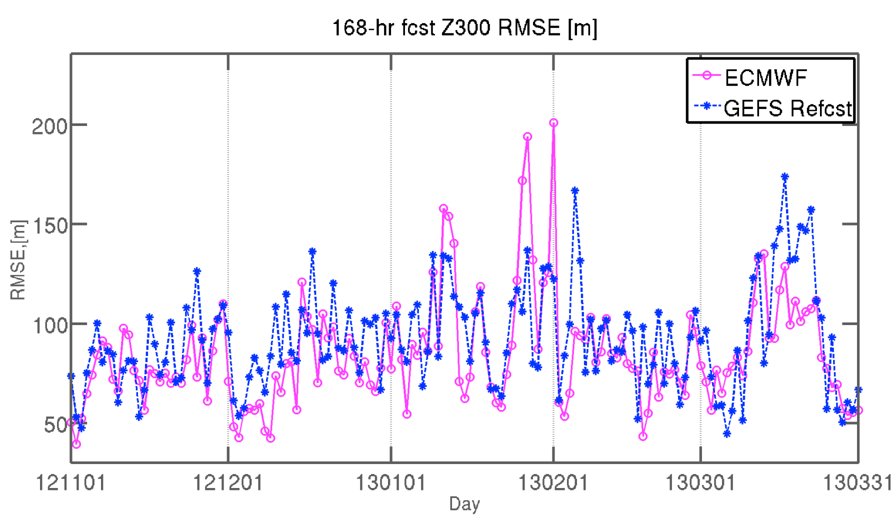
iii. Event-based verification of high impact weather – Rossby Wave Packets

Rossby wave packets (RWPs) have been linked to high impact weather, so it is important to understand how well ensembles can predict RWPs. This builds on our previous CSTAR that focused more on deterministic predictions of these RWPs.

The ensemble data used is the GEFS (NCEP Global Ensemble Forecast System) reforecast v2 data, which is similar to the GEFS operational configuration on February 2012. This set of data is generated once daily at 0000 UTC, producing an 11-member forecast, 1 control + 10 perturbed. The variables used in this work are geopotential height (Z300), zonal wind (U300), and meridional wind (V300) at 300hPa at 1°x1° resolution. The time period is from 1 Nov 1985 to 31 Mar 2013, which only includes cool seasons (NDJFM). We have mainly used day 1 to day 7 control forecasts as well as ensemble forecasts, which will be extended to longer lead times in the future.

a) <u>Verification of ensemble data quality in reforecast data over U.S. East coast and western Atlantic</u>

Since GEFS reforecast data only have 11 members, we also used ECMWF ensemble, which has 50 members and are often considered to be the best ensemble data, to verify the quality of the reforecast ensembles. We have compared RMSE of ensemble mean of Z300 in 7-day forecasts, ensemble spread, and anomaly correlation (ACC) in both data sets for five cool seasons from 1 Nov 2008 to Mar 2012 (Fig. 15). The correlations between two data sets for these three quantities are 0.609, 0.608, and 0.659. On average, the RMSE of ensemble mean Z300 (88.9 m) in ECMWF data is around 9% lower than GEFS reforecast (97.1 m); the ACC in ECMWF (0.785) is also slightly higher than that in reforecast data (0.740). The ensemble spread of Z300 in ECMWF (78.2 m) is around 11% higher than that in reforecast data (70.5 m). Figure 15 shows an example of RMSE, ACC, and ensemble spread for the 2012-2013 cool season. Although those two datasets are not perfectly correlated, the maxima and minima in general match each other. Since we are interested in extreme cases with very poor predictability and good predictability, we believe the GEFS reforecast data meets our needs.



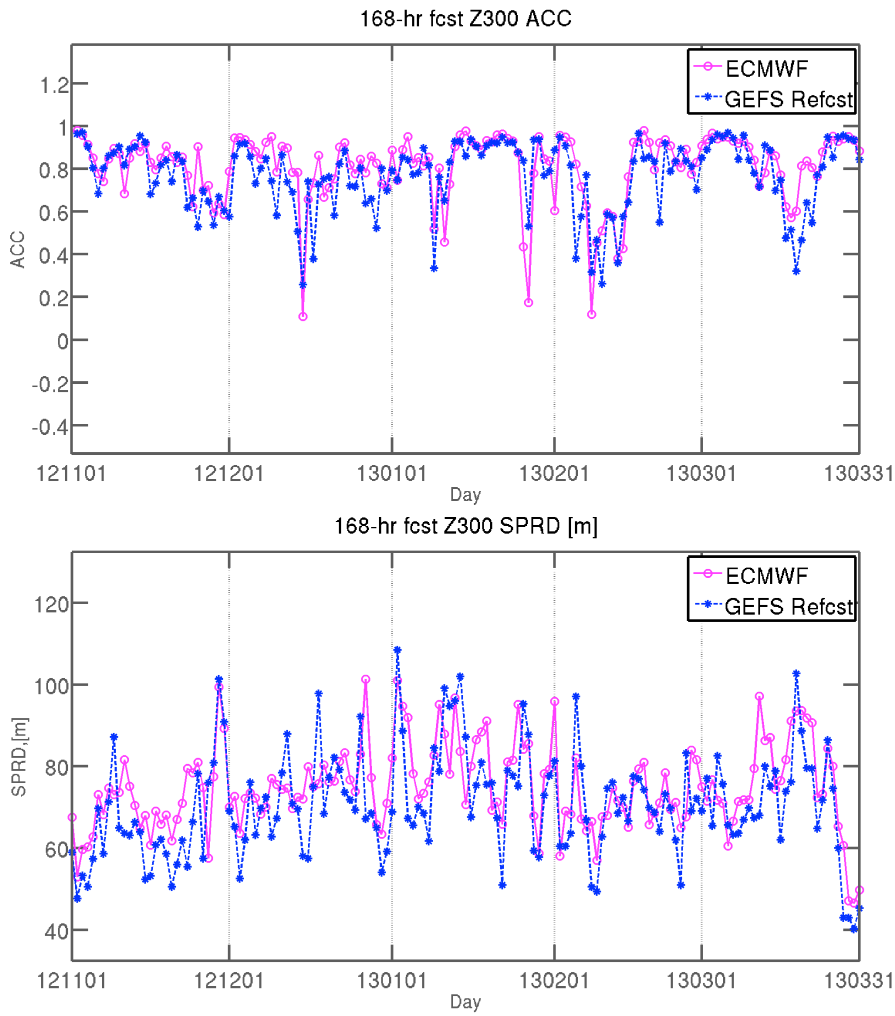


Figure 15. RMSE (top) and ACC (middle) of ensemble mean of Z300; ensemble spread (bottom) of Z300 from Nov 1st 2012 to Mar 31st 2013. All these three quantities are calculated over the region 22°N-62°N, and 100°W to 45°W (Verification region).

Empirical Orthogonal Functions (EOFs) have been employed in a variety of meteorological analyses due to its capability to reduce the dimensionality and to efficiently describe the coherent variability in large datasets. When applying this method to the datasets, which contain the same variable at different times, a modified form EEOF analysis (Weare and Nasstrom 1982) is often referenced. In an extended EOF analysis, successive patterns of the same variable within one EEOF are interpreted as the propagation or evolution in the time.

We have applied EEOF method to find dominant variance patterns of RWP amplitude anomaly during initial times and verification times corresponding to large forecast error cases. By applying this method, different types of wave packets are determined, which propagate along different paths. Based on the top and bottom principal coefficients (PCs) of each EEOF pattern, we have also composited the corresponding forecast errors and mean flow regimes.

We have also applied composite analysis to the wave packet characteristics of large error cases and large spread cases. The RWPs are represented by their amplitude, which is calculated using

Zimin's method (Zimin 2006) based on U300 and V300. In one example to show multi-model ensemble, we have used fuzzy clustering method (Bezdek 1981) to group ensemble models based on their leading two PCs within their forecasts.

b) <u>Different types of Rossby wave packets in large forecast error cases and the corresponding flow regimes</u>

Based on the time series of RMSE of 7-day Z300 control forecast (Fig. 16) as well as the anomaly correlation (not shown), we have selected 203 large error cases and 213 good forecast cases. We have also used RMSE and ACC of 7-day ensemble mean of Z300, the large error cases matched very well. The correlation between RMSE time series of control forecast and ensemble mean is 0.873.

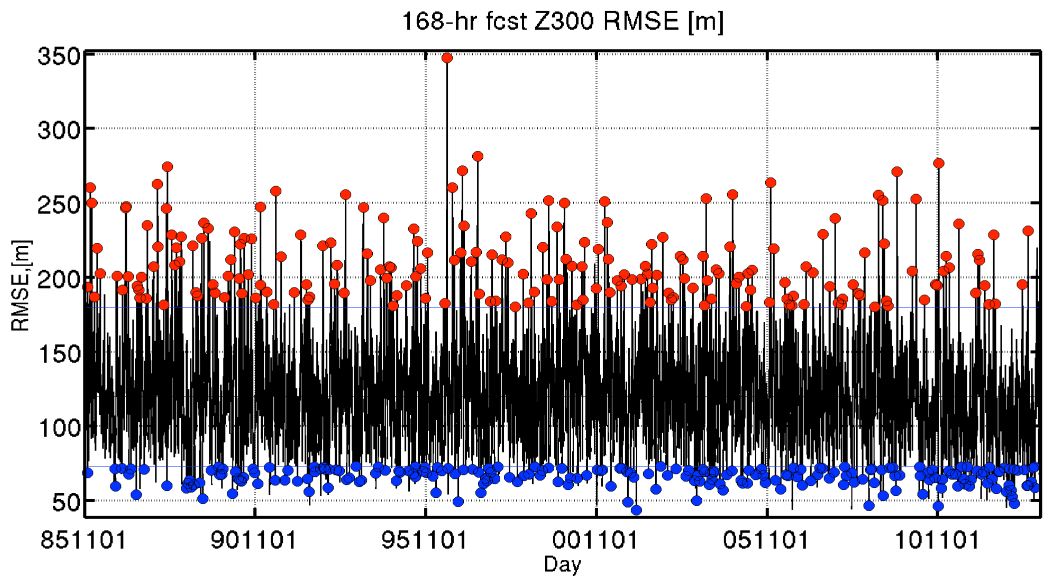


Figure 16. RMSE of 7-day control forecast of Z300 over east coast and western Atlantic oceans. Red dots show the large error cases while the blue dots show the good forecast cases.

Based on 203 large errors cases, we have composited Rossby wave packet amplitude anomaly relative to climatology as can be seen in figure 17. A positive wave packet anomaly propagates from day 2 to day 7, which suggests medium range forecast busts are associated with the existence of enhanced Rossby wave packets propagating across central and eastern Pacific. Large forecast errors tend to follow enhanced wave packets. From day 4 to day 7, large errors also tend to be maximized over the leading regions of positive wave packets amplitude anomaly.

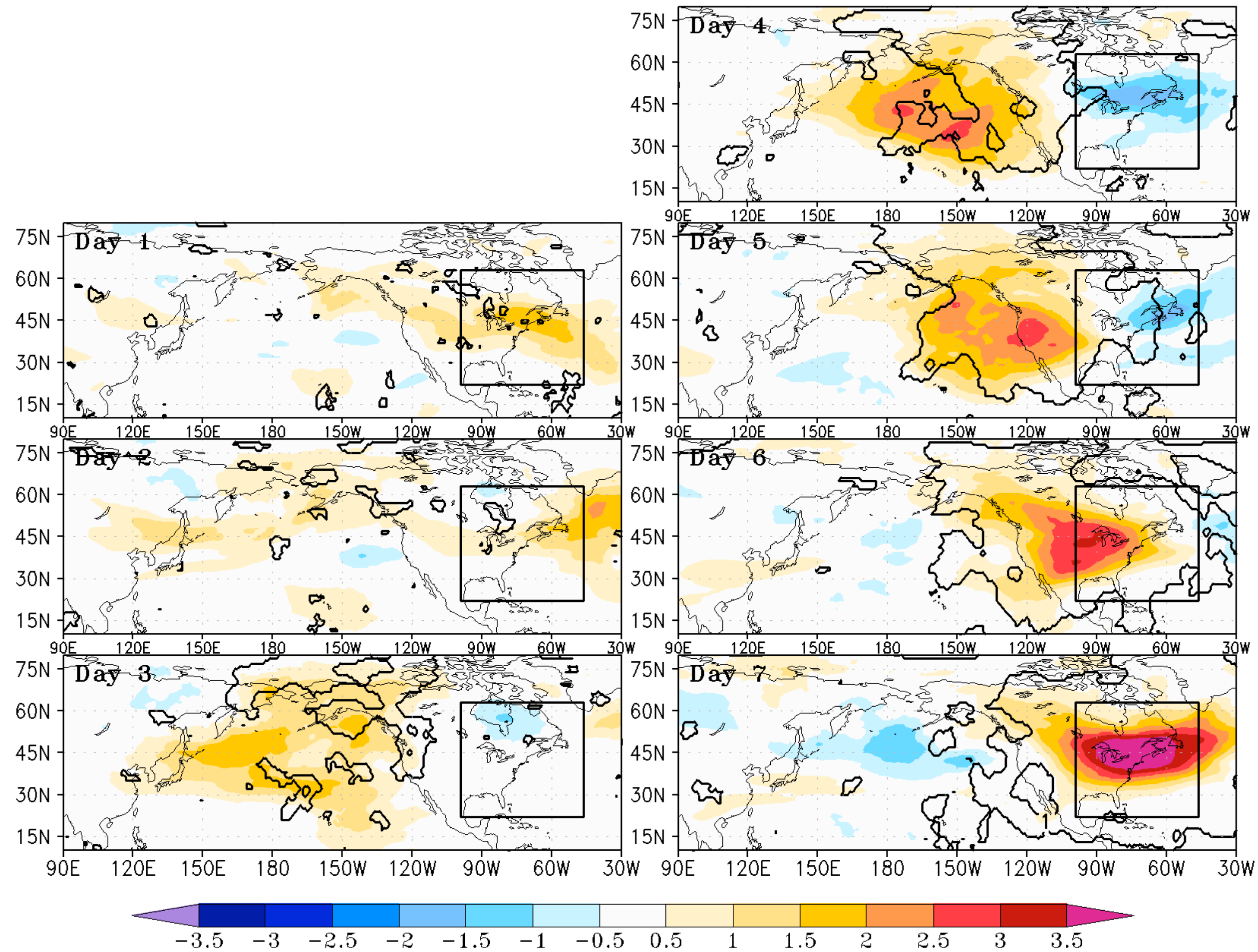


Figure 17. Composited RWPA anomaly (shaded) conditional on large 7-day forecast errors from day 1 to day 7. Day 7 is the verification time for large error cases. The black contour shows regions with significant larger absolute forecast errors than climatological errors.

iv. Climatology of ensemble sensitivity in multi-model ensemble

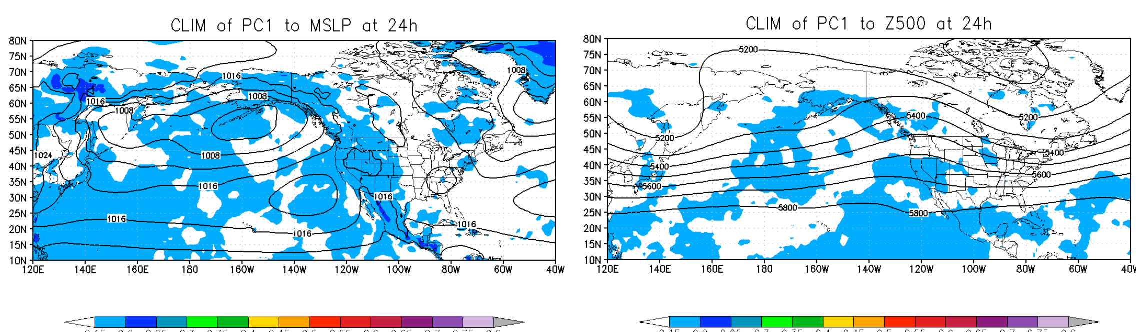
1. Motivation
    - Derive ensemble sensitivity to determine areas where perturbations produce the largest impact on the U.S. East Coast winter storms forecast
    - Investigate different paths of sensitivity for different development scenarios of cyclones
2. Data and methodology
    - A combination of NCEP, CMC and ECMWF ensemble data from TIGGE archive
    - U.S. East coast cyclone cases (102) are defined as minimum pressure <1000 hPa in GFS analysis at verification time over a verification box (32N-45N, 78W-62W)
    - 3-day and 6-day MSLP and geopotential height at 500 hPa (z500) forecast used
    - Ensemble sensitivity analysis using EOF PCs of MSLP as forecast metrics

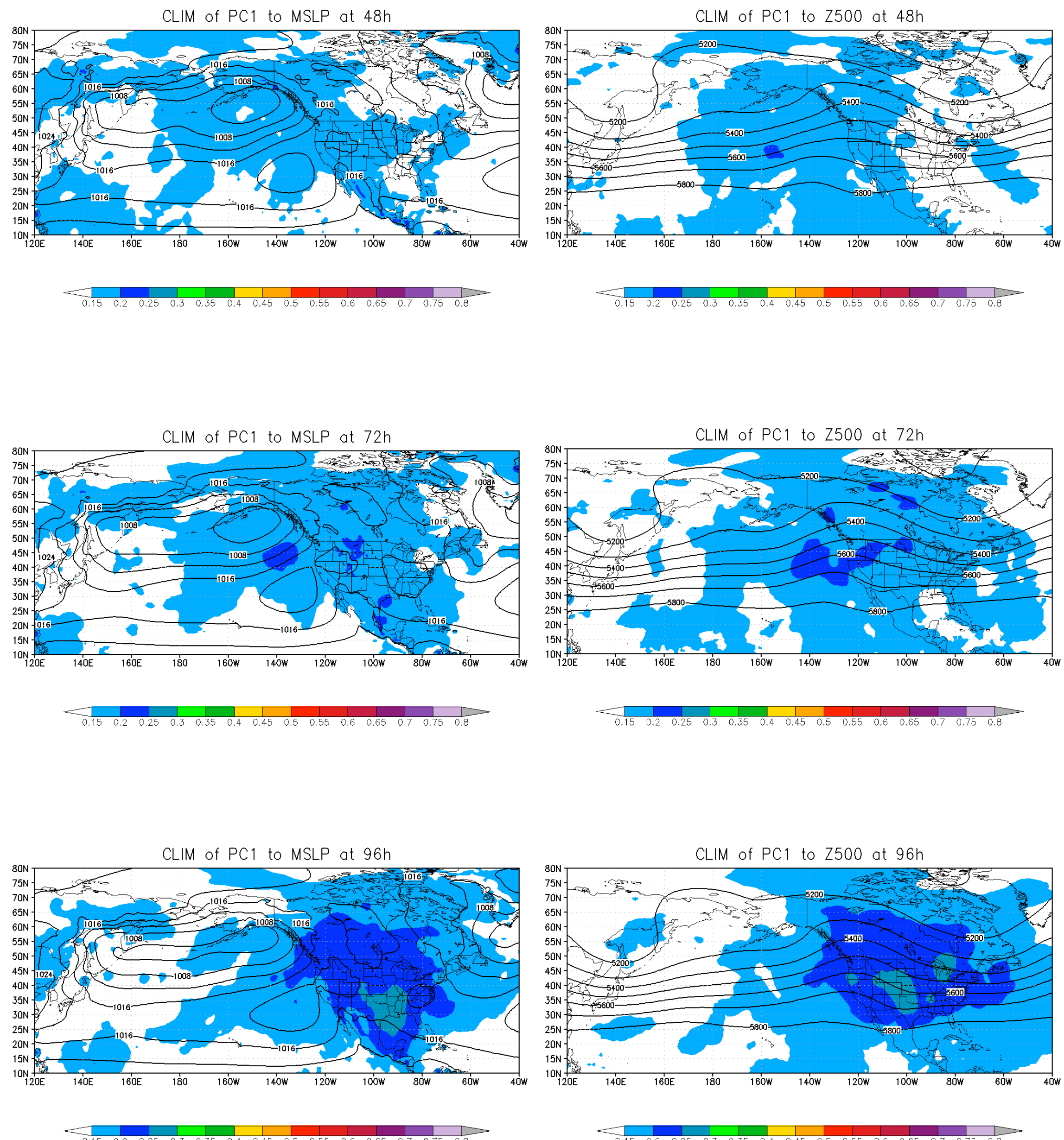
- Composite analysis
- EEOF method

Sensitivity analysis has proven to objectively quantify the relation between forecast metrics and state vectors at initial time or earlier forecast time steps (Torn and Hakim 2008; Ancell and Hakim 2008; Chang et al. 2013). Garcies and Homar (2009) applied ensemble sensitivity to the real atmosphere and derived climatologically sensitivity regions for Mediterranean intense cyclones. Our work follows our previous studies on applying ensemble sensitivity analysis to operational ensemble, and derives the model climatological sensitivity regions for U.S. East Coast cyclones to provide insights into the predictability of winter storms in medium range forecast.

Figure 18 shows the 6-day forecast sensitivity using EOF PC1 of MSLP at 144 h. Before 96h, the ensemble sensitivity is mainly over the North and East Pacific oceans, associated with the Aleutian low over northeastern Pacific. From 96h to 144h, surface sensitivity is associated with the cyclone originating from U.S. South, increasing from <0.3 at 96h to >0.8 at 144h. The surface sensitivity is mainly a monopole at 144h, indicating that it is associated with the cyclone intensity uncertainty. Meanwhile, the upper level sensitivity shows a dipole around the East Coast trough and its downstream ridge at 144h, again suggesting the strength of this trough and its downstream ridge have an impact on the surface cyclone intensity forecast.

Ensemble sensitivity for EOF PC2 (figure 19) shows robust signals at 48h associated with short wave troughs over central Pacific at both surface and upper levels. These sensitivity signals seemed to propagate eastward across western coast of U.S. at 72h. From 96h to 144h, the surface sensitivity signals showed a dipole associated with the surface cyclone, suggesting the EOF2 pattern is on average associated with the forecast uncertainty in cyclone positon. Meanwhile, the upper level sensitivity showed a triple-center structure, suggesting the shift of the East Coast trough and its adjacent systems could affect the cyclone position forecast uncertainty.




CLIM of PC1 to MSLP at 48h
CLIM of PC1 to Z500 at 48h
CLIM of PC1 to MSLP at 72h
CLIM of PC1 to Z500 at 72h
CLIM of PC1 to MSLP at 96h
CLIM of PC1 to Z500 at 96h
0.15 0.2 0.25 0.3 0.35 0.4 0.45 0.5 0.55 0.6 0.65 0.7 0.75 0.8

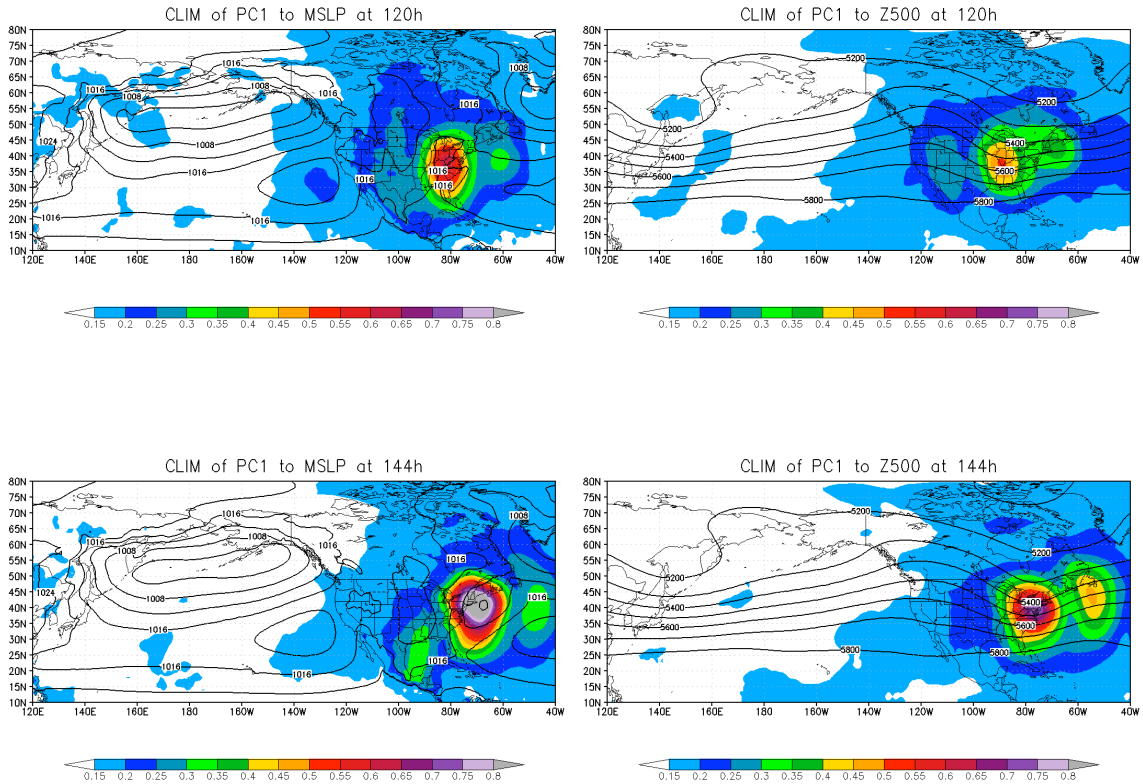
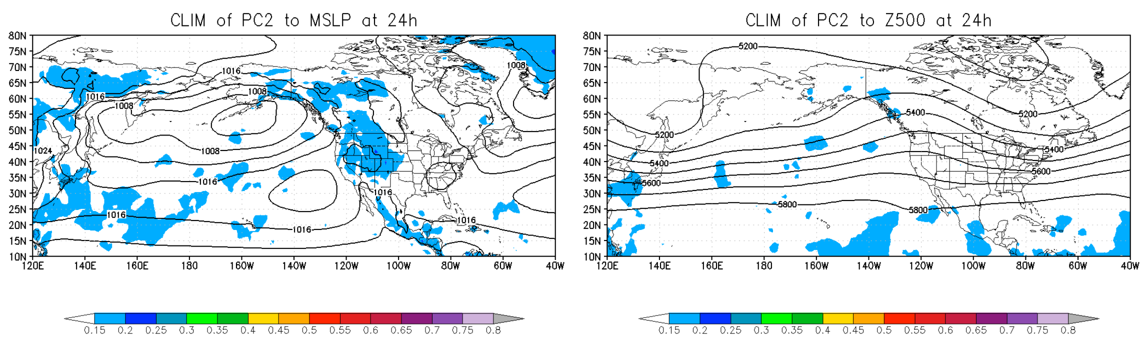
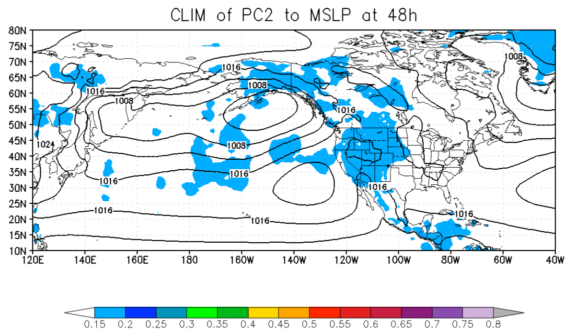
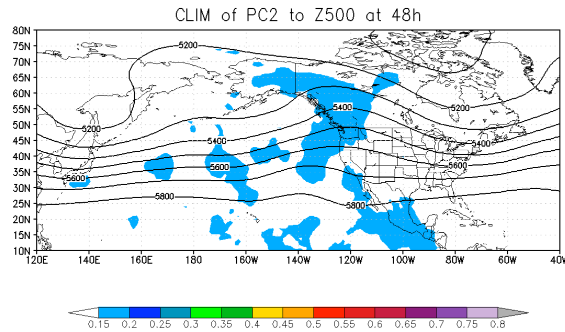
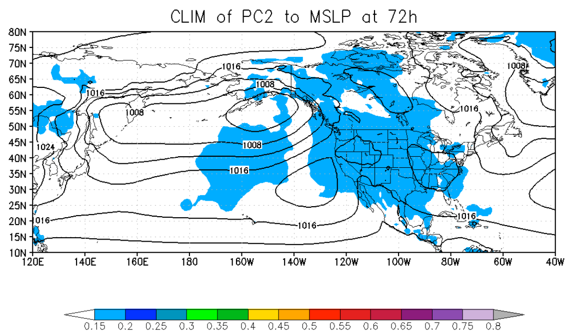


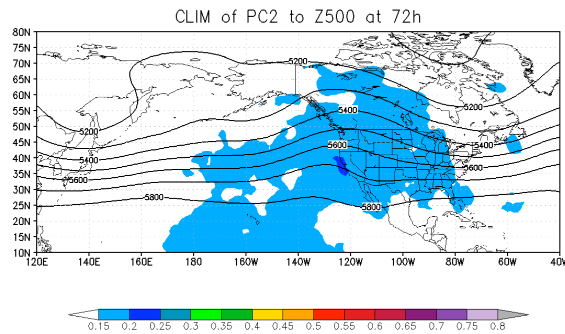
Figure 18. Mean of absolute ensemble sensitivity using EOF PC1 for MSLP (left, shaded) and Z500 (right, shaded) over 102 cyclone cases and ensemble mean MSLP (left, black contours) and Z500 (right, black contours) at different lead times. The verification time is 144h.

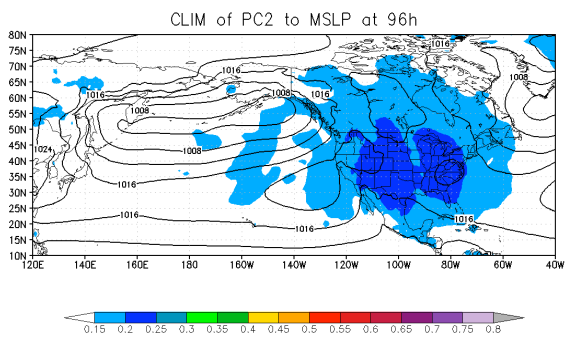


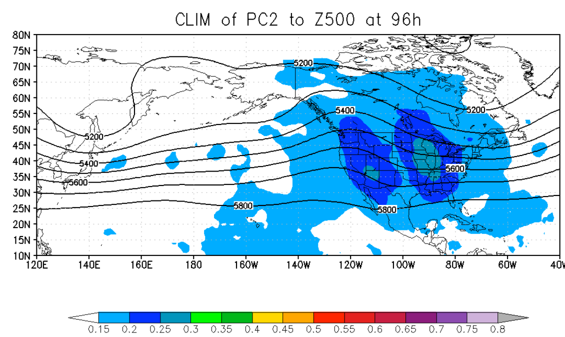

CLIM of PC2 to MSLP at 48h


CLIM of PC2 to Z500 at 48h


CLIM of PC2 to MSLP at 72h


CLIM of PC2 to Z500 at 72h


CLIM of PC2 to MSLP at 96h


CLIM of PC2 to Z500 at 96h

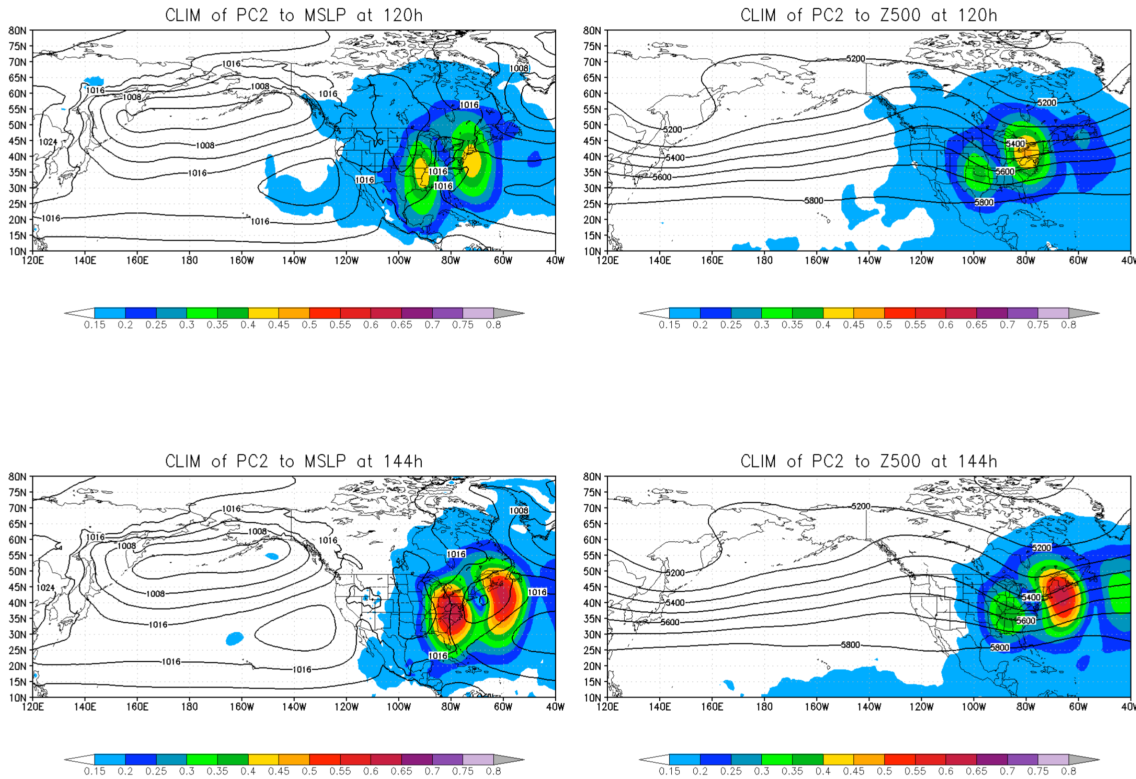


Figure 19. Mean of absolute ensemble sensitivity using EOF PC2 for MSLP (left, shaded) and Z500 (right, shaded) over 102 cyclone cases and ensemble mean MSLP (left, black contours) and Z500 (right, black contours) at different lead times. The verification time is 144h.

In ensemble forecast, each ensemble run represents a flow-dependent output. Also, there exists large inhomogeneity in U.S. East Coast cyclones. Therefore, the ensemble sensitivity could show different characteristics for different types of cyclone. Extended EOF (EEOF) has been performed to find different sensitivity paths as well as the corresponding cyclone development scenarios.

Figure 20 shows one EEOF pattern for sensitivity using EOF PC1. This pattern has large sensitivity associated with the eastern side of the surface low over Gulf of Alaska and its downstream ridge. The initial sensitivity develops to a downstream sensitivity pair associated with a surface low east of mid-Atlantic coast at 72h. Meanwhile, the upper-level trough also shows a pair of sensitivity around the East Coast trough. This pattern suggests that the northeastward of the Alaska low and the upper level trough are associated with the weakening and eastward shift of the surface cyclone at verification time.

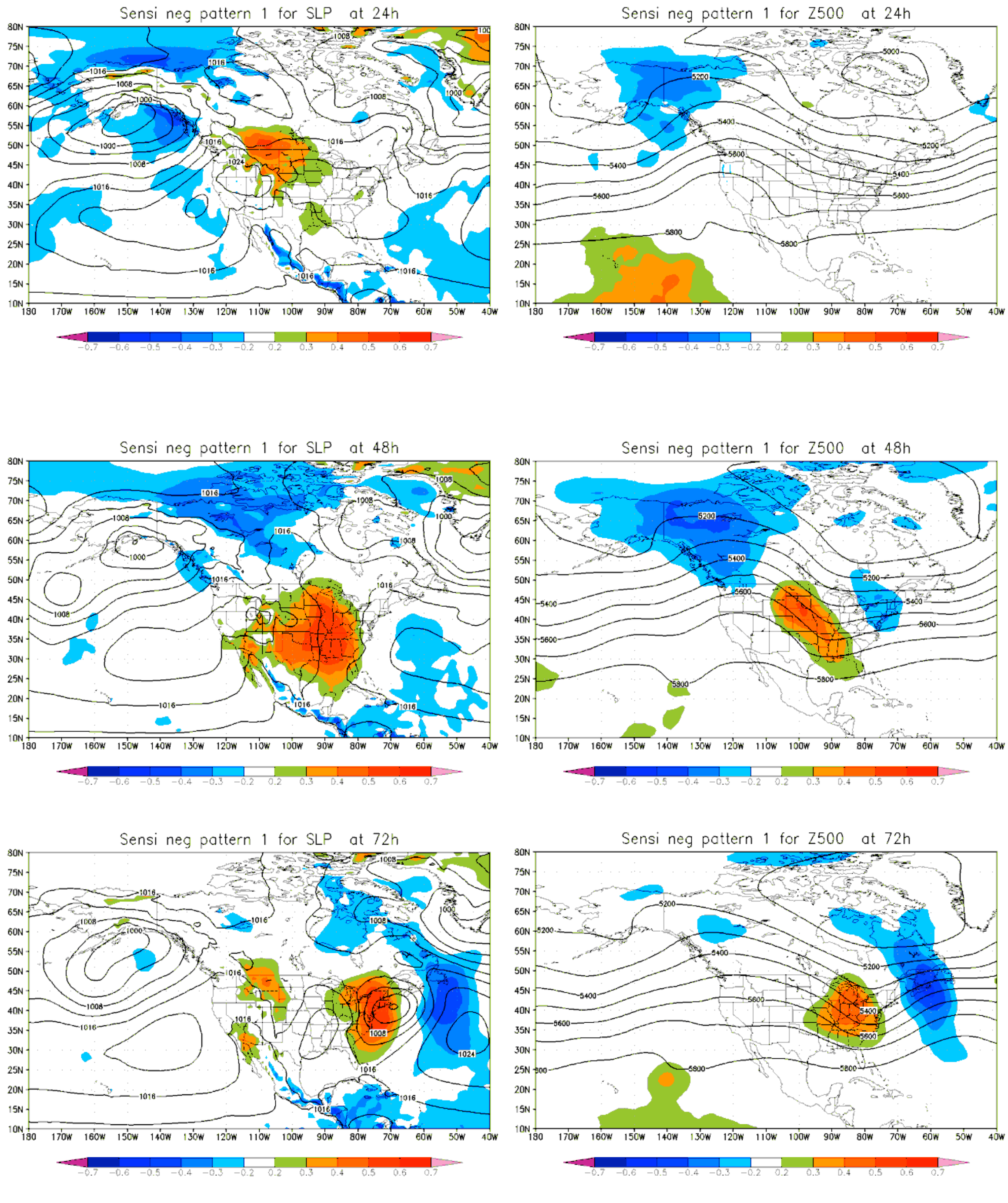


Figure 20. EEOF negative pattern 1 of ensemble sensitivity for EOF PC1 of MSLP (left, shaded) and Z500 (right, shaded) and the corresponding mean MSLP (left, black contours) and Z500 (right, black contours).

v. Spread anomaly tool

Mean and spread (standard deviation) are the most common viewed outputs of statistics from an ensemble. These two statistics provide a "first-order" assessment of the model solution, detailing the "best-guess" (mean) of all the ensemble members as well as the amount of variation (spread) represented. These statistics are often plotted onto a spatial 2D field for outputted variables, as well as post-processed variables.

Although spread provides a basic assessment of the variation or uncertainty of a model run, it does not necessarily help to detail a more in depth understanding of what makes the forecast uncertain. It also does not provide an explanation as to how a particular anomaly pattern in a forecast compares to similar anomaly patterns in space, orientation, and/or magnitude. The issue arises as to how we can best assess ensemble variability in the reference of similar events in the model climatology ("m-climate"). The Ensemble Situational Awareness Table (ESAT) managed by the National Weather Service compares forecasts from the North American Ensemble Forecast System (NAEFS) and Global Ensemble Forecast System (GEFS) to reanalysis climatologies (r-climate) and m-climates. Standardized anomalies, percentiles and return intervals are calculated to assist in identifying potentially significant weather events.
The objectives of this spread tool are to:

- Determine how to assess spread of a forecast in relation to similar events and develop a "metric" of confidence, or a method to convey the uncertainty, of a forecast in relation to events of similar anomaly.
- Determine the best practices of applying the tool to the forecast process.
- Find reasons behind the causes of anomalous spread in high spread events.
- Determine differences between the GEFS and GEFS Reforecast.

<u>Data/Methods</u>
The ESRL PSD GEFS Reforecast 2 (1 degree x 1 degree resolution) is utilized as the m-climate for this project from 1984-2015 for December, January, and February (DJF). The next version of the program will use a 3-month m-climate centered about the month of the forecast to expand the range beyond DJF (i.e. for March and November cyclones) as well as expand to 2016 for more data. Forecast case studies are taken from the GEFS retrospective run, TIGGE, and GEFS real time runs via NOAA. The real time product uses the THREDDS data server to access the latest GEFS runs' ensemble mean and spread. For the purposes of this project, we utilize mean sea level pressure, 500mb geopotential heights, and surface temperature to assess the spread anomaly and run the tool on. Future variables will include 850mb temperature and precipitable water.
The forecast are loaded as well as the m-climate for the variable and forecast hour. The data is input as a 3-dimensional array, which has dimensions of time, longitude, and latitude of the continental U.S. Each grid-point for the 30-year climatology has a distribution on which statistics can be performed upon. The distributions of variables are essentially Gaussian in nature, with a slight extreme in the tails (due to the rare extreme low or high pressure, for instance). Due to this, we can calculate a z-score, or standardized anomaly, for the forecast and all of the values within the m-climate at each grid-point,

$$SA = \frac{(F_{x,y} - C_{x,y})}{\sigma C_{x,y}} \qquad (1)$$

where F is the forecast, C is the m-climate, and sigma is the standard deviation of the m-climate. Given the z-score, we can use the values as a range to subset a new m-climate of similarly anomalous events, and a subsequent m-climate of spread for those same forecasts. For the purposes of obtaining a large enough sample size, we find that a standard different range will be discussed as well). When a valid value in the mean climatology is found the spread of that forecast data-point is put into the new m-climate. The resulting outcome is a modified ensemble mean m-climate, including only the spread from the m-climate for events of similar magnitude. In the cast of insufficient sample size, a 2x2 box is initiated for less than 20 sample data values, and a 5x5 for less than 10 about the grid-point. The resulting modified spread m-climate can be input into equation (1), and a new standardized spread anomaly product can be created (Fig. 21).

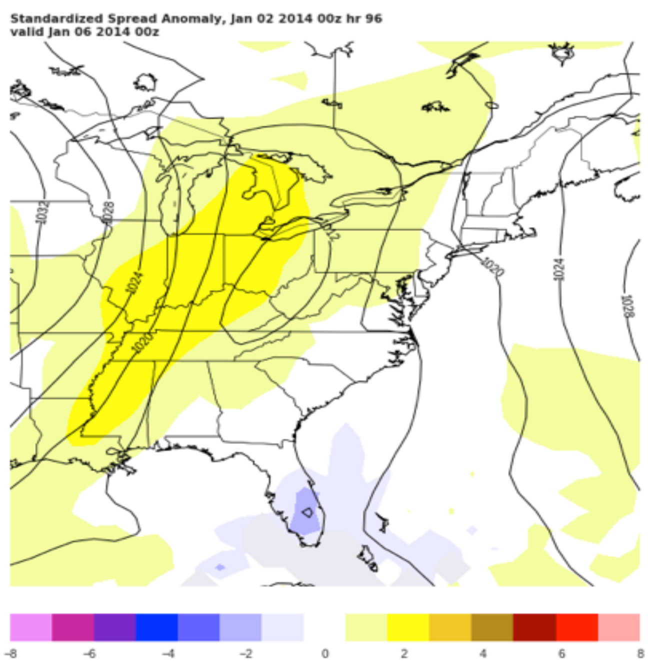


Figure 21: Standardized spread anomaly for the 2 Jan 2014 00z 96 hour forecast valid 06 Jan 2014 00z. Values are log normalized to shift the distribution closer to the Gaussian distribution. Note a 2+ z-score in a non-negligible number of locations in the spatial field.

<u>References:</u>

Ancell, B., and G. J. Hakim. "Comparing adjoint-and ensemble-sensitivity analysis with applications to observation targeting." *Monthly Weather Review* 135.12 (2007): 4117-4134.

Bougeault, P., et al. (2010). "The THORPEX Interactive Grand Global Ensemble." *Bulletin of the American Meteorological Society,* **91**, 1059-1072.

Chang, KM, M. Zheng, and K. Raeder. "Medium-range ensemble sensitivity analysis of two extreme Pacific extratropical cyclones."*Monthly Weather Review* 141.1 (2013): 211-231.

Garcies, L., and V. Homar. "Ensemble sensitivities of the real atmosphere: application to Mediterranean intense cyclones." *Tellus A* 61.3 (2009): 394-406.

Harr, P. A., et al. (2008). "Predictability Associated with the Downstream Impacts of the Extratropical Transition of Tropical Cyclones: Methodology and a Case Study of Typhoon Nabi (2005)." *Monthly Weather Review,* **136**, 3205-3225.

Keller, J. H., et al. (2011). "Characteristics of the TIGGE multimodel ensemble prediction system in representing forecast variability associated with extratropical transition." *Geophysical Research Letters* **38**, L12802.

Korfe, N.G., and B.A. Colle, 2017: Evaluation of cool season extratropical cyclones in a multi-model ensemble for eastern North America and the western Atlantic Ocean. In press to *Wea. Forecasting.*

Torn, R. D., and G. J. Hakim. "Ensemble-based sensitivity analysis."*Monthly Weather Review* 136.2 (2008): 663-677.

Zheng, M., K.M. Chang, and B.A. Colle. Ensemble sensitivity tools for assessing extratropical cyclone intensity and track predictability." *Weather and Forecasting* 28.5 (2013): 1133-1156.

Zheng, M., Chang, E.K., Colle, B. A., Luo, Y., & Zhu, Y., 2017: Applying fuzzy clustering to multi-model ensembles for validation and scenario identification of U.S. East Coast winter storms. *Wea. Forecasting,* **32**, 881-903.

3. Recent Interaction with Operational CSTAR Partners

For this project several operational tools were developed to interpret multi-model ensemble predictions for extratropical cyclones along the U.S. East coast for the medium range (days 3-6). The tools include:

1. An ensemble sensitivity tool to understand the upstream origins of some of the forecast uncertainty in the models (http://breezy.somas.stonybrook.edu/CSTAR/Ensemble_Sensitivity/EnSense_Main.html; Zheng et al. 2013);
2. Fuzzy clustering tool to objectively identify different East coast storm scenarios/clusters: (http://breezy.somas.stonybrook.edu/CSTAR/Ensemble_Sensitivity/FC_Main.html; Zheng et al. 2017);
3. A multi-model ensemble cyclone tracking page, which includes bias correction for the GEFS system: (http://smokey.somas.stonybrook.edu/cyclonetracks/).
4. An ensemble Rossby wave packet page for the GEFS system: http://smokey.somas.stonybrook.edu/wavepackets/index.html.
5. An ensemble spread anomaly tool (http://blue.somas.stonybrook.edu/ssa/index.html) for similar anomalous weather days as defined using other reforecast tools, such as the Situational Awareness Tool.

As part of our CSTAR project, Dr. Colle gave webinars on how to interpret/use many of the above tools. We have also introduced some of them to the Weather Prediction Center (WPC) Winter Weather Experiment (WWE). We also have a CSTAR list-serve to discuss the tools during East coast winter storms. However, one challenge has been to determine how to use these tools in the forecast process. Dr. Colle visited NCEP/WPC from 9 January 2017 to 20 January 2017 to discuss the CSTAR project and shadow the forecasters there to determine how best to use these tools (strengths and limitations) and think of new ways to display ensemble data. There were four main goals of the two-week visit to WPC:

1. Current Use of Ensemble Products in the Forecast Process. One goal is to learn how forecasters use existing NWP models and tools in the forecast process for various winter weather phenomena during the cool season, ranging from heavy orographic precipitation, southern U.S. cool season convective storms, and extratropical cyclones. As a researcher, there is often a disconnection between knowing what tools/models may be available and how they are used operationally. Understanding this helps transition any research to operations. This will be accomplished by shadowing forecasters on the medium range and short-range WPC desks each day. Dr. Colle will learn about current pros and cons of the ensemble products from a forecaster perspective.
2. CSTAR Tool Demonstration. As part of the daily forecaster shadowing, Dr. Colle will provide his insight how the CSTAR tools listed above may be able to be used in the forecast process. This interaction will help the forecaster(s) better understand these tools, while at the same time providing feedback from the forecasters on how to improve these tools, which also will help the R2O process. Dr. Colle will work with the forecasters to see if a protocol can be developed to incorporate these tools into the forecast process. Case examples will be saved for future reference and collaboration.
3. Transition of Products to WPC. Many of the tools listed above are driven by web pages run at Stony Brook University. However, the calculations (for ensemble sensitivity and fuzzy clustering) are done at EMC, where the GEFS, EC, and CMC ensemble data resides. It may be more efficient if some of these tools were implemented directly at WPC, so Dr. Colle will work with the staff there to determine how best to port over. We have precedence in doing this, since Mike Bodner at WPC helped to get Stony Brook's Rossby Wave Packet amplitude code working at WPC.
4. Seminar and Tutorial. Dr. Colle will provide a seminar on some of the latest understanding of single and multi- snowbands within extratropical cyclones (from his NSF project), as well as a tutorial overview of the tools in his CSTAR project.

Below was the schedule for B.Colle:

Monday 1/9/2017:
9-10am: Tour and Introductions
10-10:30: Weather Discussion
10:30-1045: Brief Meeting with the Winter Weather Experiment (WWE) Group
10:45-noon: Discuss SBU CSTAR tools and output with Mike Bodner et al.
1-2:30: Met with Dan Petersen at the medium range desk
2:45-3:45: Mike Bodner presented North Pacific Jet Tool for 8-10 day forecast

Tuesday 1/10/2017:
8:45-9:45am: Meet with Dan Petersen at the winter weather desk
10-10:30: Weather Discussion
11-noon: Met with Jim Nelson to discuss AWIPS update, coordination with WFOs, and precipitation type issues
Noon-1:30: Met with Mike Erickson to discuss verification (MET software)

1:30-3pm: Met with Mary Beth at medium range desk

Wednesday 1/11/2017:
9-10am: Seminar about EMC GEFS updates
10-10:30: Weather Discussion
10:30-noon: Meet with medium range group to discuss 8-10 day and how to extend Stony Brook tools to this lead time.
1:30-2:30pm: Met with Bruce Veenhuis to discuss post-processing (National Blend, MOS, bias correction, and RTMA).
2:30-3:30pm: Met with Bruce Sullivan at winter weather desk

Thursday 1/12/2017:
9-10am: Seminar about EMC GEFS updates
10:10:30: Weather discussion
10:30-noon: 8-10 day experiment test day
2:30-3:30pm: Met with Marc S. on day 2-3 QPF desk

Friday 1/13/2017:
9-10am: Met with Yuejian Zhu at EMC to discuss CSTAR collaboration
10-10:30: Weather discussion
10:30-11:00: Met with Bill Laberson about Situational Awareness Table
11-11:30am: Met with David Novak to discuss ensembles within WPC

Tuesday: 1/17/2017:
10-10:30: Weather Discussion
10:30-noon: Participate in Winter Weather Experiment
1-2pm: Met with Mark Klein about WWE, flash flood expt, and forecaster training
2-3pm: Met with forecaster (Mike) at the Alaska desk.
3-4pm: Met with Chris Bailey about Autoblend

Thurs: 1/19/2017:
9-10 am: Met with David Novak and Jim Nelson about my visit
10-10:30am: Weather Discussion

Some of the issues discussed:
1. Ensemble tools are still limited:
    * Medium range: spaghetti plots separated by model (EC, GEFS, and CMC)
    * GEFS stand anomaly to day5
    * Use anomaly in CPC teleconnection tool, see what members match teleconnection pattern (very static – need to combine with evolution and clusters).
    European Center (EC) in 3 clusters (1-17; 18-24,35-50)
2. 8-10 day: extend CSTAR tools; combine with other new tools, such as Pacific jet tool: (ensemble sensi targets upstream jet region, or Rossby wave packet mean/spread – is there a dynamical link to the location.
3. Winter Weather: NAM freezing rain prob, but not GFS and GEFS.
4. NAWIPS used at WPC, but AWIPS at WFOs

5. No ECs at WFOs

3.1 Winter Weather Experiment

Stony Brook University actively participated in the Winter Weather Experiment (WWE), which was held from mid-January to mid-February 2015 at the Weather Prediction Center (WPC). The goal of the WWE is to test various ensemble products related to winter weather, such as probabilistic maps for snowfall, freezing level, precipitation amount, etc…, and products related to snow liquid water ratio. Our involvement also meant that we got to incorporate our ensemble sensitivity tools into the medium range part of the experiment, which typically took place during the afternoon.

Three Stony Brook students participated for two weeks of the 4-week experiment (Matthew Sienkiewicz and Sara Ganetis in week 1, and Nathan Korfe in week 3). One week of SBU participation (week 2) was cancelled because of the lack of significant winter weather. The students not only participated in the discussions, but each day they wrote a summary of discussions each day, and sent it to the Stony Brook CSTAR listserve (cstar_stony_brook@infolist.nws.noaa.gov). Below are a couple of example write-ups.

Day 1 of WWE (13 January 2015):

Experiment Forecast Challenge

A shortwave disturbance is forecast to swing into the Mid-Atlantic region during the first half of Wednesday, 14 Jan. The objective of the winter weather experiment in the short range is to quantify how much snow is expected to fall in this region in a 24-h period. For the period ending at 00Z 15 Jan, the group used a variety of operational and experimental guidance for the forecast. The main challenge regarding this event was with precipitation type and how that would affect the snowfall amounts.

Highlighted Tools & Rationale

The conventional guidance that we used was the 09Z 13 Jan operational SREF and the 12Z 13 Jan NAM. The SREF mean 24-h snowfall valid at 00Z 15 Jan showed a swath of 1-4" extending from central VA to the coast of the Delmarva Peninsula so that was where we focused our attention. The SREF probability of >1" in 24 hours (Fig. 22a) ending 00Z 15 Jan exceeded 80% just south of the Potomac River. Figure 22b shows the individual members colored by model core (ARW, NMM, NMMB), the ARW members (labeled EM on the plot) were the coldest due to a deeper shortwave and thus had a larger areal extent (and low spread) of 1" of snow.

Moving our attention to more experimental guidance, we decided to weight our forecast on the 12Z NAM and used an additional NAM product that uses the instantaneous rime factor (from the lowest model level) to decrease the snow-to-liquid ratio (SLR) if the rime factor is large. As such, areas of precipitation type transition may become clearer if the modeled snowfall decreases compared to the operational guidance. The rime factor and percent of frozen precipitation (reaching the ground) provided slightly increased confidence to our snowfall amounts in northern VA but gave us the grounds to trim down some of the amounts in southern VA (Fig. 23). To conclude, we did forecast >1" of snow north and west of D.C. extending into Northern Virginia. The microphysical products helped increase our confidence (slightly) for this event but it remains a relatively low-confidence forecast.

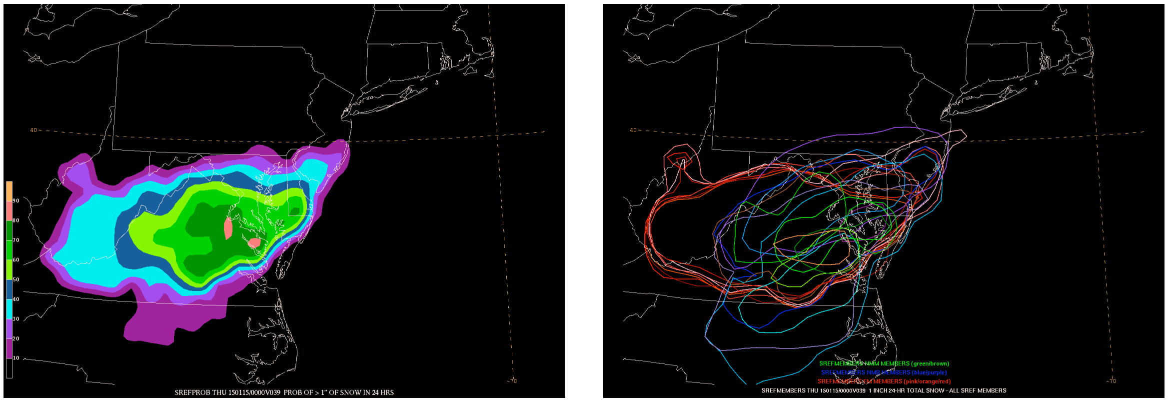


Figure 22. (a) Probability exceeding 1 inch in 24 h using all SREF members ending at 0000 UTC 15 January 2015 (hour 39). (b) Same as (a) except for spaghetti plot of the different SREF cores (see color legend at bottom of plot) showing the outline of 1 inch in).

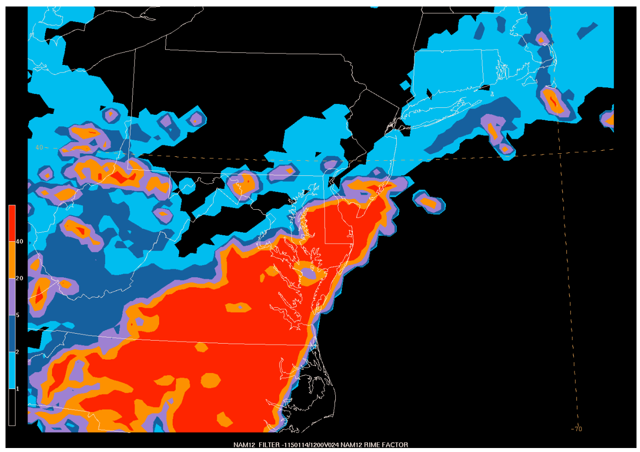


Figure 23. Rime factor for the 12-km NAM at 1200 UTC 14 January 2015 (hour 24).

3.2 Environmental Model Center (EMC) at NCEP

We collaborated with EMC (Yuejian Zhu and Yan Luo) to get our clustering tool operational for this winter season using the NCEP, CMC, and GEFS ensembles. Our original code was in Matlab to do the clustering, but we converted the code to FORTRAN, so it can run at EMC. The clustering tool was made operational in late January 2015. Examples products were shown above.

3.3 Weather Prediction Center

Stony Brook was the primary university involved in the 2015 Winter Weather Experiment in January – February 2015. Three graduate students participated at WPC, and posted daily reports (see above) on the Stony Brook CSTAR Listserve, which were followed by other online discussions. The PI (Colle) also participated in most of the 12-12:30pm WWE weather briefings, which were broadcast online for universities and NWS offices.

3.4 National Weather Service Forecast Offices

The National Weather Service–New York City continues to maintain the ALPS (ensemble viewer) installation as changes are made to the NWS computer systems and data ingest. The ALPS software is described in more detail in previous CSTAR reports. The latest instructions to install ALPS can be found on our CSTAR web site:
http://dendrite.somas.stonybrook.edu/CSTAR/Tutorials.html.

Stony Brook also collaborated with the KOKX forecast office (as well as WPC) on the analysis of the 8-9 February 2013 blizzard on Long Island, during which nearly a meter of snow fell. The rapid precipitation phase transitions around the primary snow band are a good test for the dual-Pol radar data and models. Therefore, a Maproom article has been published in the *Bulletin of the American Meteorological Society* (see products below).

The field offices have been using the ensemble sensitivity tool in operations for some major storm events. Below is an example from an AFD from the New York City NWS office:

AREA FORECAST DISCUSSION NATIONAL WEATHER SERVICE NEW YORK NY 422 PM EST **SAT** JAN 31 2015 .SYNOPSIS... HIGH PRESSURE BRIEFLY BUILDS IN TONIGHT...BEFORE A COLD **FRONT** APPROACHES AND CROSSES ON SUNDAY. LOW PRESSURE APPROACHES FROM THE OHIO VALLEY AND PASSES JUST SOUTH OF LONG ISLAND MONDAY. THE LOW TRACKS WELL TO THE EAST MONDAY NIGHT WITH HIGH PRESSURE BUILDING IN ITS PLACE. HIGH PRESSURE WILL LARGELY PREVAIL FOR THE REMAINDER OF THE WEEK WITH THE EXCEPTION OF A COLD FRONTAL PASSAGE LATE WEDNESDAY. A DEVELOPING LOW OVER THE GULF COAST STATES LATE IN THE WEEK LOOKS TO TAKE A TRACK WELL SOUTH AND EAST OF THE AREA. && .NEAR TERM /THROUGH SUNDAY/... NEAR **ZONAL FLOW** TONIGHT WITH WEAK HIGH PRESSURE BUILDING IN FROM THE WEST THIS EVENING...AND THEN A POLAR **FRONT** APPROACHING FROM THE GREAT LAKES. WINDS AND GUSTS GRADUALLY SUBSIDE THROUGH MIDNIGHT...WITH INCREASING **HIGH CLOUDS** OVERNIGHT. TEMPS WILL NOT BE QUITE AS COLD WITH LOWS GENERALLY IN THE TEENS...EXCEPT 20S NYC/NJ **METRO**...WITH NEUTRAL TEMPS **ADVECTION** AND LIMITED OPPORTUNITY FOR **RADIATIONAL COOLING**. COULD SEE SOME SINGLE DIGITS IN PINE BARRENS AND INTERIOR TRI-STATE WITH DECOUPLING BEFORE CLOUDS MOVE IN. PHASING NORTHERN STREAM AND SW US **SHORTWAVE** ENERGY BEGINS TO DIG INTO THE CENTRAL US ON SUNDAY WITH GENERALLY **WSW** UPPER **FLOW**. A WEAKENING POLAR **FRONT** APPROACHES THE REGION ON SUNDAY...BUT OTHERWISE INCREASING AND LOWERING CLOUDS AHEAD OF NEXT APPROACHING STORM SYSTEM. A FEW SNOW FLAKES MAY BE OBSERVED OVER NJ VERY LATE IN THE DAY. TEMPS GENERALLY IN THE UPPER 20S TO LOWER 30S. && .SHORT TERM /SUNDAY NIGHT THROUGH MONDAY/... MAIN STORY WILL BE NEXT IN A SERIES OF WINTER STORMS AFFECTING THE REGION SUNDAY NIGHT INTO MONDAY. MODELS ARE IN GENERAL AGREEMENT WITH LOW PRESSURE TAKING SHAPE OVER THE CENTRAL PLAIN TONIGHT...TRACKING INTO THE OHIO VALLEY ON SUNDAY AND THEN JUST SOUTH OF THE REGION SUNDAY NIGHT. OVERALL TREND HAS BEEN SLIGHTLY FURTHER NORTH AND WETTER WITH THIS SYSTEM. BASED ON COUPLED **JET** STRUCTURE...TIGHT LOW /MID- LEVEL **THERMAL** STRUCTURE...AND GULF **MOISTURE** FEED OF 2-3 STD PWATS...CONFIDENCE IS HIGH IN AT LEAST A 3/4 TO 1 1/2 INCH **QPF** EVENT FOR THE ENTIRE AREA. THIS IS BOURNE OUT IN SREF/GEFS/CMCE 1 INCH/24 **HR** **QPF** PROBS. SIGNIFICANT UNCERTAINTY LIES IN P-TYPE...PARTICULARLY ALONG THE COAST...WITH A SPREAD OF ABOUT 50 MILES BETWEEN OPERATIONAL SOLUTIONS AND **ENSEMBLE** MEANS IN HOW CLOSE TO THE COAST THE LOW GETS AND THE POSITION OF THE TIGHT **THERMAL** **GRADIENT** ACROSS THE REGION. SBU **ENSEMBLE** SENSITIVITY SHOWING THIS N/S TRACK UNCERTAINTY IN THE 00Z GEFS/CMCE/**ECMWF** **ENSEMBLES**...WITH SENSITIVITY LYING IN DEGREE OF PHASING/INTERACTION BETWEEN NORTHERN STREAM **SHORTWAVE** ENERGY DROPPING DOWN THROUGH MONTANA AND THE **CUTOFF** UPPER LOW OVER THE SW US...AND RESULTANT STRENGTH OF **TROUGH** APPROACHING THE EAST COAST SUNDAY NIGHT/MONDAY AND **UPSTREAM** RIDGING. A STRONGER **TROUGH**...MEANING MORE **UPSTREAM** RIDGING...AND A FURTHER NORTH TRACK. THIS INTERACTION IS NOT COMPLETE FOR ANOTHER 24 HOURS...SO COULD TAKE UNTIL THEN FOR CONFIDENCE IN P-TYPE FORECAST INCREASES. WITH CONFLUENT **FLOW** ACROSS NORTHERN NEW ENGLAND/SE CANADA...ARCTIC HIGH SHOULD BE LOCKED INTO THAT AREA. PREFER TIGHTER **THERMAL** STRUCTURE ALOFT OF **NAM**/**ECMWF**/CAN VERSUS MORE DIFFUSE **THERMAL** STRUCTURE OF **GFS** IN THIS SITUATION...BUT QUESTION IS LOCATION OF THIS **BAROCLINIC ZONE**. MODEL SOLUTIONS POINTING TO DIFFERING DEGREES OF WARMING IN THE 800-950 LAYER MON MORNING INTO EARLY **AFT**...BUT TREND IS TOWARDS MORE WARMING.

4. Products and Presentations

4.1 Products

In addition to the several ensemble tools developed on our CSTAR web page we also explored some of the other tools available to forecasters.

4.1 AWIPS ensemble display

Several ensemble products were tested at ER with the new AWIPs build. Figures 24 and 25 shows two sample products. Figure 24 shows an example in which the user can select the ensembles used in the ensemble mean as well as a distribution of precipitation amounts for a point of interest. This helps put the ensemble mean in perspective by also illustrating the maximum and minimum values at this point. Figure 25 shows the AWIPS capability in which the user can display the different member solutions of interest as a series of post-it charts.

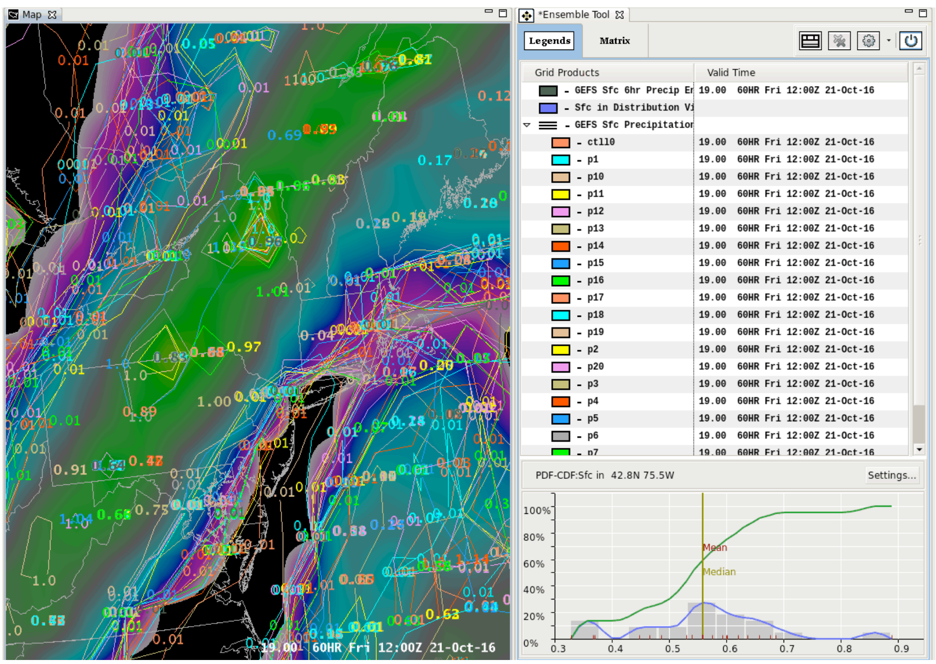


Figure 24. An example mean precipitation plot and histogram for a point using the GEFS ensemble and the AWIPS ensemble system.

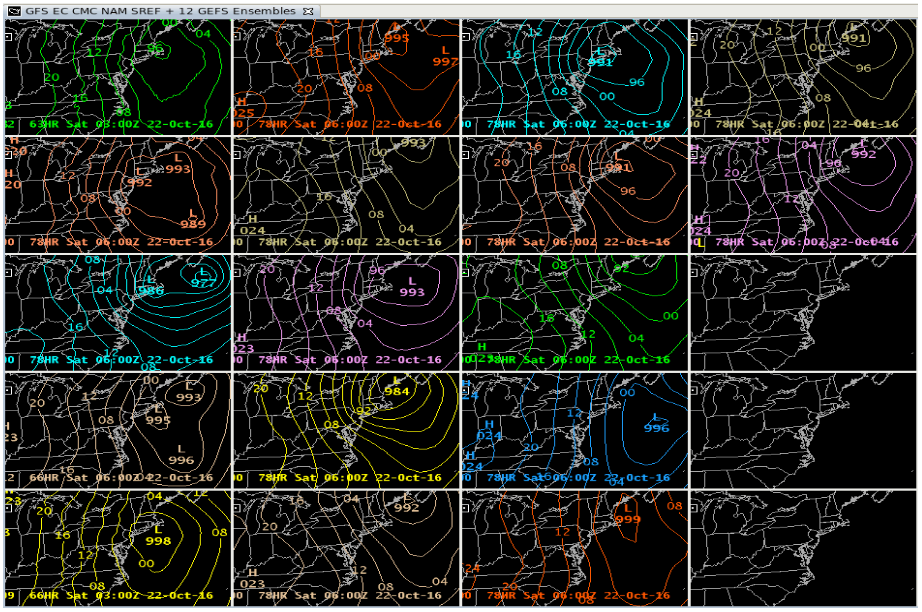


Figure 25. An example of the AWIPS ensemble tool showing sea level pressure forecasts from 9 different GEFS members.

4.2 CSTAR Tutorials

Dr. Brian Colle gave a webinar/tutorial to the KBOX (Taunton, MA) office on 24 January 2014 on RWPS and ensemble sensitivity.
https://drive.google.com/file/d/0B_ap3QL2sodKb280Z3Vka3dnRVE/edit?usp=sharing

There are other detailed tutorials were added to the ensemble sensitivity web page, with the direct link:
http://dendrite.somas.stonybrook.edu/CSTAR/Ensemble_Sensitivity/Extras/ReadMe_PlotInterpretations.pdf

The following three online tutorials are online: http://dendrite.somas.stonybrook.edu/CSTAR/Tutorials.html. As a result, there are now 2-3 hours of training material developed and available for forecasters.

4.4 Theses, Papers, and Presentations

The following formal papers have resulted from this CSTAR effort and support. Much effort was done this past 6 months working on papers in preparation given the Ph.D. thesis by Minghua Zheng.

Formal papers:

Zheng, M., Chang, E.K., Colle, B. A., Luo, Y., & Zhu, Y., 2017: Applying fuzzy clustering to multi-model ensembles for validation and scenario identification of U.S. East Coast winter storms. *Wea. Forecasting,* **32**, 881-903.

Korfe, N.G., and B.A. Colle, 2017: Evaluation of cool season extratropical cyclones in a multi-model ensemble for eastern North America and the western Atlantic Ocean. In press to *Wea. Forecasting.*

Zheng, M., Chang, E.K., Colle, B. A., Luo, Y., & Zhu, Y., 2017: Using EOF and fuzzy clustering methods to evaluate U.S. East Coast winter storms in multi-model ensembles. In preparation for *Wea. Forecasting.*

Zheng, M., Chang, E.K., & Colle, B. A., 2017: Impacts of upper level Rossby wave packets on medium-range forecast errors and uncertainties. In preparation.

Zheng, et al, 2017: Ensemble sensitivity of U.S. East Coast winter storms: the multi-model climatology and paths of forecast uncertainty in medium range. In preparation.

Wirth, V., M. Riemer, E. K. M. Chang, and O. Martius, 2016: Rossby wave packets on the mid-latitude Rossby waveguide -- a review. In press to Monthly Weather Review.

Other published papers:

Layer, M. and B. A. Colle, 2015: Climatology and ensemble prediction of non-convective high wind events in the New York City Metropolitan Region. *Wea. Forecasting*, **30**, 270-294.

Picca, J.C., D.M. Schultz, B. A. Colle, S. Ganetis, D.R. Novak, and M. Sienkiewicz, 2014: The value of dual-polarization radar in diagnosing the complex microphysical evolution of an intense snowband. *Bull. Meteor. Soc., 95, 1825-1834.*

The Picca et al. (2014) paper was the cover story for the December 2014 BAMS:

It was also highlighted on the AMS Front Page Blog: http://blog.ametsoc.org/news/dual-pol-radar-shedding-light-on-wintry-mix/.

Souders, M.B., B. A. Colle, E. K.-M., Chang, 2014: A description and evaluation of an automated approach for feature-based tracking of Rossby wave packets. *Mon. Wea. Rev.*, **142**, 3505-3527.

Souders, M.B., B. A. Colle, E. K.-M., Chang, 2014: The climatology and characteristics of Rossby wave packets using a feature-based tracking technique. *Mon. Wea. Rev.*, **142**, 3528-3548.

Zheng, M., K. Chang, and B.A. Colle, 2013: Ensemble sensitivity tools for assessing extratropical cyclone intensity and track predictability. *Wea. Forecasting*, **28**, 1133–1156.

Colle, B. A., K. A. Lombardo, J. Tongue, W. Goodman, and N. Vaz, 2012: Tornadoes in the New York Metropolitan Region: Climatology and multiscale analysis of two events. *Wea. Forecasting*, **27**, 1326-1348.

Chang, K. M., Zheng, M., and K. Raeder, 2012. Medium-Range Ensemble Sensitivity Analysis of Two Extreme Pacific Extratropical Cyclones. *Mon. Wea. Rev.*, **141**, 211-231.

Novak, D. and B. A. Colle, 2012: Diagnosing snowband predictability using a multi-model ensemble system. *Wea. Forecasting*, **27**, 565-585.

Colle, B. A., and M. E. Charles, 2011: Spatial distribution and evolution of extratropical cyclone errors over North America and adjacent oceans in the NCEP Global Forecast System model. *Wea. Forecasting,* **26**, 129-149.

Novak, D., and B. A. Colle, 2010: Climatology and composite analysis of mesoscale precipitation band formation in the comma head of mid-latitude cyclones., *Mon. Wea. Rev.* **138***,* 2354-2374.

The following CSTAR-related presentations in the past 12 months:

Brian Colle: "Evaluation of Cool-Season Extratropical Cyclones in a Multi-Model Ensemble for Eastern North America and the Western Atlantic Ocean" NCEP Ensemble Workshop June 2016.

Nathan Korfe and Brian Colle: "Evaluation of Cool-Season Extratropical Cyclones in a Multi-Model Ensemble for Eastern North America and the Western Atlantic Ocean" Model Evaluation Group (MEG) Seminar. At EMC, 22 January 2016.

Brian Colle: "Evaluation of Cool-Season Extratropical Cyclones in a Multi-Model Ensemble for Eastern North America and the Western Atlantic Ocean" Northeast Storms Conference, Saratoga Springs, NY 5 March 2016.

Nathan Korfe: "Evaluation of Cool-Season Extratropical Cyclones in a Multi-Model Ensemble for Eastern North America and the Western Atlantic Ocean", 23rd Conference on Numerical Weather Prediction, Chicago, IL. 29 June to 2 July 2015.

Brian Colle: "Stony Brook Fuzzying Clustering Tool", NROWXV, Albany, NY. 4-5 November 2015.

Brian Colle: "Orographic Precipitation" Invited Presentation at COMET Winter Weather Class, Boulder, CO. January 2016.

Taylor Mandelbaum: "Development of the Ensemble Spread Anomaly Tool", NROWXV, Albany, NY. 4-5 November 2015.

Taylor Mandelbaum: "Ensemble Spread Anomaly Tool for East Coast Winter Storms" Northeast Storms Conference, Saratoga Springs, NY 5 March 2016.

Stony Brook CSTAR graduates (alum)/students:
David Stark (M.S., 2012) – NWS General Forecaster at Upton, NY
Matthew Souders (M.S., 2013) –Weather Analytics, New Hampshire
Michael Layer (M.S., 2014) – Weatherworks, Hackettstown, NJ
Michael Erickson (Ph.D., 2015) – NOAA Contractor (Weather Prediction Center)
Minghua Zheng (Ph.D. -2016, starting post-doc at Scripps in January 2017)
Nathan Korfe (M.S. 2016 – Research Meteorologist at WindLogics, MN)
Taylor Mandelbaum -- current M.S. CSTAR student (will graduate May 2018)
Ryan Connelly -- current M.S. CSTAR student (will graduate May 2018)

4.5 CSTAR Group Meetings and List Serve

The Stony Brook List serve was active during 2016 hurricane season during Hermine and Matthew. There are over 50 participants on the list serve: cstar_stony_brook@infolist.nws.noaa.gov.