

Applying Fuzzy Clustering to a Multimodel Ensemble for U.S. East Coast Winter Storms: Scenario Identification and Forecast Verification

MINGHUA ZHENG, EDMUND K. M. CHANG, AND BRIAN A. COLLE

*School of Marine and Atmospheric Sciences, Stony Brook University, State University of New York,
Stony Brook, New York*

YAN LUO AND YUEJIAN ZHU

NOAA/NCEP/NWS/EMC, College Park, Maryland

(Manuscript received 8 June 2016, in final form 21 November 2016)

ABSTRACT

This article introduces a method for objectively separating and validating forecast scenarios within a large multimodel ensemble for the medium-range (3–7 day) forecasts of extratropical cyclones impacting the U.S. East Coast. The method applies fuzzy clustering to the principal components (PCs) of empirical orthogonal function (EOF) analysis on mean sea level pressure (MSLP) from a 90-member combination of the global ensembles from the National Centers for Environmental Prediction, the Canadian Meteorological Center, and the European Centre for Medium-Range Weather Forecasts. Two representative cases are presented to illustrate the applications of this method. Application to the 26–28 January 2015 event demonstrates that the forecast scenarios determined by the fuzzy clustering method are well separated and consistent in different state variables (i.e., MSLP, 500-hPa geopotential height, and total precipitation). The fuzzy clustering method and an existing ensemble sensitivity method are applied to the 26–28 December 2010 event to investigate forecast uncertainty, which demonstrates that these two methods are complementary to each other and can be used in the operations together to track the evolution of forecast uncertainty. For past cases one can define a cluster close to the analysis based on the projection of the analysis onto the PC base of clustering. This analysis group is validated using conventional validation metrics for both cases examined, and this analysis group has fewer errors than the other groups as well as the multimodel ensemble mean and individual model means.

1. Introduction

Extratropical cyclones along the U.S. East Coast are often associated with high-impact weather events (HIWs), such as heavy precipitation (Frankoski and DeGaetano 2011; Colle et al. 2013), strong winds (Booth et al. 2015), and coastal flooding (Colle et al. 2008, 2010). Considering the high population density of the eastern United States, accurate forecasts of these storms are very important to reducing human and economic loss.

The forecast skill of numerical weather prediction (NWP) models for wintertime storms over U.S. East

Coast varies with different event lead times. Some storms, such as the 24–25 January 2000 “surprise” snowstorm and the 26–28 December 2010 storm, have been relatively poorly predicted even 1–2 days in advance (Zhang et al. 2002; Zheng et al. 2013). Some previous studies have also pointed out systematic errors and biases in the forecasting models. For instance, Colle and Charles (2011) found that cyclones occurring off the U.S. East Coast are underpredicted by the Global Forecast System (GFS) model for the 72–120-h forecast during the 2002–07 cool seasons.

With the advancement in computational resources, a large set of ensemble output from multiple NWP models is available during the forecast process, such as 50 perturbed members from the European Centre for Medium-Range Weather Forecasts (ECMWF) global model (Molteni et al. 1996), 20 members from the National Centers for Environmental Prediction (NCEP) Global Ensemble Forecast System (GEFS) global

Supplemental information related to this paper is available at the Journals Online website: <http://dx.doi.org/10.1175/WAF-D-16-01112.s1>.

Corresponding author e-mail: Minghua Zheng, ming.h.zheng@gmail.com

DOI: 10.1175/WAF-D-16-01112.1

© 2017 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](#) (www.ametsoc.org/PUBSReuseLicenses).

model (Toth and Kalnay 1993), 20 members from the Canada Meteorological Centre (CMC) global model (Houtekamer et al. 1996), and 23 members from the Met Office (UKMO) global model (Bowler et al. 2008). Spaghetti plots and ensemble mean/spread are often used to display the ensemble outputs. The predictability can be discerned when the ensemble spread is large or small. However, the ensemble mean/spread often removes useful information because the method tends to filter out the different possible ensemble prediction system (EPS) scenarios. Therefore, there is a need to assess and display the ensemble uncertainty more effectively and objectively in order to consolidate the important multimodel ensemble information. In our previous work, we have applied an ensemble sensitivity tool to the operational environment (Zheng et al. 2013). Ensemble sensitivity can reveal upstream sensitive regions at earlier forecast times associated with specific forecast metrics over a verification region at a particular lead time. However, it is difficult to verify the sensitivity signals and provide forecasters the details of individual ensemble members using this sensitivity tool.

Cluster analysis has been used to examine similar forecast scenarios among ensemble members for the short range [0.5–2.5 day; Johnson et al. (2011)], medium range (Ferranti and Corti 2011), and extended range [beyond 10 days; Palmer et al. (1990)]. Brill et al. (2015) introduced a new divisive clustering algorithm to medium-range forecasting based on the one-dimensional discrete Fourier transformation. The divisive method is easy to develop and maintain and can efficiently produce plausible clusters in an operational forecasting process. One main disadvantage of this approach is that it often excludes around a dozen of the combined 70 ensemble members by not including them in any cluster. In addition, this approach reduces the geopotential data to one spatial dimension at the large scale, which may miss considerable two-dimensional information in complex weather systems. Nevertheless, the application of this method can provide meteorologically coherent pictures of outcomes not captured sharply by ensemble means or deterministic model output. To take full advantage of the ensemble product and provide consistent guidance for national centers and weather forecast offices (WFOs), the U.S. National Weather Service (NWS) has developed a National Blend of Models (NBM) by postprocessing and combining (blending) guidance from models and EPSs from multiple centers (Gilbert et al. 2016; Tew et al. 2016). To get the best possible science, it is important to document the variability among different EPSs, and to evaluate the benefit of combining models, especially in the sense of providing a sufficient variety of forecast scenarios. The

intracluster similarity and intercluster contrasts in the cluster analysis can be used to examine the consensus forecast and the contrasting possibilities, respectively.

Harr et al. (2008) introduced the fuzzy clustering method for identifying groupings of forecast scenarios of the extratropical transitioning (ET) of tropical cyclones within a collection of ensemble members. Keller et al. (2011) compared the forecast scenarios associated with 10 ET cases in The Observing System Research and Predictability Experiment (THORPEX) Interactive Global Grand Ensemble (TIGGE; Bougeault et al. 2010) data by applying fuzzy clustering analysis and found that some EPSs are confined to a few scenarios while others contribute to almost all scenarios. Other studies (e.g., Grams et al. 2011) have also shown that fuzzy clustering is a suitable diagnostic method for detecting the physical processes associated with different weather systems. In this study, the fuzzy clustering method of Harr et al. (2008) is applied to winter storm cases over the U.S. East Coast to separate forecast scenarios in an objective and efficient way and to validate ensemble forecast output, with a goal of providing guidance to forecasters to improve our understanding of scenarios as well as interpreting model biases.

This paper will address the following motivational questions:

- Can the fuzzy clustering method efficiently separate forecast scenarios associated with extratropical cyclones and thus condense the useful information from the large ensemble?
- How can one apply the method to the forecasting process for a winter storm over the East Coast?
- How can the fuzzy clustering method be used as a verification of ensemble predictions of winter storms?

This paper introduces the EOF/fuzzy clustering method to ensemble forecasting and illustrates it with two representative cases. Section 2 of this paper discusses the methodology and data used. Section 3 describes the synoptic evolution of an extratropical cyclone, which brought HIW to the U.S. East Coast on 26–28 January 2015. Section 4 will explore the 2010 Christmas blizzard case and show the linkage between fuzzy clustering analysis and the ensemble sensitivity tool. The discussions will be provided in section 5. Section 6 presents the conclusions.

2. Data and methodology

a. Data

This paper will analyze two East Coast winter storm cases: the 26–28 January 2015 blizzard (case 1) and the 26–28 December 2010 blizzard (case 2). The ensemble

data were retrieved from the TIGGE archive. Some ensemble data for the 26–28 January 2015 case study are missing from the TIGGE archive, and the data were retrieved from NCEP’s Environmental Modeling Center (EMC) archive. The three operational models considered in this study are the CMC, the ECMWF, and the NCEP EPSs. These provide us with a large ensemble of 90 members, interpolated onto a 1° latitude × 1° longitude grid, in 12-h increments. The GFS operational analysis interpolated onto a 1° latitude × 1° longitude grid is used to verify the ensemble forecasts. We have also used the ECMWF analysis for verification, and the results are very similar. The forecast lead times we examined are 3 and 6 days for case 1, and 3.5 days for case 2. The main forecasting parameters used are mean sea level pressure (MSLP) and 500-hPa geopotential height Z500. We first use the runs initialized at 1200 UTC 21 and 24 January 2015 [6 and 3 days before the verification time (VT) of 1200 UTC 27 January 2015] as an illustrative example before applying the method to examine the other cyclone case.

b. EOF analysis

Empirical orthogonal function (EOF) analysis is often used in climate studies as a technique for determining the spatial patterns that most efficiently explain the variability of a multivariate dataset (Richman 1986). Zheng et al. (2013) applied the EOF analysis to an ensemble sensitivity study to reveal the dominant spatial patterns of model uncertainty in an ensemble forecast of MSLP over prescribed regions. These EOFs are calculated across the model ensemble member dimension—as opposed to the more conventional time dimension—and the resulting modes show the dominant patterns of the difference between individual ensemble members and the model ensemble mean. The leading principal components (PCs) are the projections of the dominant EOF patterns onto the difference between each of the ensemble members and the ensemble mean. Since the leading PCs and the associated EOF patterns contain the main uncertainty information across the entire ensemble, we use them here as the basis for performing cluster analysis. In this study, each PC is normalized to have a variance of 1.

c. Fuzzy clustering analysis

Once the EOF analysis has been conducted, the first and second PCs for the 90 ensemble members are used as input into a fuzzy clustering routine (Scott and Symons 1971; Harr and Elsberry 1995; Harr et al. 2008), which is utilized to group ensemble members with similar forecast scenarios. Other than the fuzzy clustering method (partitional clustering), the hierarchical clustering method has often been used to cluster data (e.g., Johnson

et al. 2011). We have found that the clustering solutions based on the two methods are moderately consistent for the two cases (not shown). The main advantages of the fuzzy clustering method over the hierarchical clustering method are the linear complexity in computation, simplicity of implementation, and flexibility in cluster assignments (Maimon and Rokach 2005). Therefore, we choose the fuzzy clustering method to compute clusters in this study.

Following Harr et al. (2008), the main steps of the fuzzy clustering algorithm are as follows:

- 1) place a predefined number of clusters (initial guess) in the EOF PC1–PC2 phase space,
- 2) assign each ensemble member represented by the pair of PCs to the nearest group center,
- 3) compute new centers by minimizing an objective function that represents the distance from each point to each new cluster,
- 4) reexamine each point relative to the updated cluster centers, and
- 5) repeat steps 2–4. If no points can be reassigned because they lie closer to another center, the iterations stop.

Each ensemble member has a weighted value that identifies its relative strength of membership to each cluster. For a point *k*, the weight associated with the *i*th cluster is defined by

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

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 coefficient *q* determines the level of cluster fuzziness. A large *q* results in smaller differences between memberships *w*_{*i,k*} and, hence, fuzzier clusters. In the limit *q* = 1, the memberships converge to 0 or 1, which implies a clear partitioning. The coefficient *q* is set to 2 in this study, as suggested by previous studies (Ahmed et al. 2002). A total of *C* clusters can be calculated using the above equation and procedure. An ensemble member *k* is assigned to the cluster *i* (*i* = 1, . . . , *C*) for which *w*_{*i,k*} has the largest value.

In this study, fuzzy clustering was applied to the PCs corresponding to the leading two EOF patterns. Conceptually, there should be an optimal number of population clusters associated with different synoptic-scale patterns. However, as pointed out by Harr et al. (2008), it is often very difficult to determine this number objectively. Keller et al. (2011) tested a range of the cluster

number C from two to eight and found that six was a suitable cluster number for their study. We have also tested two- to eight-cluster solutions and found that six and five clusters tend to be optimal for cases 1 and 2, respectively. The optimal number of clusters is the stable solution with the highest adjusted Rand index (Yeung and Ruzzo 2001) based on 100 clustering results using random seeding points. The results for two- to eight-cluster solutions of each case using an adjusted Rand index are summarized in Table S1 (see supplementary material to this paper online). When two or more cluster results are stable, the largest number of clusters is chosen. One reason is that we would like the clustering procedure to produce a group that can represent the ensemble mean, which is often hypothesized to be a good estimation of the truth (Du and Zhou 2011). Our experience suggests that a group clustered around the ensemble mean is more likely for larger numbers of clusters. Note that the ensemble mean (the origin on a PC1–PC2 coordinate) is used in the clustering process.

When examining past cases, after we group forecast ensemble members based on their EOF PCs at VT, the difference between the analysis field and the ensemble mean at VT can be projected onto the leading EOF patterns and hence occupies one point on the EOF PC1–PC2 space, just like one extra ensemble member, and this can be used to verify the scenarios in the ensemble forecasts (see section 5).

3. 26–28 January 2015 event

The mid-Atlantic to Northeast major winter storm during 26–28 January 2015 is used to illustrate the application of fuzzy clustering. This storm impacted the northeastern United States, resulting in snowfall accumulations of from 30 cm (12 in.) to 91 cm (36 in.) over the central part of Massachusetts, and blizzard conditions were prevalent from Long Island to southern and eastern New England (Winkler 2015).

a. Synoptic overview

At 1200 UTC 25 January, a surface low pressure system was located near the Iowa and Missouri border with an associated upper-level short-wave trough approaching the Ohio valley (Fig. 1a). The upper-level cyclonic potential vorticity (PV) anomaly (2-PVU dynamic tropopause < 305 K, where $1 \text{ PVU} = 10^{-6} \text{ K kg}^{-1} \text{ m}^2 \text{ s}^{-1}$) was located over the central part of the Great Plains. During the following 12 h, the surface low progressed eastward along the Kentucky–Tennessee border and weakened as it approached a preexisting ridge, as indicated by warm potential temperatures on the tropopause (Fig. 1b). At 1200 UTC 26 January (Fig. 1c), the

short-wave trough and the associated surface low moved eastward across the central Appalachians. A new surface cyclone formed off the east coast of North Carolina at 1800 UTC 26 January, between the left exit of an upper-level jet stream over the southeastern United States and the right entrance of a downstream jet east of New England (not shown).

Between 1200 UTC 26 January and 0000 UTC 27 January, the upper-level trough became negatively tilted as it approached the East Coast, and the surface low intensified just off the mid-Atlantic coast (Fig. 1d). The upper-level PV maximum was over the northwest of the trough axis; hence, there was positive PV advection near the trough axis off the mid-Atlantic coast, as well as the corresponding surface low pressure system. During the following 12 h (Fig. 1e), a closed 500-hPa low developed and strengthened off of eastern Long Island. The surface low continued tracking north-northeastward to the western Atlantic east of Long Island, and deepened rapidly from 992 hPa at 0000 UTC 27 January to 980 hPa at 1200 UTC 27 January. Over the 24-h period between 1200 UTC 26 January and 1200 UTC 27 January, heavy snow fell over southeastern New England with maximum precipitation near Boston, Massachusetts (Fig. 2a).

At 0000 UTC 28 January (Fig. 1f), the 500-hPa low was situated over Cape Cod and became neutrally tilted while the surface cyclone was centered east of Boston with a minimum pressure of 985 hPa (Fig. 1f). An occluded front extended northeastward from the surface low into the Atlantic Ocean (not shown). The surface low weakened as the upper-level low began to fill through 1200 UTC 28 January and the snowfall intensity was reduced to light and scattered over the Northeast (not shown).

b. Ensemble forecasts at different lead times

Figure 3 shows the ensemble means from each of the models, the multimodel mean, and the analysis verifying at 1200 UTC 27 January 2015 for different lead times. Between days -6 and -4 (Figs. 3a–c), all three EPSs forecasted a mean cyclone that was east or southeast of the analyzed cyclone. Between days -5 and -4 (Figs. 3b,c), the NCEP and ECMWF mean cyclone shifted west-southwestward. This trend continued between days -4 and -3 (Figs. 3c,d), and during that 24 h the CMC-forecasted cyclone also shifted westward. From days -3 to -1 (Figs. 3d–f), the ECMWF mean cyclone was the deepest and farthest west, while the CMC mean cyclone was the weakest and farthest east, with the NCEP mean cyclone located between the two. Overall, for the short-range forecasts the multimodel ensemble mean indicated a closer-to-shore cyclone than the analysis, which led the forecasters at the NWS to

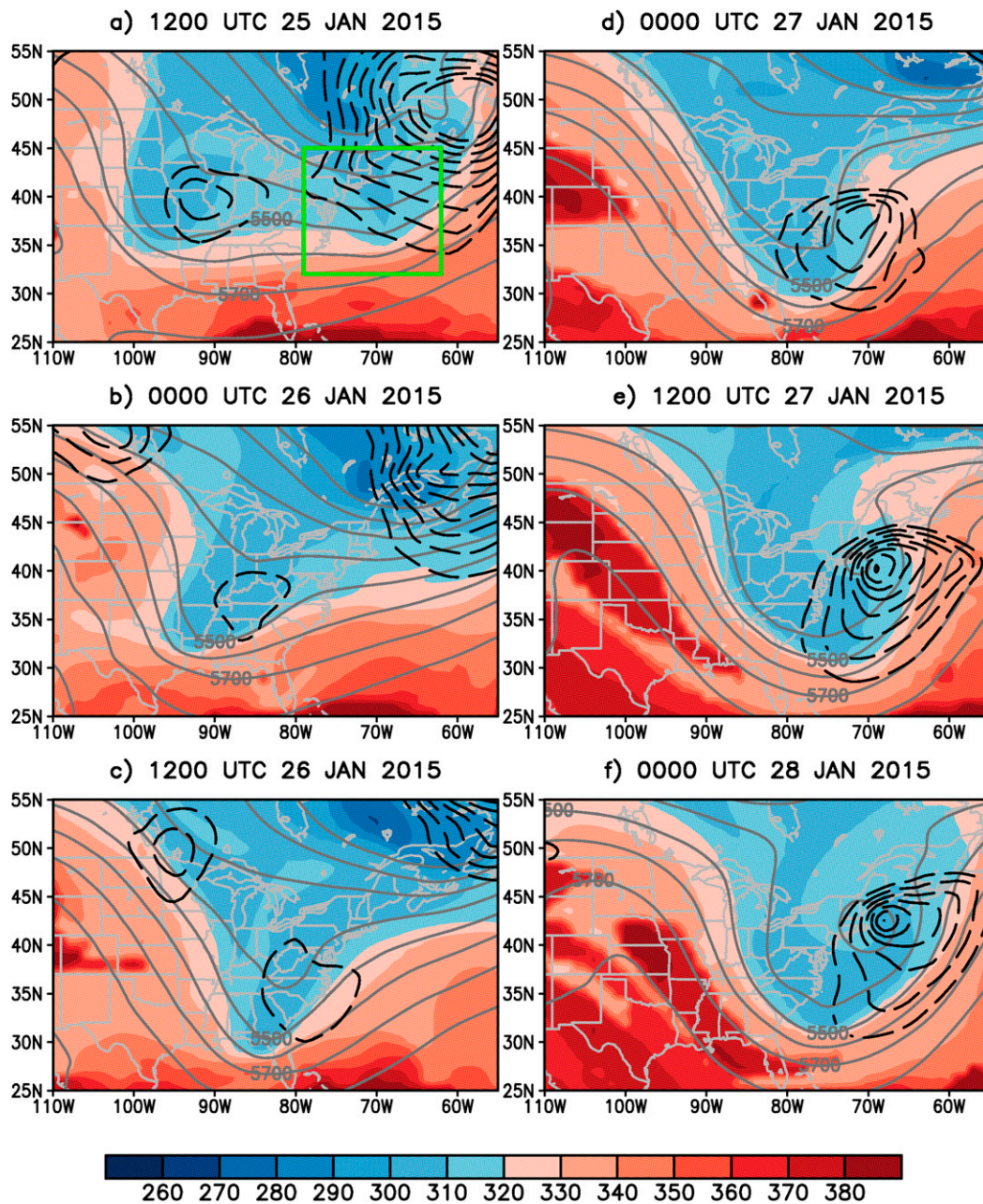


FIG. 1. (a)–(f) Geopotential height (gray solid contours, 100-m interval with the 5700- and 5500-m lines labeled) at 500-hPa level, MSLP (black dashed contour starting at 1008 hPa and decreasing in 4-hPa intervals), and potential temperature at 2-PVU dynamical tropopause (shading; K) from 1200 UTC 25 Jan to 0000 UTC 28 Jan 2015. The green solid box in (a) represents the verification region at valid time (1200 UTC 27 Jan 2015).

predict a major snowstorm and issue a blizzard warning for parts of seven states including New York City (NYC; Fig. 2), which turned out to receive far less snow than expected (Fanning 2015; Flegenhimer 2015).

In the following subsections, we will focus on the ensemble forecasts at lead times of 6 and 3 days to examine the forecast scenarios at medium range associated with this blizzard case using EOF and fuzzy clustering analyses.

c. 6-day ensemble forecasts

1) VARIABILITY CAPTURED BY LEADING EOF PATTERNS

For the 6-day ensemble forecasts (Figs. 4a,b) initialized at 1200 UTC 21 January, the multimodel ensemble mean cyclone was centered at 39.0°N, 66.0°W, which was ~280 km to the southeast of the analyzed cyclone.

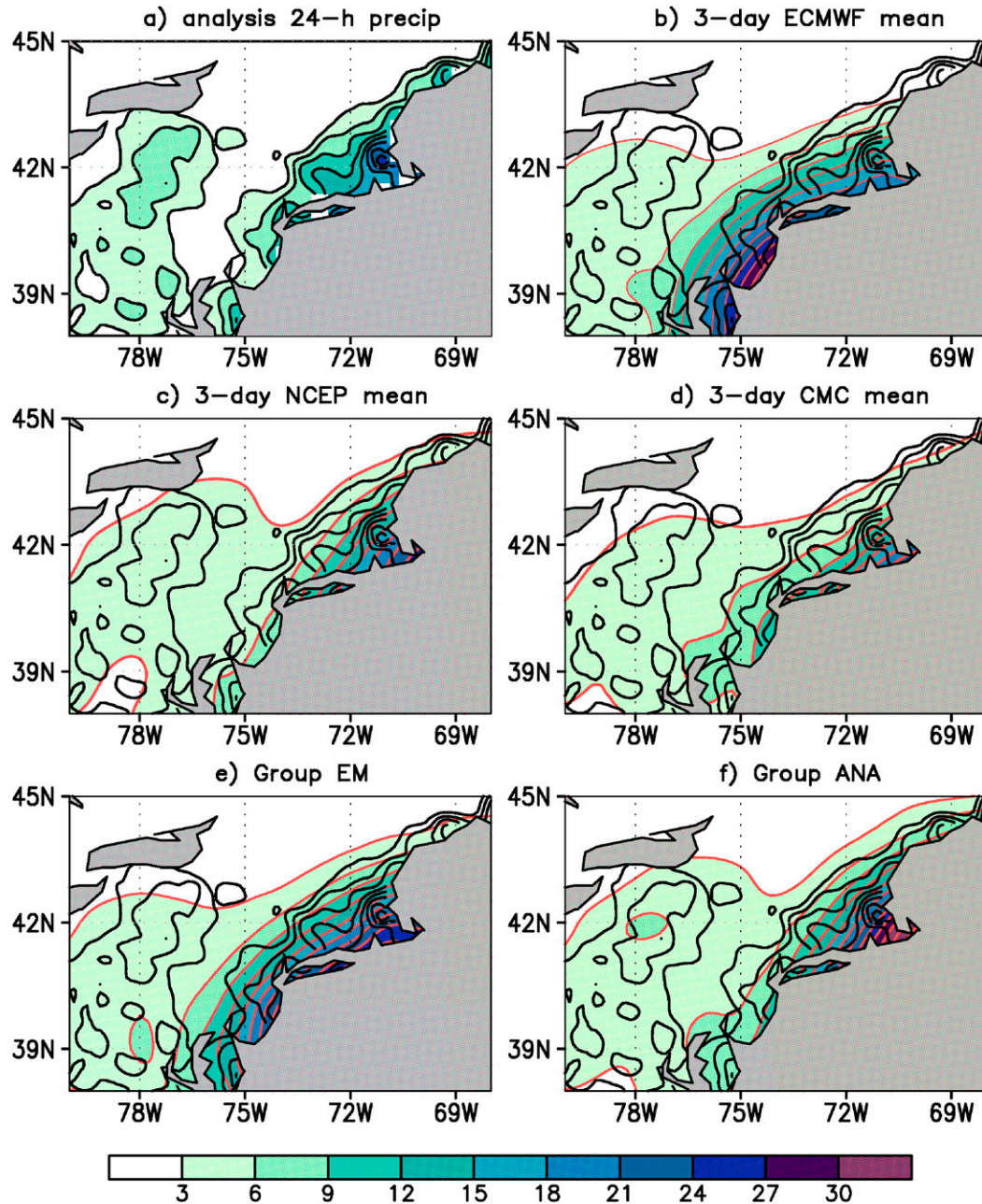


FIG. 2. (a) The CPC 24-h precipitation (shading and black contour; mm) and (b)–(d) 3-day ensemble mean forecasts (shading and red contour; mm) from three models. The means (shading and red contour; mm) from (e) group EM and (f) group ANA (or group 6) for 3-day multimodel ensemble. Time period is from 1200 UTC 26 Jan to 1200 UTC 27 Jan 2015. The black contours in (b)–(f) are the same as the contours in (a) and represent observed precipitation.

Forecast uncertainty was mainly over the north of the ensemble mean cyclone. Figure 4b shows the spaghetti plot for the 1008-hPa MSLP contour. The 1008-hPa contours from the multimodel ensemble were spread out along the east coast of North America and the adjacent Atlantic Ocean. Therefore, it could be very

difficult for forecasters to extract the possible scenarios from the spaghetti diagram plot. Useful tools, such as the fuzzy clustering approach, are needed to distill the information.

Figures 4c and 4d show the leading two EOF patterns for this 6-day MSLP forecast, which explain 62.9%

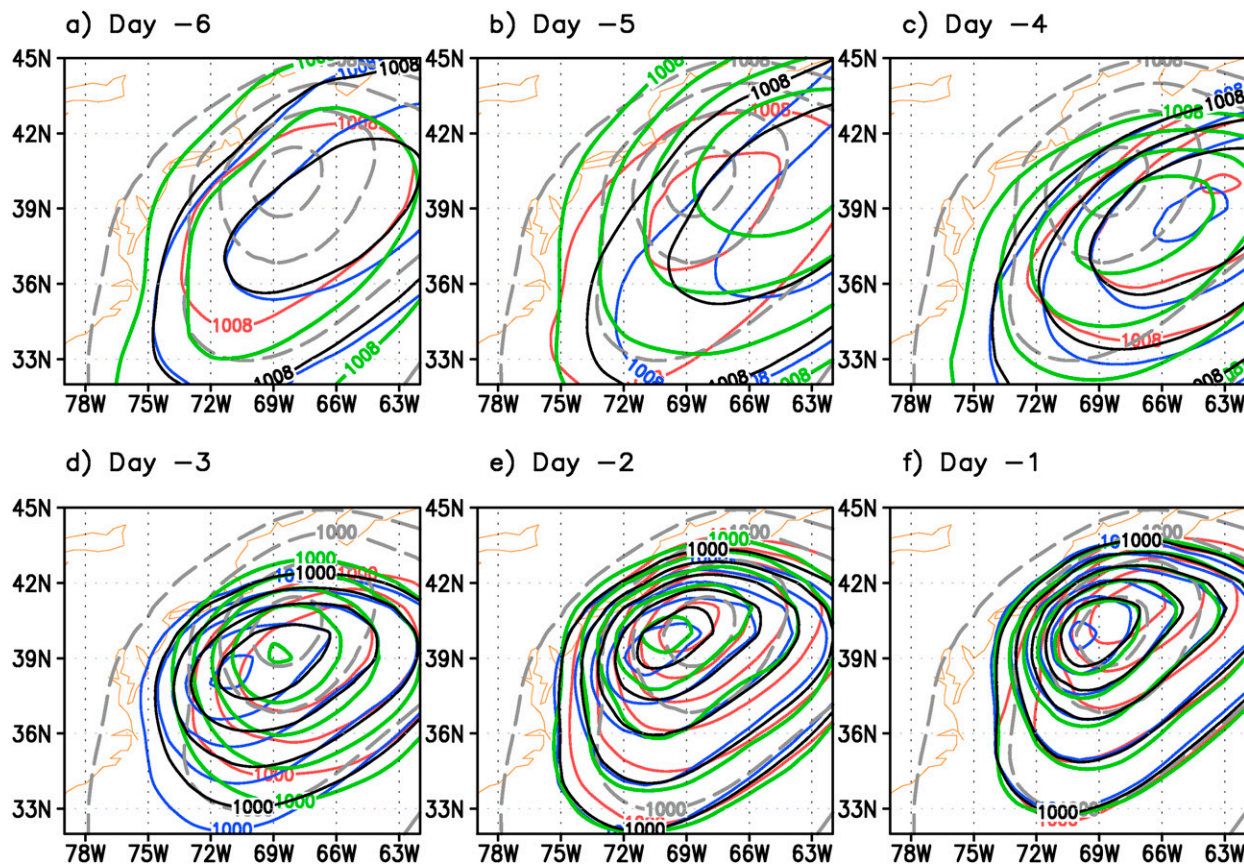


FIG. 3. (a)–(f) Ensemble mean forecasts of MSLP from a lead time of 6 days to 1 day. Red, green, and blue represent the CMC, NCEP, and ECMWF model means, respectively. Black and gray are the multimodel mean and analysis, respectively. The contours for the analysis range from 984 to 1008 hPa with a contour level of 8 hPa. Contours for the ensemble means in (a)–(c) range from 1008 to 1000 hPa and in (d)–(f) they range from 1000 to 980 hPa with a contour level of 4 hPa.

and 14.9% of the variance over the verification region, respectively. These two patterns represent the primary variability patterns of 6-day MSLP forecasts among the three EPSs. The first EOF (EOF1; Fig. 4c) is a monopole pattern centered ~650 km north of the ensemble mean position of the surface cyclone. Therefore, a positive (negative) EOF1 pattern represents a deeper (weaker) storm with a northward (southward) shift compared with the ensemble-mean cyclone on day 6. Meanwhile, EOF2 (Fig. 4d) is an asymmetric dipole pattern that has a negative center east-northeast of the ensemble mean cyclone position and a positive center to the northwest. Thus, positive (negative) EOF2 represents the deepening (weakening) and eastward (westward) shifting of the cyclone. Figure 4 shows that large forecast uncertainties exist over the north of the ensemble-mean cyclone center, which are associated with the disagreement among models in forecasting both the amplitude and proper location of the cyclone.

2) FORECAST SCENARIOS DETERMINED BY THE FUZZY CLUSTERING METHOD

Fuzzy clustering is applied to the PCs corresponding to the leading two EOF patterns. Figure 5a shows the partitions of the 90 ensemble members into the six clusters as determined by the clustering method. The group including the origin is defined as group EM. In this case, while group EM is not centered at the ensemble mean, its group center is still closest to the ensemble mean compared with the other groups. The definition of group EM has its practical applications: the members of this group are often closer to the ensemble mean. Groups 2 and 4 are centered on negative PC1 and represent weaker cyclone scenarios, with the former being stronger and more south-southwestward (large negative PC2) than the latter (slightly positive PC2). In contrast, group 6 consists of 5 members with the largest PC1 values in all 90 members, representing the deepest cyclone scenario. Both groups 3 and 5 are centered on

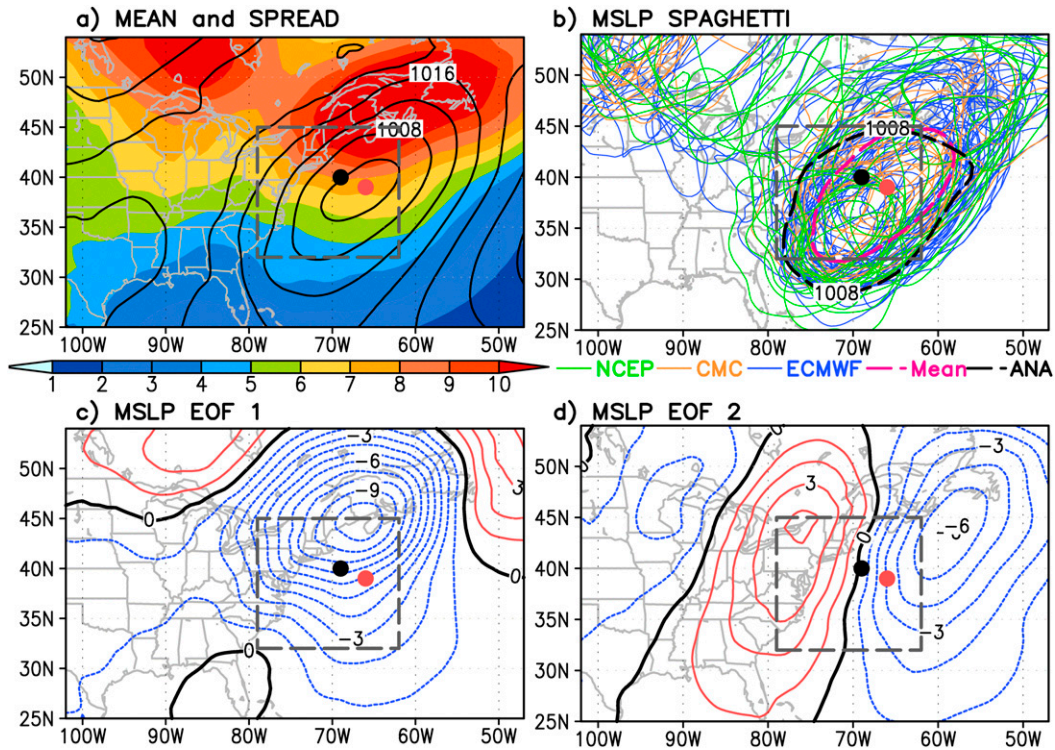


FIG. 4. (a) MSLP ensemble mean (contours; hPa) and spread (shading; hPa). (b) Spaghetti plots of 1008-hPa contour for 90 multimodel ensemble members (blue for ECMWF members, green for NCEP members, and orange for CMC members) with the dashed magenta lines and black lines showing the ensemble mean and analyzed contour line. (c) The regressed pattern of MSLP corresponding to EOF PC1 for a 6-day forecast (contours; hPa). (d) As in (c), but for the EOF2 pattern (contours; hPa). The VT is 1200 UTC 27 Jan 2015, and the IT is 1200 UTC 21 Jan 2015. The black (red) dot in each panel denotes the analyzed (ensemble mean) position of the surface cyclone at VT. The dashed gray box in all panels is the verification region for calculating EOFs.

fairly positive PC1 and represent deeper cyclone scenarios than the ensemble mean, with the former being more west-southwestward (small negative PC2) than the latter (large positive PC2).

3) SYNOPTIC CHARACTERISTICS OF DIFFERENT CLUSTERS

Figure 6 shows a summary of the six groups' means partitioned based on EOF PCs on day 6. The area/position of each closed circle in Figs. 6a–d (the 1008- or 1004-hPa MSLP contour) represents the strength/location of the surface cyclone for the mean of each group. Meanwhile, the 5100-, 5400-, and 5700-m Z500 contours (Figs. 6e–h) represent the mean upper-level features for each group. On day 3 (Figs. 6a,e), the surface cyclone forecasts in all groups seem to be similar over southwest Canada, but the cyclone and the associated upper-level trough for group 4 seem to be weaker than those of the other four groups. On days 4 and 5 (Figs. 6b,c,f,g), the surface cyclones as well as the upper-level troughs in groups 2 and 4 are weaker than those of group EM. Groups 3, 5, and 6, on the other hand, have deeper

cyclones than group EM has. Among the six groups, group 6 shows the strongest and most northwestward cyclone and upper-level trough system.

On day 6 (VT; Figs. 6d,h), the surface cyclones in groups 3, 5, and 6 continue to be much deeper than that of group EM. In contrast, cyclones in groups 2 and 4 are much weaker than that of group EM. The group 5 mean forecasts the most east-northeastward surface cyclone and the most eastward trough among the six groups. Group 6 has the deepest cyclone and upper-level trough while the group 4 mean forecasts the weakest system among the six groups (Figs. 6d,h). The synoptic features represented by both surface and upper-level variables match the interpretations of clusters in the previous subsection.

Figure 7 shows the spaghetti plot of 1004-hPa MSLP contour lines for the six groups. At this rather long lead time, there is quite a bit of scatter within each group. However, consistent forecasts are clearly clustered in the same group, especially for deep cyclone scenarios, such as group 6. By contrast, weak cyclones can be seen in groups 2 and 4. Note that a substantial number of

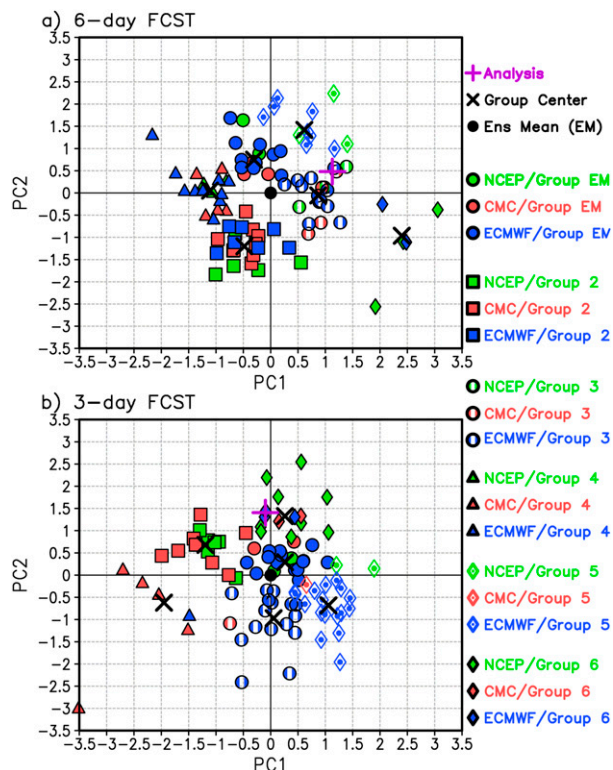


FIG. 5. The six clusters divided using the fuzzy clustering method on PC1–PC2 space from three models for (a) 6- and (b) 3-day forecasts. The VT is 1200 UTC 27 Jan 2015, and the ITs are 1200 UTC 21 Jan in (a) and 1200 UTC 24 Jan 2015 in (b).

members in group 4 do not even show the 1004-hPa contour. Note that the cluster separations within the verification regions are more distinct than those within the plotting domain.

d. 3-day ensemble forecasts

1) VARIABILITY CAPTURED BY LEADING EOF PATTERNS

The day 3 ensemble mean forecast (Figs. 8a,b) predicts a cyclone centered at 39.0°N, 69.5°W with a minimum pressure of 985 hPa, ~120 km to the south-southwest of the analyzed cyclone. The maximum ensemble variance is ~400 km to the west-southwest of the ensemble-mean cyclone center (Fig. 8a). The spaghetti plot of the 996-hPa contour lines (Fig. 8b) suggests that there is large forecast uncertainty regarding the position of the cyclone. The two leading EOF patterns for this 3-day MSLP forecast explain 42.9% and 28.7% of the variance, respectively. The positive (negative) EOF1 (Fig. 8c) pattern is a monopole centered ~260 km west of the ensemble mean cyclone, representing a deeper (weaker), as well as a more westward (eastward) shifted, cyclone. The positive (negative) EOF2 (Fig. 8d) pattern

is a dipole, representing the northeastward (southwestward) shift of the cyclone, indicating that there is a subset of ensemble members shifted to the northeast (as well as southwest) of the ensemble mean cyclone at VT.

2) FORECAST SCENARIOS DETERMINED BY THE FUZZY CLUSTERING METHOD

Figure 5b shows the six groups based on the fuzzy clustering method. Group EM is centered close to the ensemble mean. Group 2 represents a negative EOF1 pattern and weakly positive EOF2, with the cyclone in this group anticipated to be weaker and more eastward than the ensemble mean. Group 3 represents negative EOF2 characteristics: the southwestward shift of the cyclone position relative to the ensemble mean (Fig. 8d). Group 4 has mainly a strong negative EOF1, suggesting a cyclone much weaker and more eastward relative to the ensemble mean. One of the CMC members in this group is at (−3.5, −3.0), seemingly an outlier with respect to the rest of the ensemble. Group 5 indicates a combination of positive EOF1 and negative EOF2 anomalies. Therefore, the cyclone in this group should be more west-southwestward (more onshore) and deeper than the ensemble mean. Group 6 mainly represents a positive EOF2 pattern, indicating a northeastward shift of the cyclone position relative to the ensemble mean.

To sum up, groups 6 and 3 represent largest positive and negative EOF2 patterns, corresponding to more northeastward and southwestward shifts, respectively. Meanwhile, groups 4 and 5 represent the largest negative and positive EOF1 patterns, corresponding to weaker/offshore and deeper/onshore cyclone scenarios, respectively.

3) SYNOPTIC CHARACTERISTICS OF DIFFERENT CLUSTERS

Figure 9 shows the group means of the six clusters (partitioned based on EOF PCs on day 3) from day 1 to day 3 using the 1008-hPa MSLP and 5100-, 5400-, and 5700-m Z500 contour lines. On day 1 (Figs. 9a,d), the surface cyclones in all groups seem to be similar, but the cyclone for group 6 is more intense than that of the other five groups. The Z500 short-wave trough shows more separations among the groups, with that of group 6 being the strongest and most southeastward one. On day 2 (Figs. 9b,e), the surface cyclone as well as the upper-level trough in group 6 are still the most eastward among the six groups. Group 5, on the other hand, has the cyclone being the most westward. Group 4 shows the weakest cyclone and upper-level trough, with the ensemble mean pressure higher than 1008 hPa and thus absent from Fig. 9b.

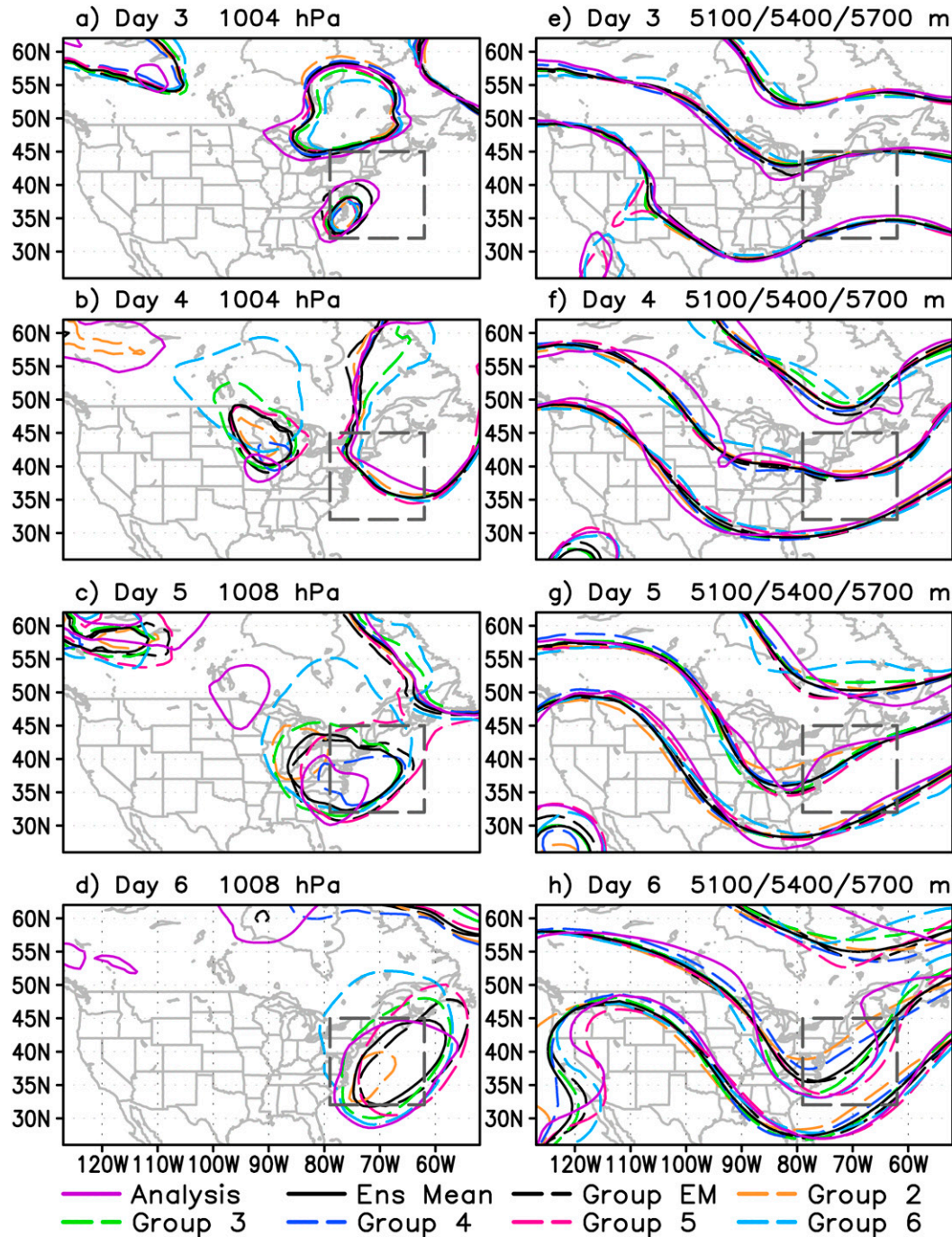


FIG. 6. (a)–(d) A summary of six cluster means using 1008- or 1004-hPa MSLP contour lines and (e)–(h) 5100-, 5400-, and 5700-m Z500 contour lines for a 6-day forecast initialized at 1200 UTC 21 Jan 2015. The dashed gray box in all panels is the verification region.

On day 3 (VT; Figs. 9c,f), group 6 has both the surface cyclone and its associated upper-level trough being ~ 120 km more northeastward than that of group EM, which is consistent with the earlier forecast time steps. Group 5 has the cyclone deeper than that of group EM and ~ 320 km more west-southwestward (Fig. 9c). Its

associated trough is the deepest among the six groups and extends more west-southwestward than the ensemble mean (Fig. 9f). By contrast, both the surface cyclone and the upper-level trough in groups 2 and 4 (Figs. 9c,f) are weaker than that of group EM. The cyclone of group 4 continues to be the weakest, with the

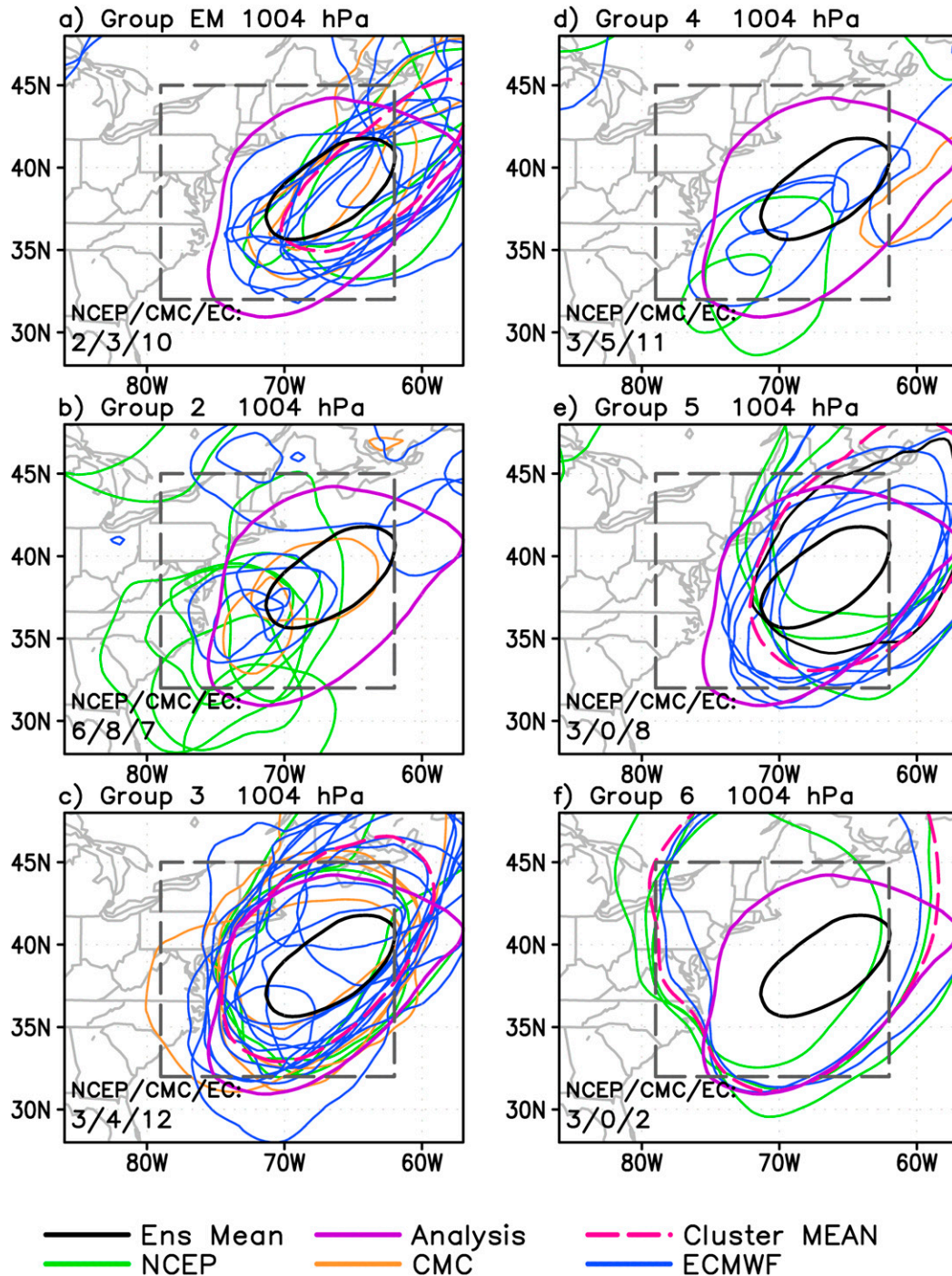


FIG. 7. (a)–(e) Spaghetti diagram plot for six groups using the 1004-hPa MSLP contour line. Red, green, and blue are contour lines from the CMC, NCEP, and ECMWF members, respectively. The dashed magenta line is for the group mean. The black and purple lines are the multimodel mean and the analysis, respectively. The VT is 1200 UTC 27 Jan 2015, and the IT is 1200 UTC 21 Jan 2015. The dashed gray box in all panels is the verification region.

center ~320 km more southeastward than the ensemble mean. The cyclone of group 2 is weaker by 3 hPa and ~170 km more southwestward than the ensemble mean.

Figure 10 shows the spaghetti plot of 1000-hPa MSLP contour lines for the six groups. ECMWF members are mainly contributing to groups EM, 3, and 5 (Figs. 10a,c,e), forecasting a more close-to-shore cyclone. However,

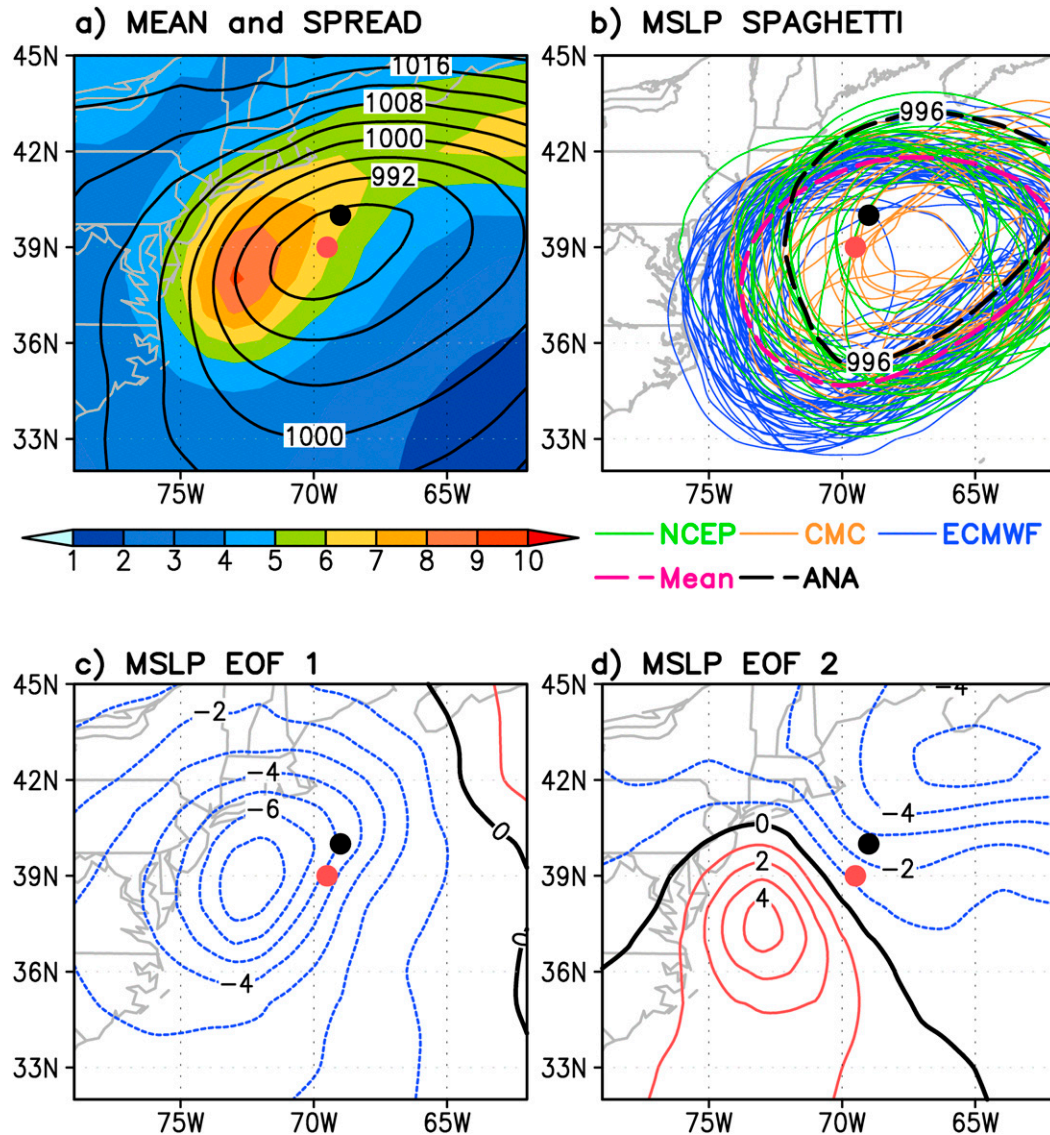


FIG. 8. (a) MSLP ensemble mean (contours; hPa) and spread (shading; hPa), (b) spaghetti plots of the 996-hPa contour for 90 multimodel ensemble members (blue are for ECMWF members, green are for NCEP members, and orange are for CMC members) with the dashed magenta lines and black lines representing the ensemble mean and analyzed contour line. Patterns shown for (c) EOF1 MSLP (contours; hPa) and (d) EOF2 MSLP (contours; hPa). The VT is 1200 UTC 27 Jan 2015, and the IT is 1200 UTC 24 Jan 2015. The black (red) dot in each panel denotes the analyzed (ensemble mean) position of the surface cyclone at VT.

three ECMWF members also contribute to group 6, together with nine NCEP and two CMC members, forecasting a deeper and more east-northeastward scenario. Thirteen CMC members contribute to groups 2 and 4, together with six NCEP and one ECMWF members, forecasting the weakest and more offshore cyclone (Figs. 10b,d). Spaghetti plots (Figs. 7 and 10) confirm that the clustering method is able to find members with similar forecast scenarios not only from the same EPS but also from different models. Figures 2b–d show that the 24-h total precipitation in the ECMWF mean is closer to shore

than in the other two models, with the heaviest precipitation regions including New York City and New Jersey. Figures 2e and 2f compares the mean precipitation for group EM and group 6, suggesting the heavy precipitation in group 6 is more northeastward and closer to the analysis than group EM. Figure 2 again demonstrates the consistency of the cluster scenarios for important weather elements. We have also investigated the clustering results for other variables including 24-h precipitation, moisture flux, and Z500 (not shown). The adjusted Rand index suggests that the

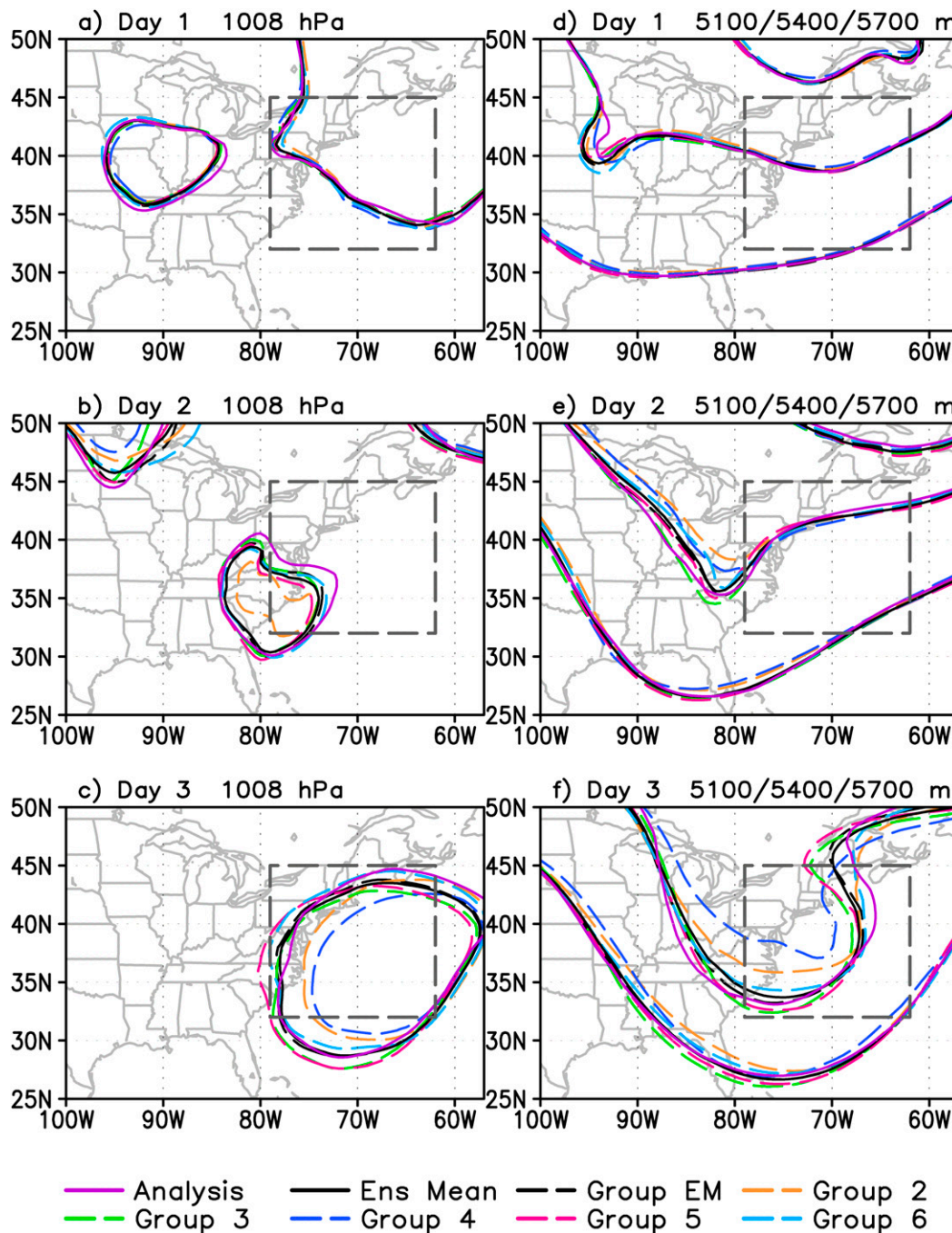


FIG. 9. (a)–(c) A summary of six cluster means using the 1008-hPa MSLP contour line and (d)–(f) 5100-, 5400-, and 5700-m Z500 contour lines for the 3-day forecast. The VT is 1200 UTC 27 Jan 2015, and the IT is 1200 UTC 24 Jan 2015. The dashed gray boxes in all panels are for the verification region.

clustering results for the precipitation and moisture flux are more consistent with the results using MSLP than those using Z500, indicating that the forecast uncertainties for the 3-day MSLP forecasts in this case could be more closely associated with the moisture process than the upper-level flow.

4. Linking clustering analysis to ensemble sensitivity

Zheng et al. (2013) investigated a blizzard case during 26–28 December 2010 and examined the forecast uncertainty associated with it using the ensemble sensitivity method. Since the ensemble sensitivity approach

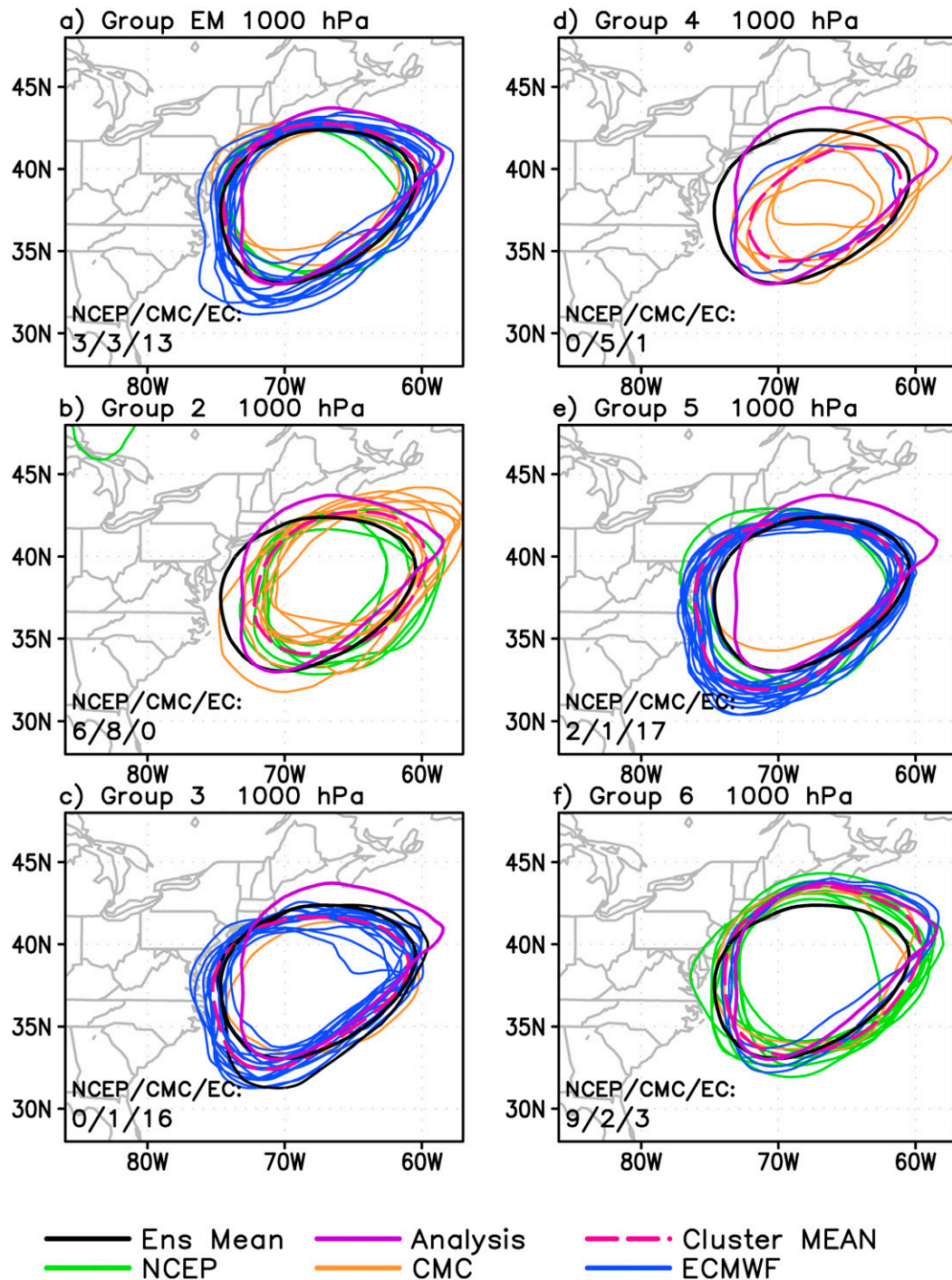


FIG. 10. (a)–(e) Spaghetti plot for six groups using the 1000-hPa MSLP contour line. Red, green, and blue contour lines are from CMC, NCEP, and ECMWF members, respectively. The dashed magenta line is for the group mean. The black and purple solid lines are the multimodel mean and the analysis, respectively, in all panels. The VT is 1200 UTC 27 Jan 2015, and the IT is 1200 UTC 24 Jan 2015.

also employs EOF PCs as a forecast metric and has been applied to the operational ensemble forecasting, here we want to relate these two methods and explore the linkage between them in ensemble forecast uncertainty studies.

a. A brief overview of the 2010 December blizzard and forecast uncertainty

On 26–28 December 2010, a powerful winter storm impacted the northeast United States, resulting in blizzard conditions across the region, including the New York City metropolitan area, New Jersey, and portions of New England (Zheng et al. 2013). The ECMWF operational ensemble forecasts indicated a potential U.S. East Coast snowstorm a week in advance (not shown), but the ensemble mean suggested that the storm would stay well to the east of the East Coast at the VT (1200 UTC 27 December; see Figs. 1a and 1e in Zheng et al. 2013). However, during the subsequent 3 days, the ensemble-mean cyclone track shifted back and forth. By 0000 UTC 24 December, the ECMWF ensemble mean suggested a solution for the cyclone center ~550 km to the east of the coast (see Figs. 1c and 1g in Zheng et al. 2013). Meanwhile, between 0000 and 1200 UTC 24 December, the operational NCEP ensemble mean correctly trended ~700 km farther northwestward, indicating the potential for heavy snowfall from New York City to Washington, D.C. (not shown). Finally, at 0000 UTC 25 December, the operational models converged on the storm track and indicated the New York City–Boston corridor was in line with the potential for heavy snow and strong winds associated with a rapidly deepening surface low. This consensus for a snowstorm event emerged only 36–48 h before the onset (1200 UTC 26 December–0000 UTC 27 December) of the heaviest snows.

In the following subsections, 1200 UTC 27 December 2010 is chosen as the VT. The 3.5-day ensemble run initialized at 0000 UTC 24 December 2010 based on 90 multimodel members is selected to compare the clustering analysis and the ensemble sensitivity tool in assessing the evolution of forecast uncertainty.

b. The forecast groups partitioned by the fuzzy clustering for 3.5-day forecast

For the 3.5-day MSLP ensemble forecasts (Fig. 11a), the multimodel ensemble mean cyclone is centered at 42.0°N, 64.0°W, and the maximum ensemble variance is over the west-southwest of the ensemble mean cyclone. Figures 11b and 11c show the leading two EOF patterns for this 3.5-day MSLP forecast, which explain 53.8% and 21.4% of the variance, respectively. Positive (negative) EOF1 corresponds to a weaker (stronger) cyclone shifted eastward (westward) of the ensemble-mean cyclone. Meanwhile,

positive (negative) EOF2 represents more of a slightly weaker (stronger) and southwest- (northeast) shifted cyclone.

For this case, the five-cluster solution is found to be the optimal choice. Figure 11d shows the partitioning of the 90 ensemble members into five clusters. Group EM is centered close to the ensemble mean. Group 2 represents positive EOF1 and weak positive EOF2 characteristics; hence, there is a weaker and more east-southeastward-shifted cyclone relative to the ensemble mean. Group 3 has negative EOF1 and slightly positive EOF2 characteristics, and is expected to be more west-southwestward and stronger than the ensemble mean. Group 4 has mainly a small positive EOF1 anomaly and a small negative EOF2 anomaly, which suggests a slightly weaker and more east-northeastward cyclone relative to the ensemble mean. Group 5 indicates a combination of negative EOF1 and negative EOF2 anomalies. Therefore, the cyclone in this group on average should be north-northwestward and deeper than the ensemble mean. Figure 11d shows that CMC forecasts mainly contribute to groups 2 and 4; NCEP forecasts to groups EM, 2, and 4; and ECMWF forecasts to groups EM, 3, and 5.

c. Time evolution of different groups and the ensemble sensitivity

Ensemble sensitivity has been applied to explore the ensemble relation between a forecast metric **J** and a state vector **X**. Zheng et al. (2013) showed that when using the EOF PCs as a forecast metric, the ensemble sensitivity reduces to the correlation between the state vector **X** and **J** since EOF PCs are normalized to the unit variance. Therefore, the ensemble sensitivity for an ensemble with *M* members is expressed as

$$\begin{aligned} \text{Ensemble Sensitivity} &= \frac{\text{cov}(\mathbf{J}, \mathbf{X}_i)}{\sqrt{\text{var}(\mathbf{X}_i)}} = \frac{\text{cov}(\mathbf{J}, \mathbf{X}_i)}{\sqrt{\text{var}(\mathbf{X}_i)}\sqrt{\text{var}(\mathbf{J})}} = \text{Correlation}, \end{aligned} \tag{2}$$

where **J** and **X_i** are the 1 × *M* ensemble estimates of the forecast metric **J** and the *i*th state vector **X**, respectively; cov denotes the covariance between **J** and **X_i** across the ensemble; and var is the variance.

To study the relation between each of these two forecast uncertainty patterns on day 3.5 (Figs. 11b,c) and the upstream conditions at earlier forecast times, the PC of each EOF pattern was used as the forecast metric, and the sensitivity in Eq. (2) was calculated using the 500-hPa geopotential height as the state vector at different

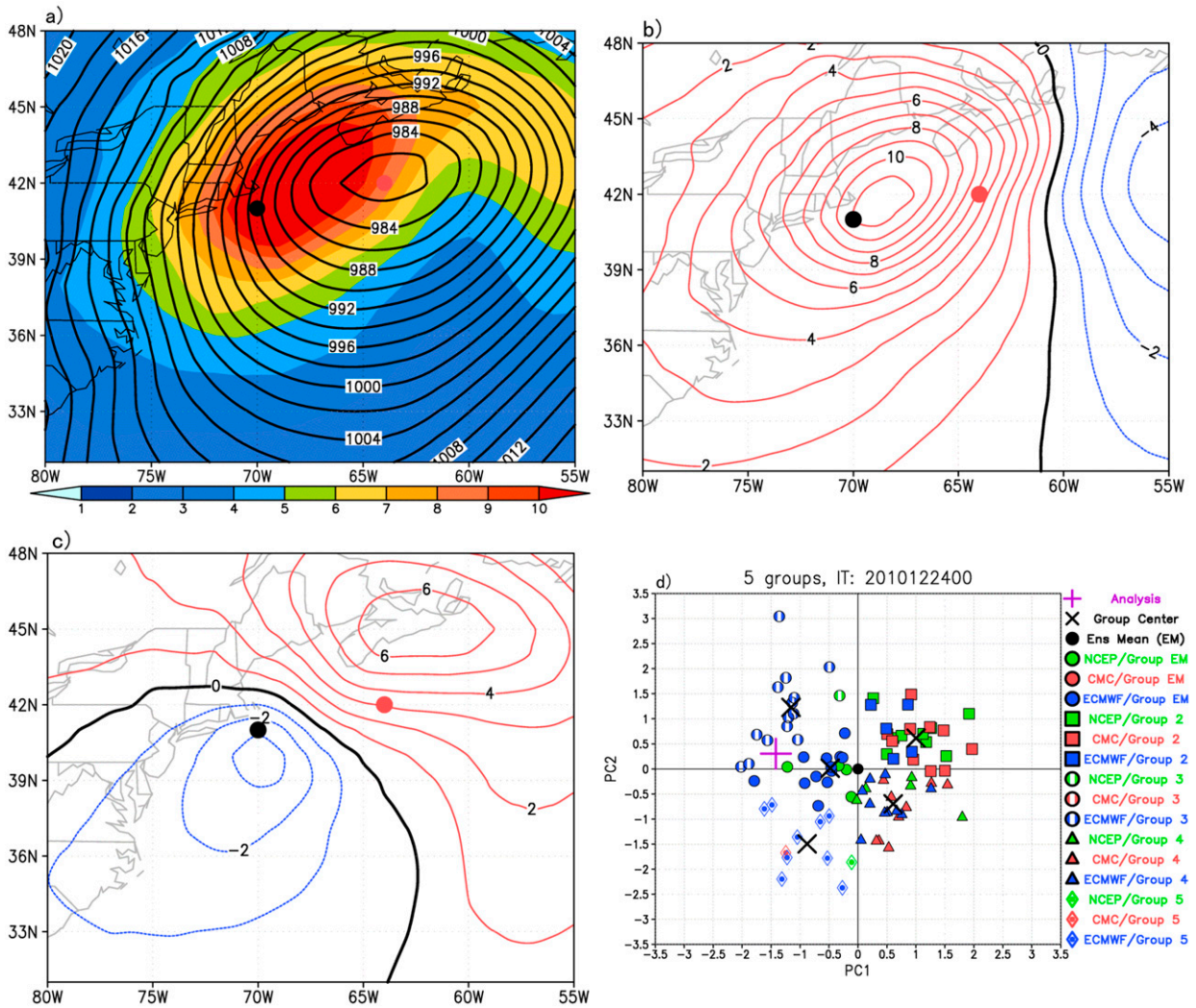


FIG. 11. (a) Mean and spread, as in Fig. 8a, but for the 3.5-day forecast initialized at 0000 UTC 24 Dec 2010. (b) EOF1 and (c) EOF2 MSLP patterns, as in Figs 8c and 8d, but for the 3.5-day forecast initialized at 0000 UTC 24 Dec 2010. (d) Five-cluster scatterplots, as in Fig. 5a, but for the 3.5-day forecast initialized at 0000 UTC 24 Dec 2010.

forecast times. Positive (negative) sensitivity indicates areas where an increase (decrease) in geopotential height is associated with an increase in the PCs and therefore an enhancement of each of the two EOF patterns.

Figure 12 compares the time evolution of the ensemble sensitivity of EOF PC1 to Z500 and the group means of Z500 for the five clusters determined by the cluster analysis at VT (0h). At a lead time of 60 h, or the day 1 forecast (Fig. 12a), there are two sensitivity maxima (>0.5) over the continental United States, one associated with an upper-level short-wave trough over Texas and the other associated with the eastern side of an amplified ridge spreading from the western United States to western Canada. In addition, there is another area with a large positive sensitivity over northeast Nova Scotia. Meanwhile, a large negative sensitivity (<-0.6)

is found over the aforementioned ridge. Another negative region (<-0.5) is located over the downstream eastern U.S. ridge. Such a distribution suggests that a weakened wave group from western North America to the western Atlantic Ocean is associated with an increase in PC1 and thus an enhancement of the EOF1 pattern at VT, which is a weaker and eastward-shifted cyclone.

Among the five groups, groups 2 and 3 (Fig. 11d) are associated with the largest positive and negative EOF PC1 values, representing the weakest/most eastward and the strongest/most westward surface cyclones at VT, respectively. At this time step (-60h; Fig. 12a), the differences between the five groups are subtle because the total geopotential field is used. Nevertheless, over regions where the sensitivity to PC1 is large, one can see

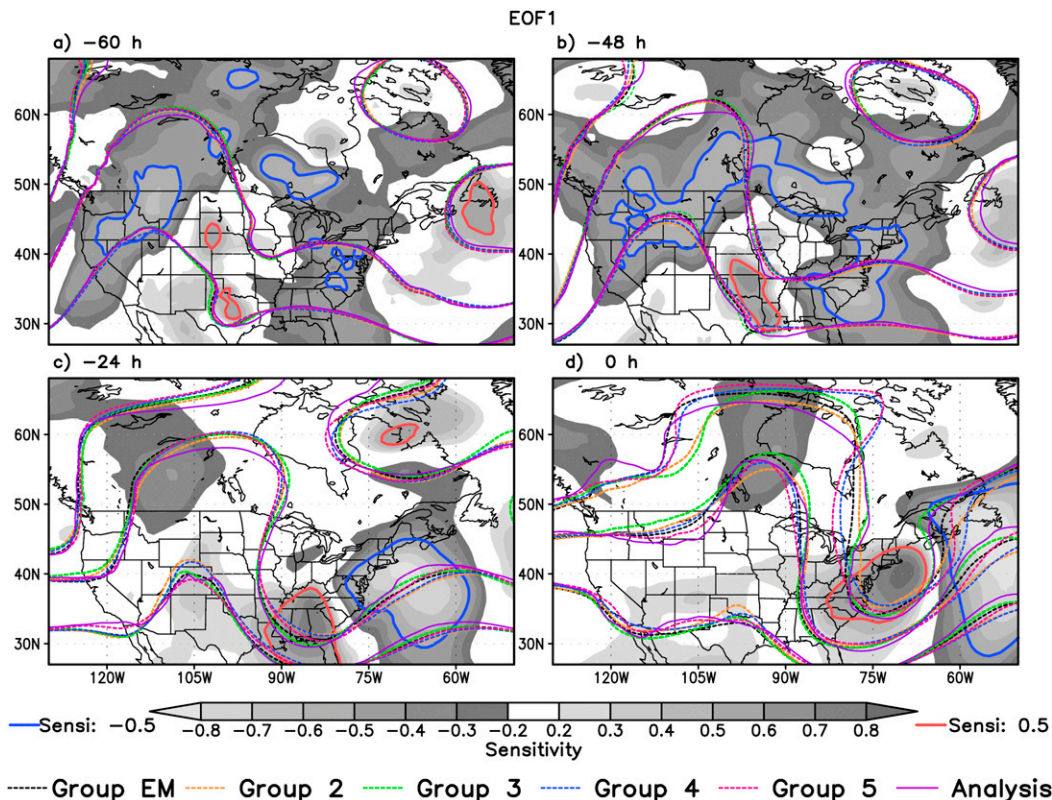


FIG. 12. (a)–(d) Ensemble sensitivity of EOF PC1 to Z500 (gray shading and solid blue/red contours) and five cluster means of three Z500 contour lines (colored contours: 5300, 5500, and 5700 m) from a lead time of 60 h (1-day forecast) to VT (3.5-day forecast). Black, orange, green, blue, and magenta dashed lines are for group EM and groups 2–5, respectively. The VT is 1200 UTC 27 Dec 2010, and the IT is at 0000 UTC 24 Dec 2010.

the evident differences between groups 2 and 3. For example, the Texas short-wave trough of group 2 is weaker and more eastward than that of group 3.

As we look forward in time from -48 to 0 h (Figs. 12b–d), the ridge–trough–ridge pattern generally shifted eastward with the amplitude of the pattern increasing especially over the trough and downstream ridge. The sensitivity signal also generally shifted eastward, suggesting that a weaker ridge–trough–ridge pattern favored positive EOF1 (a weaker and eastward-shifted cyclone). At the VT (Fig. 12d), the sensitivity is maximized (>0.8) near the center and the southwest of the U.S. East Coast trough, where group 2 has the weakest trough while group 3 has the deepest one among the five groups. Meanwhile, the negative sensitivity is over the center of its downstream ridge, where group 2 has the weakest ridge and group 3 has the most amplified ridge among the five groups.

The sensitivity using PC2 as the forecast metric is shown in Fig. 13. At -60 h (Fig. 13a), there is negative sensitivity (<-0.3) to the southwest of the short-wave trough over Texas, suggesting a southwestward-shifted trough is associated with positive EOF2, which

represents the southwestward shift of the cyclone at VT. Groups 3 and 5 represent positive and negative EOF2, respectively (Fig. 11d). Over regions where the sensitivity to PC2 is large, the contrasts between groups 3 and 5 are identifiable, with one example being northwestern Canada.

When going forward from -48 to 0 h (Figs. 13b–d), the sensitivity signal moved eastward with the corresponding systems and increased in amplitude. The positive sensitivity developed rapidly from >0.3 at the base of the trough over the midwestern United States at -48 h (Fig. 13b) to >0.8 over the western Atlantic upper-level ridge (Fig. 13d), which is consistent with the downstream development scenario suggested by Zheng et al. (2013). Meanwhile, the five groups showed large differences in forecasting the downstream ridge over the western Atlantic, again demonstrating the large downstream development of the forecast uncertainty. The distribution of the sensitivity centers corresponding to the East Coast trough and its upstream and downstream ridges suggested that the southwest shift of the ridge–trough–ridge system favors the EOF 2 pattern (i.e., the southwestward shift of the cyclone). Note that over

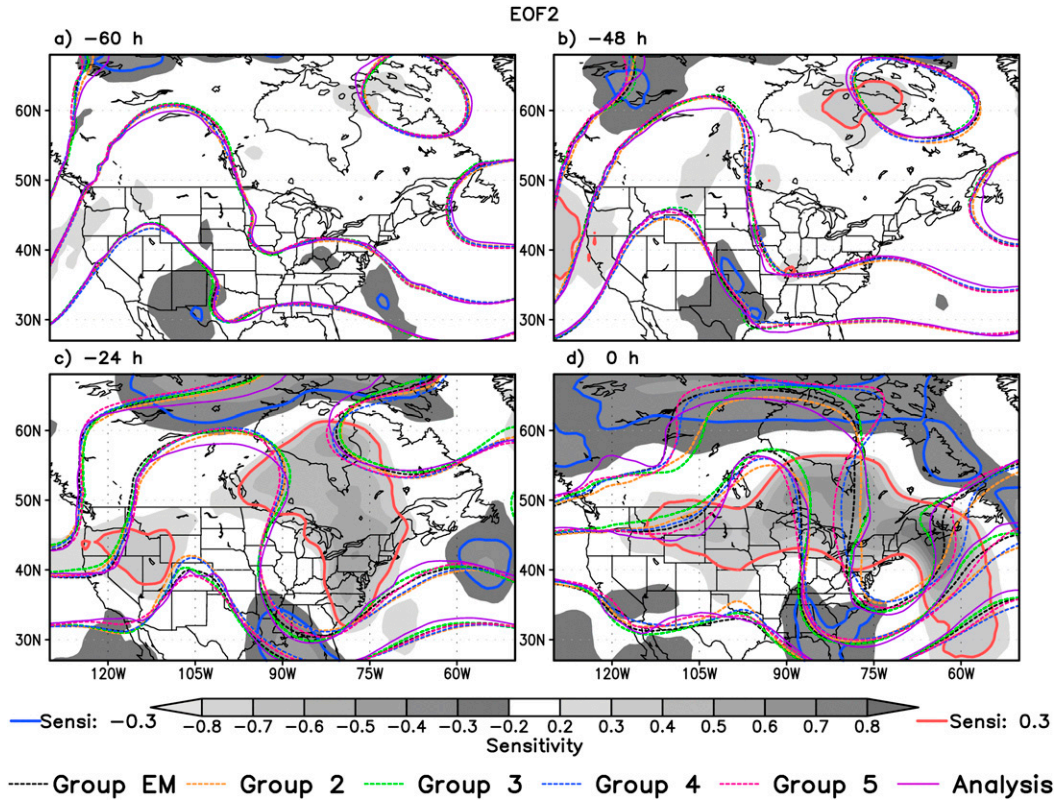


FIG. 13. As in Fig. 12, but for EOF PC2.

regions where sensitivity to PC2 is large while sensitivity to PC1 is small (e.g., over southern Canada between Hudson Bay and the Great Lakes, near Alabama and Wyoming), we can see clear differences between the means of groups 5 and 3 (Fig. 13d).

The time evolutions based on sensitivity and cluster means (Figs. 12 and 13) both suggest that the southern plains short-wave trough has the most consistent and robust signal in affecting the intensity and position shift of the day 3.5 surface cyclone. This is consistent with the results in Zheng et al. (2013) using the same run but only with 50 ECMWF members. Their forward ensemble regression results suggest that other sensitive regions, such as the area of large sensitivity over high-latitude Canada, only play a secondary role in modifying the EOF2 pattern.

d. The benefits of using the two methods together

Figures 12 and 13 qualitatively compares the ensemble sensitivity and the clustering analysis based on the leading two EOF PCs. The ESA method in either correlation form [Eq. (2)] or regression gradient form (Ansell and Hakim 2007) establishes the linear relation between a state vector (e.g., Z500) and the forecast metric. In our applications here, the forecast PC metrics

form the basis for performing clustering calculations. Thus, these two methods were related through the similar forecast metrics used. The clusters at VT can be interpreted as the rotated EOF patterns or a linear combination of PC1 and PC2. Since we have demonstrated that the ensemble sensitivity method can be applied to medium-range ensemble forecasts (Chang et al. 2013; Zheng et al. 2013), the linkage between these two methods suggests that both methods can be utilized in operations to investigate the forecast uncertainty among ensemble outputs for medium-range forecasts.

Each of these two methods has its advantages and limitations. The ensemble sensitivity can tell forecasters where the forecast uncertainty is coming from and how it evolves from the initial times (IT) to the VT. The limitations for using the ensemble sensitivity include that 1) it is difficult to verify the ensemble sensitivity, 2) it is often hard to disclose the details of individual members, and 3) the sensitivity patterns are only associated with either EOF1 or EOF2, and are treated separately in calculations, whereas the real forecasts can be a combination of two patterns. The fuzzy clustering method assigns each member to a particular cluster, which often represents an independent scenario. This scenario can be the positive/negative EOF1 or EOF2, or

any combination of two EOF patterns. The details of each forecast scenario (e.g., a more onshore surface cyclone) can also be explored by examining the spaghetti plots in the cluster. Also, the forecast scenarios in the fuzzy clustering analysis can be easily verified by projecting the analysis onto the EOF PCs coordinate (see section 5). At the initial times, the amplitude of the differences in the ensemble members is small. Ensemble sensitivity highlights the linearly regressed signal, thus removing the noise from that process; hence, the signal is clearer. Composites based on the clusters encompass the full field; thus, during initial times when the signal is small, cluster separations could be contaminated by noise and become less clear. These two methods can serve different purposes in operations. The ensemble sensitivity method provides an overall linear relation between the state vectors and each of the two EOF patterns while fuzzy clustering can show the possibility of different development scenarios. Thus, we believe that these two methods are complementary and should be used together.

5. Discussion

a. Applying EOF/fuzzy clustering to real-time forecasting

The EOF/fuzzy clustering method has several benefits when applied in real-time forecasting situations. 1) It can quickly separate different forecast scenarios among a large multimodel output dataset. The NBM aims to create a “best” product from multimodels in the medium-range time frame. This objective way of separating scenarios among multimodel EPSs could provide another perspective for validating the best product. 2) The time evolution of different clusters can help forecasters determine the critical systems associated with the largest model disagreement, for example, the short-wave trough over the Midwest on day 1 for the January 2015 case (Fig. 9d). 3) Together with the ensemble sensitivity method, it can indicate where the forecast uncertainty is coming from during the initial forecast steps, for example, the short-wave trough over Texas on day 1 for the December 2010 case (Fig. 12a). 4) It can clearly show outlier members on the scatterplot. For example, Fig. 5a shows most of the NCEP members are located over positive PC2 quadrants. The CMC model has an outlier member around $(-3.5, -3.0)$ with respect to the rest of the ensemble. 5) It can reinforce the “equally likely” nature of ensemble forecasts, which will be further discussed in an upcoming paper. Based on over 100 cyclone cases, we have found that the analysis scenario does not have a strong tendency to fall into any one of the clusters beyond the short range. This confirms that in a well-constructed ensemble, all scenarios have an

equal chance to happen; in other words, even the scenario away from the ensemble mean should not be ignored.

b. The analysis group for past cases

When analyzing past cases in which the analysis is available, after the ensemble members have been grouped into an optimal number of clusters based on their EOF PCs at VT, the analysis field can be projected onto the leading EOF patterns and hence occupies one point on the EOF PC1–PC2 space to represent analysis or validation (see the purple plus sign in Figs. 5a, 5b, and 11d). By doing this, the cluster closest to the analysis can be defined as the “analysis group” (“group ANA”). Based on this definition, for the January 2015 case, group 6 in Fig. 5b is the analysis group. As can be seen in Figs. 9 and 10, group 6 (or group ANA) is the closest cluster to the analysis (purple solid line in Figs. 9 and 10), which forecasts a cyclone more north-northeastward than the ensemble mean. The average precipitation corresponding to group 6 is also more north-northeastward than that of the multimodel ensemble mean (Figs. 2e,f). The heaviest precipitation in group 6 is much closer to the observed than other clusters as well as individual model means (Fig. 2). For the December 2010 case, group 3 in Fig. 11d is assigned as the analysis group. From Figs. 12 and 13, the 5400-m Z500 contour in group 3 is the closest to the analyzed contour over the eastern United States and the western Atlantic Ocean. Note that in both cases, group ANA is distinct from group EM; that is, the analyses do not lie within the space occupied by the ensemble mean groups.

Here, the analysis group has been defined based on the projection of the analysis onto PC1–PC2 space. Two more conventional metrics, the root-mean-squared error (RMSE) and pattern correlation coefficient (Corr), are used to verify that the analysis group defined this way is indeed closest to the analysis.

Figures 14 and 15 show the RMSE and Corr values of Z500/MSLP from IT to VT for cases 1 and 2, respectively. At VT (72 h for case 1 and 84 h for case 2), the analysis group (group 6 for case 1 and group 3 for case 2) has smaller RMSE errors (Figs. 14a,b and 15a,b) and higher correlation (Figs. 14c,d and 15c,d) than the other four groups. In addition, they also have smaller RMSE and higher Corr values than the multimodel mean as well as each individual model mean. The superiority of the analysis group is significant at the 95% level based on 100 random samplings of groups.

The superiority of the analysis group is not only limited at VT. The time range of its robustness can also be seen from Figs. 14 and 15. For the 2015 case (Fig. 14), the analysis group is significantly better from 48 h (VT – 30 h; Figs. 14a,b) to 84 h (VT + 12 h; not shown), spanning around 36 h. For the December 2010 case (Fig. 15), the

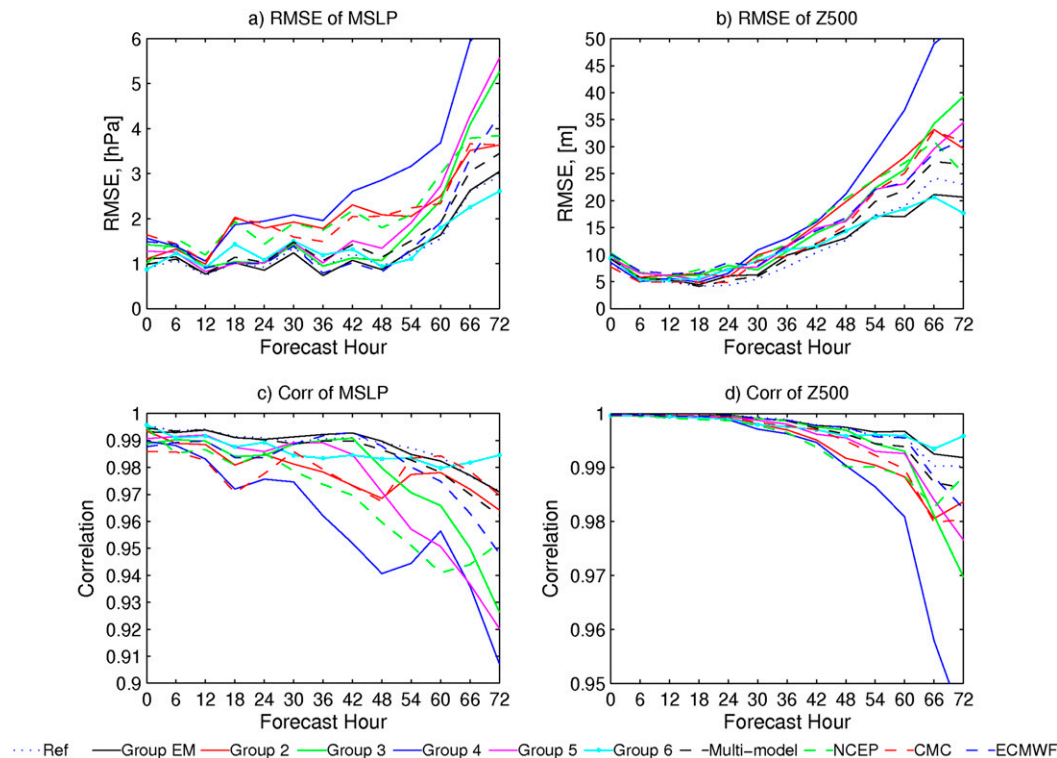


FIG. 14. RMSEs of (a) MSLP and (b) Z500 for each cluster mean and model means at different forecast times. Correlations of (c) MSLP and (d) Z500 between each mean and the analysis at different forecast times. The clusters are selected based on the EOF PCs of the 72-h forecast initialized at 1200 UTC 24 Jan 2015. The black, red, green, blue, purple, and cyan solid lines are for group EM and groups 2–6, respectively. The cyan solid line with dots is for group 6 (analysis group). The black dashed line is for the multimodel ensemble mean. The green, red, and blue dashed lines are for the NCEP, CMC, and ECMWF model means, respectively. The blue dotted line shows the 95% significance line based on 100 randomly selected clusters. The verification area is a box (20° latitude \times 24° longitude) moving with the cyclone (for MSLP) and the corresponding upper-level trough (for Z500) system.

analysis group is significantly better from 30 h (VT – 54 h; Figs. 15b,d) to 120 h (VT + 36 h; not shown), spanning over 3.5 days. Therefore, the time range of the dominant benefit of using the analysis group varies case by case, but it can be robust, spanning as long as over 3.5 days. Nevertheless, Figs. 14 and 15 show that at early forecast times, other groups are closer to the analysis than the analysis group, indicating that one cannot easily pick a “best” forecast group based on its agreement with observations during early forecast hours, consistent with the findings of Smith (2001) and Orrell et al. (2001).

The verifications using RMSE and Corr confirm that the analysis group has relatively lower errors and higher pattern correlations among the six (five for case 2) groups over a time range spanning the VT. The selection of the analysis group is hence reliable and can be applied to event-based scenario verifications. The result also indicates that by combining members from different models, an optimal cluster (group ANA) can be more valuable than both the multimodel ensemble mean and individual model means. In operations, the optimal

cluster is not known in advance. However, the statistics of optimal clusters for specific events (e.g., significant winter storms) can help forecasters to sharpen their forecasts, which will be further explored in an upcoming paper.

The projection of the analysis onto the PC space and the definition of the analysis group can be used as a verification tool. Using TIGGE data, we have applied this method to examine over 100 cases of East Coast cyclones, and the results demonstrate that in the majority of cases, the analysis does not lie within group EM, similar to the two cases discussed here, suggesting that one should not put too much emphasis on the ensemble mean and ignore the other groups. This and other statistics will be discussed in more detail in an upcoming paper.

6. Conclusions

This study introduces an EOF/fuzzy clustering method for separating different development scenarios when forecasting East Coast storms out to medium range.

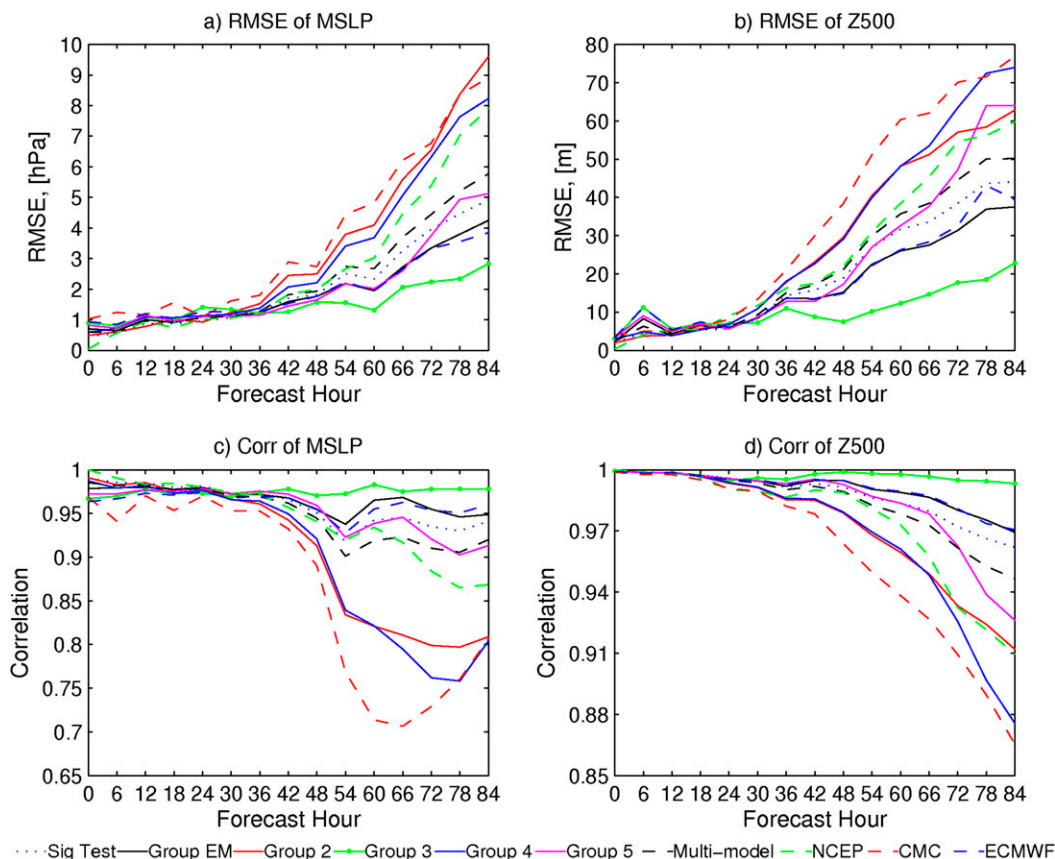


FIG. 15. As in Fig. 14, but for the five-cluster results of case 2 initialized at 0000 UTC 24 Dec 2010 and verified at 1200 UTC 27 Dec 2010. Group 3 (green solid line with dot) is the analysis group.

In real-time operations, this work can meet the following needs: 1) to quickly partition different forecast scenarios, 2) to provide a new scenario-based method for NBM to validate the “best” product, 3) to indicate where and when the forecast uncertainties are from, and 4) to increase forecasters’ situational awareness of the “equal chance to happen” for each scenario.

Two winter storm cases are investigated. For the January 2015 case, the separation into six groups represents storms with different locations or amplitudes. The verified analysis group shows more of an offshore solution, with the heaviest snowfall region not including NYC. In the 3-day forecast, the ECMWF ensemble, which has the largest number of members, does not contribute to the analysis group in proportion to its number of members. The NCEP model, in contrast, has around half of its members contributing to the analysis group.

The December 2010 case is used to compare the clustering method and ensemble sensitivity analysis. These two methods are complementary to each other. Early in the forecasts, when differences among the clusters are small, the ensemble sensitivity signals can show the sensitive regions more clearly. However, closer

to VT, the clusters can show the evolution of forecast uncertainty and different scenarios more clearly.

The validation using RMSE and Corr results confirms that our definition of the analysis group is quantitatively reasonable. Out of the scenarios represented by the optimal number of clusters, there is in general one group closest to the analysis. For both cases examined, the analysis groups are distinct from the ensemble mean group, and have fewer errors than both the multimodel mean and individual model means.

In this paper, we have only shown the clusters using MSLP parameter. Since high-impact weather events (e.g., heavy precipitation, high winds) are strongly associated with the cyclone position, it has motivated us to use the mass field as a first-step parameter for our project. We are expanding the applications to other fields too. We have also explored different ways of calculating clustering by utilizing other meteorological parameters, for example, the total precipitation, the temperature at different levels, and the potential vorticity at the dynamical tropopause. The cluster means show physical consistency when using these different parameters. However, more work needs to

be done to combine different weather elements to perform clustering and interpret the synoptic relations among scenarios.

This tool, when applied to past cases, can also be used as an ensemble verification tool. We have applied this tool to examine over 100 cases involving East Coast cyclones, and the results will be presented in an upcoming paper. Other than winter storm events, this tool can also be utilized to investigate scenarios in different HIW conditions (i.e., Atlantic hurricanes or tropical cyclones in any basin). We have successfully applied the fuzzy clustering tool to interpreting the real-time forecast uncertainty in two recent hurricanes (Hermine in September 2016 and Matthew in October 2016). This method can provide insights into improving our understanding of HIW predictability and model bias for both operational and research purposes.

Acknowledgments. The authors thank the TIGGE data archives for the availability of the ensemble data. We also thank three anonymous reviewers for their deep insights and helpful comments on an earlier version of the manuscript. This work is supported by NOAA-CSTAR (NA13NWS4680002).

REFERENCES

- Ahmed, M. N., S. M. Yamany, N. Mohamed, A. A. Farag, and T. Moriarty, 2002: A modified fuzzy C-means algorithm for bias field estimation and segmentation of MRI data. *IEEE Trans. Med. Imaging*, **21**, 193–199.
- Ancell, B., and G. J. Hakim, 2007: Comparing adjoint- and ensemble-sensitivity analysis with applications to observation targeting. *Mon. Wea. Rev.*, **135**, 4117–4134, doi:10.1175/2007MWR1904.1.
- Booth, J. F., H. E. Rieder, D. E. Lee, and Y. Kushnir, 2015: The paths of extratropical cyclones associated with wintertime high-wind events in the northeastern United States. *J. Appl. Meteor. Climatol.*, **54**, 1871–1885, doi:10.1175/JAMC-D-14-0320.1.
- Bougeault, P., and Coauthors, 2010: The THORPEX Interactive Grand Global Ensemble. *Bull. Amer. Meteor. Soc.*, **91**, 1059–1072, doi:10.1175/2010BAMS2853.1.
- Bowler, N. E., A. Arribas, K. R. Mylne, K. B. Robertson, and S. E. Beare, 2008: The MOGREPS short-range ensemble prediction system. *Quart. J. Roy. Meteor. Soc.*, **134**, 703–722, doi:10.1002/qj.234.
- Brill, K. F., A. R. Fracasso, and C. M. Bailey, 2015: Applying a divisive clustering algorithm to a large ensemble for medium-range forecasting at the Weather Prediction Center. *Wea. Forecasting*, **30**, 873–891, doi:10.1175/WAF-D-14-00137.1.
- Chang, E. K. M., M. Zheng, and K. Raeder, 2013: Medium-range ensemble sensitivity analysis of two extreme Pacific extratropical cyclones. *Mon. Wea. Rev.*, **141**, 211–231, doi:10.1175/MWR-D-11-00304.1.
- Colle, B. A., and M. E. Charles, 2011: Spatial distribution and evolution of extratropical cyclone errors over North America and its adjacent oceans in the NCEP Global Forecast System model. *Wea. Forecasting*, **26**, 129–149, doi:10.1175/2010WAF2222422.1.
- , F. Buonaiuto, M. J. Bowman, R. E. Wilson, R. Flood, R. Hunter, A. Mintz, and D. Hill, 2008: New York City's vulnerability to coastal flooding. *Bull. Amer. Meteor. Soc.*, **89**, 829–841, doi:10.1175/2007BAMS2401.1.
- , K. Rojowsky, and F. Buonaiuto, 2010: New York City storm surges: Climatology and an analysis of the wind and cyclone evolution. *J. Appl. Meteor. Climatol.*, **49**, 85–100, doi:10.1175/2009JAMC2189.1.
- , Z. Zhang, K. A. Lombardo, E. Chang, P. Liu, and M. Zhang, 2013: Historical evaluation and future prediction of eastern North American and western Atlantic extratropical cyclones in the CMIP5 models during the cool season. *J. Climate*, **26**, 6882–6903, doi:10.1175/JCLI-D-12-00498.1.
- Du, J., and B. Zhou, 2011: A dynamical performance-ranking method for predicting individual ensemble member performance and its application to ensemble averaging. *Mon. Wea. Rev.*, **139**, 3284–3303, doi:10.1175/MWR-D-10-05007.1.
- Fanning, A., 2015: Storm summary number 09 for mid-Atlantic to Northeast major winter storm. NCEP/Weather Prediction Center, College Park, MD. [Available online at http://www.wpc.ncep.noaa.gov/winter_storm_summaries/2015/storm2/stormsum_9.html.]
- Ferranti, L., and S. Corti, 2011: New clustering products. *ECMWF Newsletter*, No. 127, ECMWF, Reading, United Kingdom, 6–12. [Available online at <http://www.ecmwf.int/en/elibrary/14596-newsletter-no127-spring-2011>.]
- Flegenhimer, M., 2015: Leaders in New York and New Jersey defend shutdown for a blizzard that wasn't. *New York Times*, 28 January. [Available online at <http://www.nytimes.com/2015/01/28/nyregion/new-york-blizzard.html>.]
- Frankoski, N. J., and A. T. DeGaetano, 2011: An East Coast winter storm precipitation climatology. *Int. J. Climatol.*, **31**, 802–814, doi:10.1002/joc.2121.
- Gilbert, K. K., and Coauthors, 2016: The national blend of global models, version one. *23rd Conf. on Probability and Statistics in the Atmospheric Sciences*, New Orleans, LA, Amer. Meteor. Soc., 1.3. [Available online at <https://ams.confex.com/ams/96Annual/webprogram/Paper285973.html>.]
- Grams, C. M., H. Wernli, M. Böttcher, J. Čampa, U. Corsmeier, S. C. Jones, and L. Wiegand, 2011: The key role of diabatic processes in modifying the upper-tropospheric wave guide: A North Atlantic case-study. *Quart. J. Roy. Meteor. Soc.*, **137**, 2174–2193, doi:10.1002/qj.891.
- Harr, P. A., and R. L. Elsberry, 1995: Large-scale circulation variability over the tropical western North Pacific. Part I: Spatial patterns and tropical cyclone characteristics. *Mon. Wea. Rev.*, **123**, 1225–1246, doi:10.1175/1520-0493(1995)123<1225:LSCVOT>2.0.CO;2.
- , D. Anwender, and S. C. Jones, 2008: Predictability associated with the downstream impacts of the extratropical transition of tropical cyclones: Methodology and a case study of Typhoon Nabi (2005). *Mon. Wea. Rev.*, **136**, 3205–3225, doi:10.1175/2008MWR2248.1.
- Houtekamer, P. L., L. Lefavre, J. Derome, H. Ritchie, and H. L. Mitchell, 1996: A system simulation approach to ensemble prediction. *Mon. Wea. Rev.*, **124**, 1225–1242, doi:10.1175/1520-0493(1996)124<1225:ASSATE>2.0.CO;2.
- Johnson, A., X. Wang, F. Kong, and M. Xue, 2011: Hierarchical cluster analysis of a convection-allowing ensemble during the Hazardous Weather Testbed 2009 Spring Experiment. Part I: Development of the object-oriented cluster analysis method

- for precipitation fields. *Mon. Wea. Rev.*, **139**, 3673–3693, doi:10.1175/MWR-D-11-00015.1.
- Keller, J. H., S. C. Jones, J. L. Evans, and P. A. Harr, 2011: Characteristics of the TIGGE multimodel ensemble prediction system in representing forecast variability associated with extratropical transition. *Geophys. Res. Lett.*, **38**, L12802, doi:10.1029/2011GL047275.
- Maimon, O., and L. Rokach, Eds., 2005: *The Data Mining and Knowledge Discovery Handbook*. Springer, 1382 pp.
- Molteni, F., R. Buizza, T. N. Palmer, and T. Petroliagis, 1996: The ECMWF Ensemble Prediction System: Methodology and validation. *Quart. J. Roy. Meteor. Soc.*, **122**, 73–119, doi:10.1002/qj.49712252905.
- Orrell, D., L. Smith, J. Barkmeijer, and T. N. Palmer, 2001: Model error in weather forecasting. *Nonlinear Processes Geophys.*, **8**, 357–371, doi:10.5194/npg-8-357-2001.
- Palmer, T. N., C. Brankovic, F. Molteni, S. Tibaldi, L. Ferranti, A. Hollingsworth, U. Cubasch, and E. Klinker, 1990: The European Centre for Medium-Range Weather Forecasts (ECMWF) program on extended-range prediction. *Bull. Amer. Meteor. Soc.*, **71**, 1317–1330, doi:10.1175/1520-0477(1990)071<1317:TECFMR>2.0.CO;2.
- Richman, M. B., 1986: Rotation of principal components. *J. Climatol.*, **6**, 293–335, doi:10.1002/joc.3370060305.
- Scott, A. J., and M. J. Symons, 1971: Clustering methods based on likelihood ratio criteria. *Biometrics*, **27**, 387–397, doi:10.2307/2529003.
- Smith, L. A., 2001: Disentangling uncertainty and error: On the predictability of nonlinear systems. *Nonlinear Dynamics and Statistics*, A. I. Mees, Ed., Birkhäuser, 31–64.
- Tew, M. A., A. Stem, A. Horvitz, K. Gilbert, and D. T. Myrick, 2016: National Blend of Models: Transformational forecast change by the National Weather Service. *Fourth Symp. on Building a Weather-Ready Nation*, New Orleans, LA, Amer. Meteor. Soc., 6.2. [Available online at <https://ams.confex.com/ams/96Annual/webprogram/Paper286228.html>.]
- Toth, Z., and E. Kalnay, 1993: Ensemble forecasting at NMC: The generation of perturbations. *Bull. Amer. Meteor. Soc.*, **74**, 2317–2330, doi:10.1175/1520-0477(1993)074<2317:EFANTG>2.0.CO;2.
- Winkler, R. L., 2015: The importance of communicating uncertainties in forecasts: Overestimating the risks from winter storm Juno. *Risk Anal.*, **35**, 349–353, doi:10.1111/risa.12398.
- Yeung, K. Y., and W. L. Ruzzo, 2001: Principal component analysis for clustering gene expression data. *Bioinformatics*, **17**, 763–774, doi:10.1093/bioinformatics/17.9.763.
- Zhang, F., C. Snyder, and R. Rotunno, 2002: Mesoscale predictability of the surprise snowstorm of 24–25 January 2000. *Mon. Wea. Rev.*, **130**, 1617–1632, doi:10.1175/1520-0493(2002)130<1617:MPOTSS>2.0.CO;2.
- Zheng, M., E. K. Chang, and B. A. Colle, 2013: Ensemble sensitivity tools for assessing extratropical cyclone intensity and track predictability. *Wea. Forecasting*, **28**, 1133–1156, doi:10.1175/WAF-D-12-00132.1.