Operational Forecaster Uncertainty Needs and Future Roles

DAVID R. NOVAK

NOAA/National Weather Service, Eastern Region Headquarters, Scientific Services Division, Bohemia, New York

DAVID R. BRIGHT

NOAA/National Weather Service, Storm Prediction Center, Norman, Oklahoma

MICHAEL J. BRENNAN

NOAA/National Weather Service, Hydrometeorological Prediction Center, Camp Springs, Maryland

ABSTRACT

Key results of a comprehensive survey of U.S. National Weather Service operational forecast managers concerning the assessment and communication of forecast uncertainty are presented and discussed. The survey results revealed that forecasters are using uncertainty guidance to assess uncertainty, but that limited data access and ensemble underdispersion and biases are barriers to more effective use. Some respondents expressed skepticism as to the added value of formal ensemble guidance relative to simpler approaches of estimating uncertainty, and related the desire for feature-specific ensemble verification to address this skepticism. Respondents reported receiving requests for uncertainty information primarily from sophisticated users such as emergency managers, and most often during high-impact events. The largest request for additional training material called for simulator-based case studies that demonstrate how uncertainty information should be interpreted and communicated.

Respondents were in consensus that forecasters should be significantly involved in the communication of uncertainty forecasts; however, there was disagreement regarding if and how forecasters should adjust objective ensemble guidance. It is contended that whether forecasters directly modify objective ensemble guidance will ultimately depend on how the weather enterprise views ensemble output (as the final forecast or as a guidance supporting conceptual understanding), the enterprise's commitment to provide the necessary supporting forecast infrastructure, and how rapidly ensemble weaknesses such as underdispersion, biases, and resolution are addressed.

The survey results illustrate that forecasters' operational uncertainty needs are intimately tied to the end products and services they produce. Thus, it is critical that the process to develop uncertainty information in existing or new products or services be a sustained collaborative effort between ensemble developers, forecasters, academic partners, and users. As the weather enterprise strives to provide uncertainty information to users, it is asserted that addressing the forecaster needs identified in this survey will be a prerequisite to achieve this goal.

1. Introduction

Uncertainty is a fundamental characteristic of hydrometeorological (hydrologic, weather, and seasonal climate) prediction, and is a consequence of the inherent chaotic nature of the atmosphere, inadequate observations, and numerical weather prediction (NWP) defi-

ciencies (NRC 2006). Thus, the assessment and communication of uncertainty is an inherent part of any forecast process.

The assessment of uncertainty in modern operational forecasting has largely relied on the use of ensemble prediction systems (EPSs). First proposed by Leith (1974), operational EPS approaches to assessing uncertainty became practical in the early 1990s (e.g., Toth and Kalnay 1993; Brooks et al. 1995). Today, sophisticated global EPSs are run operationally at national centers worldwide (WMO 2003), and higher-resolution regional short-range EPSs are also operational at several

Corresponding author address: David R. Novak, NOAA/NWS, Eastern Region Headquarters, Suite 202, 630 Johnson Ave., Bohemia, NY 11716.
E-mail: david.novak@noaa.gov

national centers (e.g., Du et al. 2003; Marsigli et al. 2005; Bowler et al. 2008) and universities (e.g., Mass et al. 2003; Jones and Colle 2007). Model output statistics (MOS; Glahn and Lowry 1972) have also provided probabilistic guidance to help assess event uncertainty [e.g., probability of precipitation (PoP)].

The potential socioeconomic advantage of providing uncertainty information over traditional deterministic forecasts has been demonstrated (e.g., Katz and Murphy 1997; Pielke 1999; AMS 2002; Keith 2003); however, identifying effective methods of communicating forecast uncertainty has been challenging (AMS 2002; NRC 2006). Recent research studies have explored how users interpret and use forecast uncertainty (e.g., Morss and Ralph 2007; Morss et al. 2008; Roulston and Kaplan 2008), but arguably less attention has been focused on what resources *forecasters* need to assess and communicate forecast uncertainty. Both the assessment and communication of forecast uncertainty could be improved with information about what operational forecasters view as their uncertainty needs. Such information also documents the current state of the operational assessment and communication of uncertainty.

As part of an effort to improve the generation and dissemination of uncertainty information, the National Oceanic and Atmospheric Administration (NOAA)/ National Weather Service (NWS) conducted a comprehensive survey of NWS operational managers concerning guidance, training, products and services, and forecast system needs related to forecast uncertainty. To the authors' knowledge, this was the first comprehensive survey concerning operational forecaster uncertainty needs. The goal of the survey was to obtain feedback regarding

- uncertainty guidance needs, especially needs related to high-impact events (e.g., heavy snow, floods, high winds, tropical cyclones);
- training needs related to assessing and communicating forecast uncertainty;
- operational barriers to using uncertainty information in forecast preparation, and expressing uncertainty information in forecast products; and
- what current deterministic forecast processes, products, and services could benefit from the addition of forecast uncertainty information.

This article presents key findings from the survey organized into the areas of uncertainty guidance, training, products and services, and future forecaster roles. Although it is recognized that the results of the survey were likely influenced by the specific operational infrastructure and practices of the NWS, particular focus is placed on survey questions that may be applicable to the broader weather enterprise.[1]

2. Survey methodology

A small team of NWS and Office of Atmospheric Research (OAR) employees worked with Claes Fornell International (CFI) Group, Inc.—a private company specializing in customer feedback—to develop the survey. Development of the survey was guided by qualitative interviews with NWS operational forecast managers and a review of available uncertainty guidance by the team. Draft survey questions were peer reviewed for structure, content, and clarity. The final survey was composed of 21 questions, 12 of which were open ended. Specific survey questions discussed in this paper are provided in the appendix. The CFI Group programmed and hosted the Web survey on a secure server.

E-mail invitations were sent to operational forecast managers at the National Centers for Environmental Prediction's (NCEP) National Centers (NCs), Weather Forecast Offices (WFOs), and River Forecast Centers (RFCs). The NWS has five NCs, 122 WFOs, and 13 RFCs with forecast responsibility for the United States and its territories (four additional NCs serve in a forecast support capacity). NCs are responsible for providing hydrometeorological forecasts and guidance on a national scale, RFCs are responsible for providing hydrological forecasts and guidance on a regional scale, and WFOs are responsible for providing hydrometeorological forecasts on a local scale. The survey participants were directors, branch chiefs, Warning Coordination Meteorologists (WCMs), and Science and Operations Officers (SOOs) at NCs, Meteorologists in Charge (MICs), WCMs, and SOOs at WFOs, and Hydrologists in Charge (HICs) and Development and Operational Hydrologists (DOHs) at RFCs. A majority of these operational managers are experienced forecasters, which may have impacted the survey results. Future work sampling a broader forecaster population is encouraged. Data were collected from 21 August to 19 September 2007. Survey participation was voluntary and responses were anonymous.

A total of 237 responses were received out of a possible 399, equating to a 59% response rate. Of the 237 total responses, 214 were from WFOs (59% response rate), 16 from RFCs (67% response rate), and 7 from

---

[1] For the purpose of this paper, the "weather enterprise" is defined as the group of public, private, and academic entities associated with weather and climate forecasts.

NCs (54% response rate). Given the small sample size for RFCs and NCs, caution should be taken when interpreting the RFC and NC results. Similarly, given the disparity in sample size among the WFOs, RFCs, and NCs, caution should be taken when comparing results between these groups. The CFI Group provided the NWS results of the survey separated into NC, WFO, and RFC answers. Responses to the open-ended questions were analyzed by the team members to identify common themes.

3. Results

a. Uncertainty guidance

Since the missions of WFOs, RFCs, and NCs are unique, there were two primary operational systems used to view guidance and generate products at the time of the survey. WFOs and RFCs used the Advanced Weather Interactive Processing System (AWIPS) (Friday 1994, p. 47) and NCEP centers used the National Advanced Weather Interactive Processing System (N-AWIPS). Common uncertainty guidance sources, their referring acronym, and their availability in AWIPS/N-AWIPS and on Web pages are shown in Table 1.

Given the variety of uncertainty datasets available, it is of interest to learn what datasets are used in forecast preparation. Question 4 (Q4) presented a list of common uncertainty guidance sources and asked respondents to identify which uncertainty guidance datasets are used in forecast preparation in their office (see Table 1 for guidance acronyms). The results indicated that the most common uncertainty guidance used by NCs was the global ensemble forecast system (GEFS) (Fig. 1). For WFOs, the GEFS, Short Range Ensemble Forecast System (SREF), and Ensemble MOS were used nearly equally (Fig. 1), while the most common uncertainty guidance used by RFCs was the ESP, followed by the GEFS (Fig. 1). Analysis of the individual responses to this question showed that 234 of the 237 respondents (99%) chose more than one data source, testifying to the use of EPS guidance by many operational forecasters to assess forecast uncertainty. Unique approaches, such as using ensemble standardized anomaly information to anticipate high-impact events (Grumm and Hart 2001; Stuart and Grumm 2006) or using normalized ensemble confidence measures (Durante et al. 2005) to determine whether a certain time period is more uncertain than average, appeared as responses in the "other" category of Q4 and in responses to Q6.

Q5 assessed what forecasters view as the most critical uncertainty guidance issues that need to be addressed. The results (Table 2) show that NC and WFO respondents were most concerned with data availability in their respective operational systems. Ensemble underdispersion (the solution falls outside the envelope of ensemble solutions a disproportionate amount of the time) was rated second and third for the NCs and WFOs, respectively. Ensemble probability calibration was a priority for RFC respondents, followed closely by underdispersion. Given the common concern for data access and ensemble underdispersion and bias, these issues are discussed in further detail below.

1) DATA ACCESS

As shown in Table 1, uncertainty guidance is available to NWS forecast offices via operational display systems and Web pages, although the quantity and form (e.g., individual members, means, probabilities) available in AWIPS/N-AWIPS varies. Web pages are often consulted to fill the data access gaps. For example, although the NAEFS was not available in AWIPS at the time of the survey (Table 1), nearly 20% of WFO respondents reported using it (Fig. 1).

One of the reasons such gaps in data access exist is that EPS guidance datasets are notoriously large—scaled approximately by the number and resolution of members composing the EPS. Given the vast amount of information provided by an EPS and finite communication speed, it is often asked whether a forecaster needs to see every individual member of an EPS, or could be served just as well by viewing summary information, such as the 10th, 50th (median), and 90th percentile values of the ensemble distribution. Q9 and Q9.2 assessed to what degree forecasters need access to individual ensemble members in high-impact events. For NCs and WFOs the most common responses were "all of the time" or "some of the time," while the most common response for RFCs was "some of the time" (Fig. 2). Open responses to Q6 suggested that forecasters wish to view individual ensemble members since it allows them to treat the EPS as an interactive system rather than a "black box." Forecasters can view alternative scenarios noting extremes, assess member initializations, and interpret the ensemble mean and probability fields in a dynamical context (e.g., comparing the forecast evolution of members that have analyzed a deeper trough versus those that have a weaker trough). The desire of forecasters to interact with ensemble guidance is explored further in section 3d.

Given that many EPSs have over 20 members and that this number is expected to grow, can a forecaster realistically view output from numerous members? Q9.2 asked the respondent to quantify just how many members they would expect to view given a high-impact event. NC respondents were the most willing to

TABLE 1. Primary NWS operational uncertainty guidance sources, acronym names, and availability (as of Aug 2007).

| Guidance source	Acronym	WFO/RFC AWIPS availability	NCEP N-AWIPS availability	Primary Web site availability
Ensemble model output statistics	Ensemble MOS	Not available	Not available	http://www.weather.gov/mdl/synop/products.shtml
Global Ensemble Forecast System	GEFS	Individual member fields or means/spreads of 10 of the members	All individual members and mean, spread, and probability fields derived from full ensemble	http://wwwt.emc.ncep.noaa.gov/gmb/ens/
Short Range Ensemble Forecast System	SREF	Means, spreads, and probability fields derived from full ensemble; no individual members	Individual fields from all members; mean, spread, and probability fields available for entire ensemble	http://wwwt.emc.ncep.noaa.gov/mmb/SREF/SREF.html
North American Ensemble Forecast System	NAEFS	Not available	All individual members and mean, spread, and probability fields derived from full ensemble	http://wwwt.emc.ncep.noaa.gov/gmb/ens/NAEFS.html
European Centre for Medium-Range Weather Forecasts ensemble	ECMWF ensemble	Not available	Not available	Coarse-resolution spaghetti plots and ensemble mean available to the NCEP HPC via internal Web page
Ensemble streamflow prediction	ESP	Available to RFCs	Not available	Not available
Advanced Hydrologic Prediction System	AHPS	Not available	Not available	http://www.nws.noaa.gov/oh/ahps/
Climate Forecast System	CFS	Not available	Available to Climate Prediction Center	http://cfs.ncep.noaa.gov/
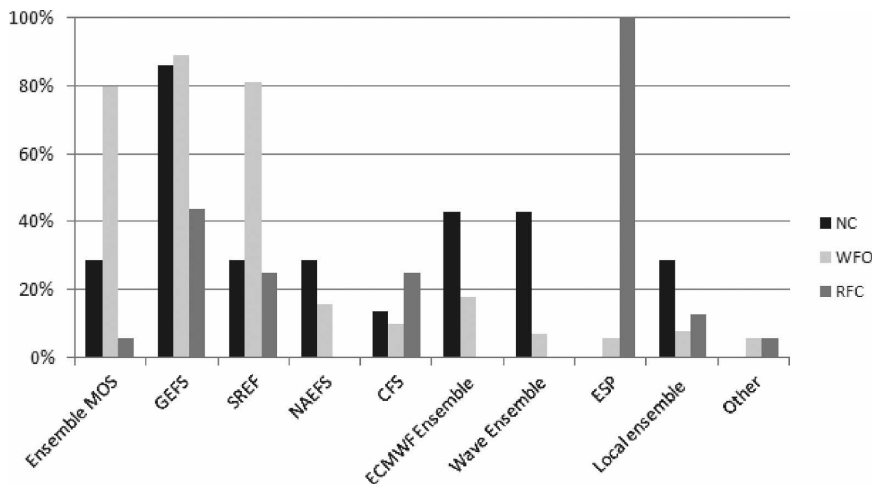
FIG. 1. Percentage of respondents in each survey class using the specified uncertainty guidance source (survey Q4, $N = 237$).

view a large number of ensemble members, with 29% responding "up to 40" and 14% with "as many as available" (Fig. 3). WFO respondents showed a clear preference for "up to 10" (62%), followed by "up to 20" (21%), and "as many as possible" (10%) (Fig. 3). RFCs showed a similar preference for "up to 10" (50%) (Fig. 3). At the time of the survey, WFOs and RFCs were able to view 10 individual GEFS members in AWIPS, while NCs were generally able to view all GEFS/SREF members in N-AWIPS (Table 1). It is possible that the WFO and RFC responses to Q9.2 were biased by display capabilities. Thus, the willingness of NC respondents to view more members may suggest that if forecasters are provided output from more members, they will use them. Visualization methods such as "postage stamp" displays, where each member forecast is displayed on one page (e.g., Leutbecher 2005, his Fig. 3; Palmer and Hagedorn 2006, their Fig. 1.9), may facili-

tate quick assessment of the variety of solutions predicted among members.

Taken as a whole, the responses to Q9 and Q9.2 show a clear desire for the ability to view and analyze output from individual ensemble members. The utilization of server–client data distribution systems such as the National Oceanic and Atmospheric Administration (NOAA) National Operational Model Archive and Distribution System (NOMADS; Rutledge et al. 2006) may provide a cost-effective means of providing operational forecasters individual ensemble member solutions in the future. In the context of ensembles, such a system could allow the user to select the variables and domain of interest, likely reducing the size of a particular EPS dataset by more than half. The utilization of a server–client data infrastructure is planned for future NWS operational systems (Lawson et al. 2007).

2) ENSEMBLE UNDERDISPERSION AND BIASES

Theoretically, EPSs simultaneously provide an estimate of forecast uncertainty through the variation in ensemble member solutions, and minimize error through the ensemble mean (Kalnay 2003, p. 26; Sivillo et al. 1997). In practice, EPSs are generally underdispersive and have biases that limit the assessment of forecast uncertainty and degrade forecast accuracy (e.g., Hamill and Colucci 1997; Stensrud and Yussouf 2003; Eckel and Mass 2005; Buizza et al. 2005; Jones and Colle 2007). Application of running-mean or weighted bias-correction schemes to EPSs can mitigate these limitations to some degree (e.g., Stensrud and Yussouf 2003; Yussouf et al. 2004; Baars and Mass 2005; Woodcock and Engel 2005; Eckel and Mass 2005;

TABLE 2. Average ranking of uncertainty guidance issues respondents would most like to see addressed, with 1 being "address first" and 5 being "address last" (survey Q5, $N = 237$).

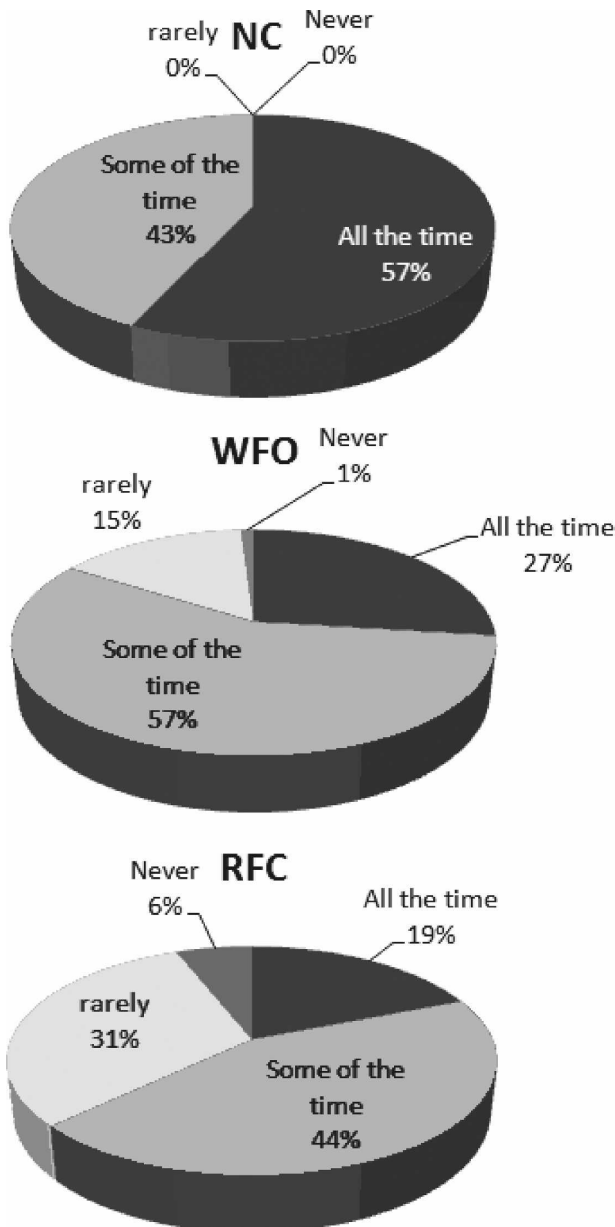|	NC	WFO	RFC
The actual solution falls too often outside the envelope of possible solutions generated by the ensemble			
Avg rank	2.18	2.84	2.43
The probabilities provided are not calibrated			
Avg rank	2.6	3.01	2.04
The data format provided is not useful			
Avg rank	3.84	2.68	2.86
Data are not available in AWIPS/GFE/N-AWIPS			
Avg rank	1.84	1.73	3.16

FIG. 2. Respondents' expected frequency of viewing individual ensemble members in a high-impact event (survey Q9.1, $N = 237$).

Jones and Colle 2007), and provide ensemble mean forecasts that are competitive with deterministic MOS (Yussouf and Stensrud 2007; Cheng and Steenburgh 2007). However, bias-correction schemes perform worst in rapidly evolving flow regimes (characteristic of high-impact events) when biases are changing rapidly (e.g., Cheng and Steenburgh 2007, 1313–1315). Other approaches, such as reforecasting (Hamill et al. 2004; Hamill and Whitaker 2006), Bayesian model averaging (e.g., Raftery et al. 2005), and the application of neural

networks (e.g., Yuan et al. 2007), have also improved ensemble bias and underdispersion; however, since high-impact events are relatively rare by definition, these approaches also suffer in these situations. Ensemble bias correction was not operational at NCEP at the time of the survey, but even after applying postprocessing approaches, operational EPSs will likely exhibit weaknesses in high-impact events.

Open responses to Q6 and Q14 suggested that these weaknesses create skepticism among some forecasters as to the operational utility of formal EPSs, especially in relation to the use of simpler techniques such as deterministic model runs or a "poor-man's ensemble," defined as a combination of independent deterministic models (Ebert 2001; Arribas et al. 2005). Open responses to Q6 and Q14 also noted that high-impact events commonly verified in the tails of the ensemble distribution, or outside the distribution altogether. In the words of one respondent, "The science of ensembles is still developing. The spread is still limited."

Verification evidence of EPS weaknesses in high-impact events is slowly being gathered. Verification of storm tracks compiled by the NCEP Hydrometeorological Prediction Center (HPC) for over 500 extratropical cyclones producing hazardous winter weather over the contiguous United States during the 2004–07 cold seasons (15 September–15 May; see http://www.hpc.ncep.noaa.gov/wwd/winter_wx.shtml) shows that at the 48-h forecast projection (similar results were found at 24- and 72-h projections), the global forecast system (GFS) was superior to the GEFS mean (Fig. 4).[2] Furthermore, a simple North American Mesoscale (NAM)/GFS model average was found to be superior to the available global deterministic models and the ensemble means of the GEFS and SREF (Fig. 4). Similar results comparing the skill of the GFS and SREF mean for extratropical cyclone track and intensity have been found by Colle and Charles (2007). In a study of early warnings of gale winds, heavy snow, and flooding rain over the United Kingdom, Legg and Mylne (2004) show that on most occasions when these high-impact events occurred, the European Centre for Medium-Range Weather Forecasts (ECMWF) EPS only predicted them with low probabilities.[3] In this respect the EPS provided a

_____

[2] HPC storm-track verification includes extratropical cyclones directly associated with a >10% probability of exceeding 10.2 cm (4 in.) of snow and/or 0.64 cm (0.25 in.) of freezing rain as forecast for HPC's probabilistic winter weather desk products. Forecast surface low positions are taken from deterministic and ensemble mean model grids available operationally at HPC. The verifying dataset is HPC's manual surface analysis.

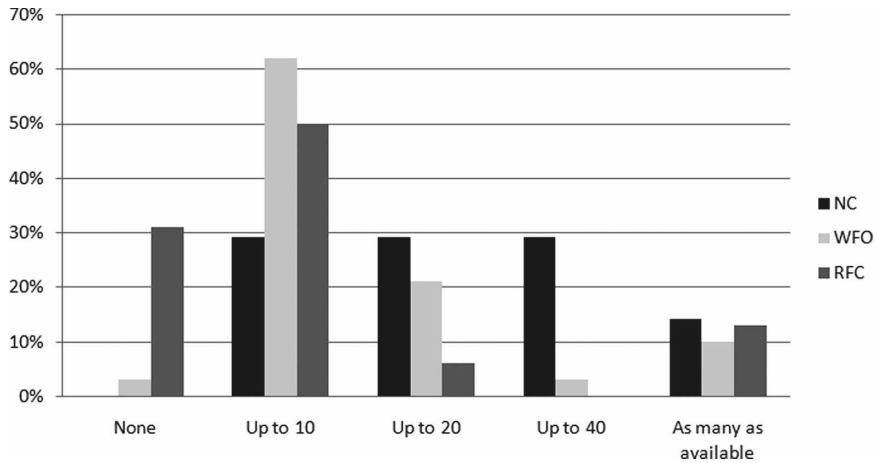[3] The verification methodology is outlined in sections 4 and 5 of Legg and Mylne (2004).

FIG. 3. Respondents' expected number of ensemble members they would view in a high-impact event (survey Q9.2, N = 237).

"heads up;" however, similar hit rates and false alarms were recorded for a high-resolution deterministic model. Legg and Mylne (2004) interpreted these results as showing that high-impact weather is an intrinsically low probability event, requiring the joint occurrence of anomalous individual events, and that an EPS should not be expected to generate high probabilities except in highly predictable states. Legg and Mylne (2004) found that human severe weather predictions (constrained to have a 60% confidence) were much more skillful than the EPS (and high-resolution model) at this confidence threshold.

Additional verification of EPS performance for other types of high-impact events is needed to further validate the above forecaster perceptions, expand previous research results, and determine a path to improvement.

Responses to Q6 and Q14 indicated the desire for event- or feature-specific (e.g., Ebert and McBride 2000; Davis et al. 2006a,b) verification approaches to be applied to EPSs to demonstrate the added value of ensemble guidance relative to single deterministic runs or a poor man's ensemble. Traditionally verification is applied over long time periods, which averages performance during active and quiescent periods. Given forecasters' experience that EPSs perform poorly for some high-impact events, a critical test is how the EPS verifies for high-impact events. Responses indicated that forecasters will quickly gravitate to formal EPSs when such verification demonstrates that EPSs are providing more accurate information than deterministic or poor man's ensemble approaches, and reliable estimates of
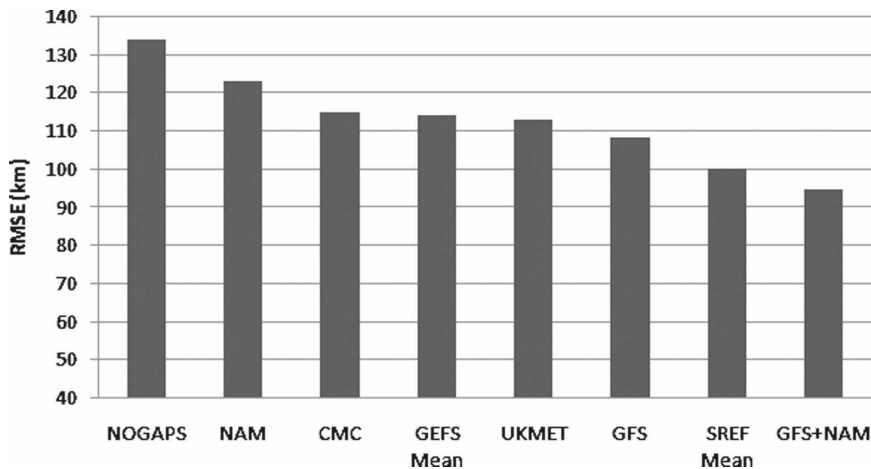

FIG. 4. The 48-h forecast surface low track RMSE (km) for extratropical cyclones producing hazardous winter weather over the contiguous United States compiled by HPC during the 2004–07 cold seasons (15 Sep–15 May). (Data provided by HPC.)
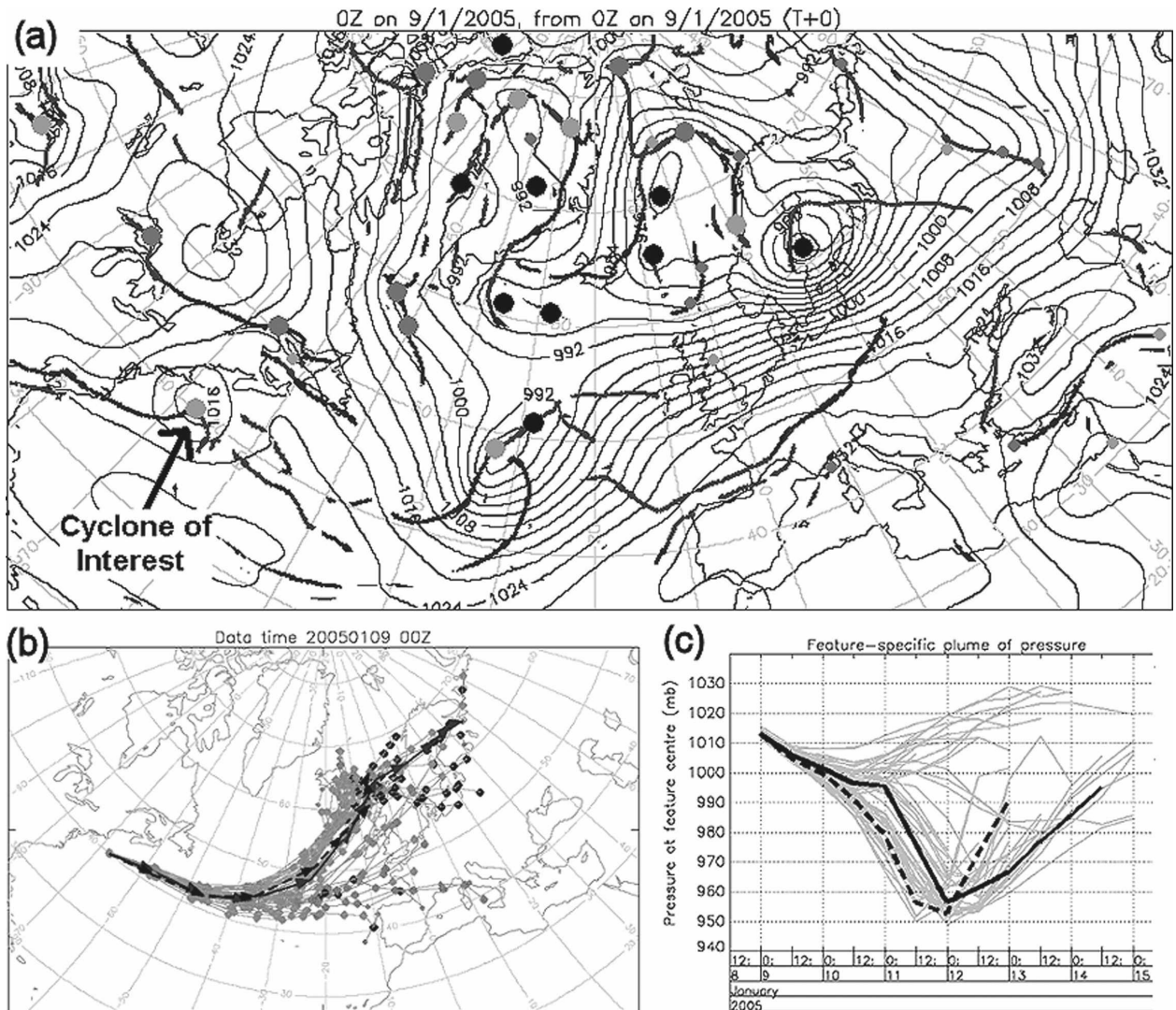
FIG. 5. Example feature-specific verification using the Cyclone Database for the ECMWF ensemble prediction system from 0000 UTC 9 Jan 2005. (a) Objectively identified fronts and cyclonic features at initial time. Key frontal wave is off the New England coast. (b) Forecast cyclone tracks from each ensemble member (mean = thick line; analysis = dashed line). (c) Associated feature-specific plumes of minimum sea level central pressure. [Adapted from Watkin and Hewson (2006).]

uncertainty for high-impact events. These results are consistent with Morss and Ralph (2007, p. 549), who note that demonstration of a large improvement is often required to convince forecasters to change forecast procedures.

An example of a feature-based EPS verification system is the Met Office Cyclone Database (Hewson 2002). This system uses objective algorithms to identify fronts and cyclones in model output and to track characteristics of these features through the forecast (Watkin and Hewson 2006). Example verification for a cyclone affecting the United Kingdom with heavy rain and hurricane-force winds is shown in Fig. 5. Although the mean cyclone track was near the observed track

(Fig. 5b), the verifying cyclone central sea level pressure was located in the lower tail of the distribution (Fig. 5c). This result illustrates the point of Legg and Mylne (2004) that extreme weather is often a low-probability event, and is consistent with the responses to Q6 and Q14 that high-impact events commonly verify in the tails of the distribution.

An EPS's horizontal resolution may contribute to underdispersion and biases, especially biases related to the magnitude of resolution-dependent features. For example, members composing the SREF at the time of the survey were run with horizontal grid spacing of 32–45 km. At this resolution, mesoscale features such as orographic flows, precipitation bands, and convective

systems are not adequately resolved, which limits the EPS's ability to provide uncertainty information for these high-impact features (NRC 2006, p 48). Open responses to Q6 indicated that forecasters are aware of this limitation and desire higher-resolution EPSs. For example, if a regional or global EPS is showing a high probability of exceeding 50 mm of precipitation, the forecaster often interprets this result as a high probability of exceeding a larger amount (i.e., 75 mm of precipitation). Tests with high-resolution EPSs illustrate the validity of such forecast adjustments (e.g., Walser et al. 2006). In the absence of available high-resolution EPSs, open responses indicated that locally run high-resolution models are often consulted to provide further confidence when making such an interpretation. Similarly, if an EPS is showing a high probability of large values of CAPE, shear, and convective precipitation, high-resolution models are consulted to determine the convective mode (e.g., Fowle and Roebber 2003; Weiss et al. 2006). This complementary approach of using high-resolution models and coarser-resolution ensembles has been recommended by Roebber et al. (2004). As computational power continues to increase, the implementation of high-resolution EPSs, which are becoming available on regional scales (e.g., Mass et al. 2003; Jones and Colle 2007), is an important step in providing reliable uncertainty information for resolution-dependent features.

b. Training

A recurrent theme in the survey results was the need for forecaster training in both assessing and communicating uncertainty. Ensemble training for NWS forecasters has traditionally drawn on Web modules developed by the Cooperative Program for Operational Meteorology, Education and Training (COMET) (e.g., Bua 2005). Q13 and Q13.1 asked whether the respondent's office had developed local uncertainty guidance training to supplement the COMET modules, and if so, what the training entailed. Only about 22% (53 of 237 responses) of the respondents indicated that their office had developed local training for uncertainty guidance (many indicated that they relied on COMET-produced training). The most common topic focused on ensemble utility and interpretation, and tended to spotlight the SREF and GEFS products available in AWIPS/N-AWIPS. Several offices developed training on the availability of Web-based ensemble guidance. Responses indicated that training often focused on high-impact events such as tropical cyclones, heavy precipitation, and convection. Responses to Q13 noted that more ensemble training would occur if the products and services that forecasters produced incorporated uncertainty.

Q14 asked respondents to rate on a scale of 1 to 10 (where 1 is poor and 10 is excellent) forecasters' knowledge of five topics (EPS design and perturbations, statistics, decision support, weather risk management, and user requirements) as they apply to preparing forecasts. If a topic area was rated below 6, the respondent was prompted to describe what additional training would better prepare them to produce uncertainty forecasts for high-impact events (Q14.1).

Respondents indicated they needed additional training in user requirements more than any other topic. Numerous respondents indicated that they needed a better understanding of the type of uncertainty information their users require, a better understanding of how their users will utilize uncertainty information, and a better method of communicating the necessary information to them. Results from recent studies and this survey (see section 3c) exploring user requirements and effective communication may guide development of such training, although clearly more work is needed in this area (AMS 2002; NRC 2006).

The respondents also identified the need for additional training in EPS design, and statistics. A common theme expressed by respondents was the need to make EPSs less of a black box by providing a more complete description of how the systems are constructed (perturbation methods, model, and physics diversity) and how associated guidance products are derived. Responses indicated the need for ensemble application and interpretation training on a range of subjects from basic interpretations of ensemble mean and spread diagrams to advanced cluster analysis. As one response noted,

"Too often ensemble training has focused on technical aspects rather than application. How to apply the output in a range of scenarios given the aforementioned strength and weaknesses of the ensemble system is the bottom line need."

Consistent with this comment, the largest request for additional training called for (Weather Event Simulator) WES-like case studies that show how ensemble guidance should be interpreted and communicated in end products and services. The WES is a software package that simulates the AWIPS software environment, providing real-time simulation of cases (Magsig and Page 2003). Respondents indicated that this type of hands-on training is more effective than distance learning approaches.

Respondents also identified the need for additional training in decision support and risk management. Specifically, needs included better training to communicate uncertainty, to express uncertainty relative to climatol-

ogy (e.g., a 2% chance of a tornado at a certain location may seem low but may be 25 times higher than climatology), and to provide decision support with incomplete information.

c. Uncertainty products and services

At the time of the survey, forecast uncertainty was conveyed in NWS products through a variety of approaches. NCs communicated forecast uncertainty through text discussions and event-specific graphical forecast products [such as the probability of a tornado within 25 miles of a point, or the tropical cyclone track cone of uncertainty (Broad et al. 2007)]. WFOs' primary means of communicating uncertainty was the Area Forecast Discussion text product, and the probability of a precipitation element integrated into gridded, graphical, and text products. RFCs' primary means of communicating uncertainty information was through the Advanced Hydrologic Prediction Service (AHPS; McEnery et al. 2005). However, a majority of NWS products were deterministic (e.g., Glahn and Ruth 2003; Mass 2003; Glahn 2003), and with the exception of AHPS, the PoP element, and select probabilistic products from NCs, NWS uncertainty information was generally presented in qualitative terms (NRC 2006, section 3.4).

Given this product suite, Q16 asked what deterministic products, services, and processes would be most enhanced with additional uncertainty information. The question was answered by 160 out of 235 respondents. Sixteen respondents (10%) simply replied "all of them," "everything," or something similar. Precipitation-related products were particularly emphasized, with 52 respondents (33%) specifically identifying precipitation amount, PoP, or precipitation types. Winter weather and snowfall were specifically identified by 30 respondents (19%), while 19 respondents (12%) identified river stages and other hydrological forecasts.

Q17 asked respondents what requests for uncertainty information they have received from users. Seventy-four percent of respondents (120 of 164) reported receiving specific requests, highlighting the user demand for uncertainty information. Among these responses, two common themes were

• most requests were made by "sophisticated users" such as emergency managers, fire weather and flood control officials, and agricultural interests; and
• most requests focused on "high impact" events.

Frequent types of uncertainty information requested included

• the likelihood of an event occurring (e.g., a freeze) or of exceeding some threshold (29%);

• some qualitative measure of forecast confidence or uncertainty (i.e., "how confident are you") (18%);
• information on the "worst case scenario" for an event (15%); and
• information on the range of possibilities (9%).

Open responses to Q17 and Q18 indicated that much of this information is relayed directly to federal, state, and local officials through briefings. Such close forecaster–user interactions have been highlighted by Morss and Ralph (2007) in the context of emergency mangers' use of forecast information for landfalling extratropical storms along the West Coast, and featured as a future role for forecasters by Mass (2003).

d. Future forecaster roles

The role of forecasters in an increasingly automated forecast process has been a topic of recent discussion (e.g., Mass 2003; Glahn 2003; Bosart 2003; Doswell 2004; Roebber et al. 2004; Baars and Mass 2005; Stuart et al. 2006). In the context of uncertainty forecasts, the reliability and sharpness of EPS distributions will improve as observations continue to expand, data assimilation and perturbation approaches are refined, better models are developed, bias correction schemes improve, and the resolution of members increases. In anticipation of such an improved EPS, Q12 asked respondents what they think the role of the forecaster should be in developing uncertainty forecasts.

Of the 158 total responses to Q12, 105 responses directly addressed to what degree forecasters should be involved in developing uncertainty forecasts. Seventy-nine respondents (75%) felt that forecasters should be significantly involved, 14 (13%) felt there should be minimal involvement (i.e., hands off), and 13 (12%) were not sure or stated that it depended on various conditions. Of those respondents that felt there should be significant involvement, nearly all felt that forecasters should be significantly involved in the communication of uncertainty. Weather enterprise–relevant responses of what this role could entail included

• interpreting forecasts for users (decision assistance),
• explaining alternative forecast scenarios,
• educating users on probabilities, and
• assisting in development of products based on user needs.

The potential expansion of the forecasters' role in interpreting and communicating uncertainty has been noted by many (e.g., AMS 2002; Mass 2003; Baars and Mass 2005; NRC 2006), and as discussed in section 3c, there is current demand for this function from sophisticated users. Perhaps more importantly, recent re-

search suggests that the general public would also be receptive to uncertainty information (Morss et al. 2008; Roulston and Kaplan 2008).

There was more disagreement regarding if and how forecasters should be involved in the modification of objective EPS guidance. Roughly half of the respondents envisioned such a role, while others felt that forecasters should be focused on communicating uncertainty and interpreting end products and services for users. Much of this divide was related to how a forecaster perceived an EPS—as a bias-corrected, calibrated dataset that provided the final *forecast* (in which case the respondent was not inclined to adjust the output), or as *guidance* used in concert with other information to develop a conceptual understanding of the meteorological situation (in which case the respondent was inclined to adjust the output). Respondents also considered the forecast projection; expressing more comfort in accepting the EPS output as the final forecast at extended time ranges.

Responses as to how forecasters envisioned adjusting objective guidance included

- local perturbation of key features (e.g., shortwaves, jets),
- selecting the most probable member,
- elimination of errant members, and subsequent recalculation of probabilities/fields, and
- subjective EPS bias correction on a local scale.

Considering the variety of approaches suggested, are all of these forecaster roles feasible?

There is reason to believe that the local perturbation of key features can be skillfully made for short-range forecasts. For example, Meteo-France forecasters draw upon water vapor–potential vorticity (PV) relationships (e.g., Appenzeller and Davies 1992; Demirtas and Thorpe 1999; Santurette and Georgiev 2005) to correct model initial analyses (Guerin et al. 2006). In the Meteo-France system the forecaster is primarily augmenting a single deterministic solution; however, Homar et al. (2006) show in an experimental context that when forecasters perturb features they identify as important to short-range high-impact convective forecasts, better probabilistic information concerning that event is produced than when the probabilistic information is obtained by traditional automated perturbation techniques. Operational perturbation techniques such as bred vectors (Toth and Kalnay 1993), singular vectors (Molteni et al. 1996), and perturbed observations (Houtekamer et al. 1996) have been shown to be deficient for short-range forecasting (e.g., Houtekamer and Derome 1995; Matthieu and Arbogast 2005; Buizza et al. 2005), and it is likely the forecaster can improve

upon these perturbation strategies. Besides improved verification, the direct involvement of a forecaster in the generation of an ensemble can further the forecaster's conceptual understanding of a weather situation through hypothesis testing (e.g., Roebber et al. 2002).

On the other hand, selecting the most probable member of an ensemble has proven difficult. In the specific context of 24–36-h convective forecasts for the 2001 National Severe Storms Laboratory (NSSL)/Storm Prediction Center (SPC) Spring Program (Kain et al. 2003b), forecast teams showed little skill in assessing how "good" individual model forecasts would be (Kain et al. 2003a). Bright and Nutter (2004) suggest that this result is due to the fact that model forecasts can be "right for the wrong reason," and that different members may provide the "best" forecast at varying times over a short forecast period.

Although selecting the most probable ensemble member may be difficult, eliminating errant members may be more feasible. In the context of tropical cyclone track prediction, Payne et al. (2007) showed that Joint Typhoon Warning Center (JTWC) forecasters were able to skillfully create a selective consensus forecast by discarding members that were subjectively identified to be in error. Whether such positive results can be achieved for other weather phenomena is not known.

The survey results suggest that subjective bias correction of EPS output is already occurring in the form of local downscaling of EPS information (i.e., adjusting fields for topography), or the adjustment of ensemble means and ensemble distributions to account for resolution-dependent features. Carroll and Hewson (2005) describe the Met Office capability to directly edit short-range NWP guidance by either modifying model analyses and forecasts in a dynamically consistent manner via PV inversion (e.g., Hoskins et al. 1985), or through subjective correction of model forecast biases of sensible weather elements. Although not specifically designed for ensemble application, the method has been used in conjunction with ensembles to create a deterministic forecast that better matches the forecaster's best estimate of the most likely outcome (K. Mylne 2008, personal communication). Verification shows that the forecaster modifications have an overall positive impact, with the ratio of improved to degraded short-range forecasts at approximately 4:1 (Carroll and Hewson 2005). Although improved objective bias-correction and downscaling schemes may reduce the opportunities for forecaster adjustments in the future, such schemes have weaknesses in changing flow regimes, when accurate guidance is needed most.

The examples above suggest that given sufficient data access, training, and supporting forecast systems,

forecasters *can* play an important role in the direct modification of objective EPS guidance, especially for short-range high-impact events. Among the proposed approaches, all appear feasible with exception of the selection of the most probable member. However, based on the forecast improvements shown at Meteo-France and the NSSL/SPC Spring Experiment, the perturbation of key features is advocated as the favored method in which forecasters responsible for short-range weather forecasts should modify objective guidance. Besides generating improved probabilistic information for the event of interest, the ability to perturb key features supports the active appraisal of conceptual understanding through hypothesis testing—a key aspect of skilled forecasters (Roebber et al. 2004, p. 946; Bosart 2003; Doswell 2004; Stuart et al. 2007). This deep conceptual understanding will assist in the communication of uncertainty.

4. Discussion and summary

Key results of a comprehensive survey of NOAA/NWS operational forecast managers concerning guidance, training, products and services, and future forecaster roles related to forecast uncertainty were presented. The major findings are

- many forecasters use uncertainty guidance to assess uncertainty;
- forecasters desire access to the output of individual ensemble members;
- forecasters view data access and EPS underdispersion and biases as critical issues related to uncertainty guidance that need to be addressed;
- ensemble underdispersion and biases have created skepticism among some forecasters regarding the added value of EPS guidance relative to single deterministic runs or a poor man's ensemble, especially for high-impact events;
- there is a desire for event- or feature-specific EPS verification to identify and improve weaknesses to address this skepticism;
- forecasters desire information concerning their users' uncertainty requirements;
- simulator-based case studies that show how uncertainty information should be applied in the forecast process, from guidance interpretation to product generation, are needed;
- there is demand for uncertainty information, primarily from "sophisticated users" such as emergency managers, most often during high-impact events; and
- there is consensus that forecasters should be significantly involved in the communication of uncertainty

forecasts; however, there is disagreement as to if and how forecasters should adjust objective guidance.

The survey results suggest that the current generation of operational EPSs provides opportunities for human intervention, especially for short-range high-impact events. Greater data access, training, and supporting forecast systems will help forecasters act on these opportunities for some time in the future. Whether forecasters directly modify objective EPS guidance and how far this is practiced into the future will ultimately depend on how the weather enterprise views EPS output (as the final forecast or as guidance supporting conceptual understanding), the enterprise's commitment to provide the supporting forecast infrastructure, and how rapidly EPS weaknesses such as underdispersion, biases, and resolution are addressed.

The survey results illustrate that forecasters' operational uncertainty needs are directly tied to the end products they produce. For example, to provide probabilistic precipitation forecasts, forecasters need precipitation uncertainty guidance, forecast systems to display and modify that guidance, and training in how to interpret and apply the guidance to create the forecast product. New or enhanced products often result from advancements in science and technology and specified user requirements. Thus, it is critical that the process to develop uncertainty information in forecasts be a sustained collaborative effort between EPS developers, forecasters, academic partners, and users. Such an effort is currently ongoing in the United States in the form of the American Meteorological Society's Ad Hoc Committee on Uncertainty in Forecasting (AMS 2007), and on the world stage in the form of The Observing System Research and Predictability Experiment (THORPEX; WMO 2007; Toth and Majumdar 2007). These and similar future efforts are needed to serve as a forum to address the forecaster uncertainty needs identified in this survey.

Regardless of whether probabilistic forecast products and services are expanded in the future, there is evidence that the expansion and improvement of uncertainty guidance, training, and supporting forecast systems can improve current deterministic forecasts. For example, ensemble spread can be used to predict the error that will exist in deterministic (and ensemble mean) forecasts (e.g., Buizza 1997; Sivillo et al. 1997). In practice, Joslyn et al. (2007) found that access to probability information improved threshold forecast decisions. In fact, the decision by the SPC to issue the first ever day-two "high risk" severe weather outlook in the United States for the significant tornado outbreak that occurred 7 April 2007 was dependent on the un-

certainty information provided by the NCEP SREF (C. Broyles 2007, personal communication). These results and the survey results in general highlight the urgent need to expand and improve uncertainty guidance, training, and supporting forecast systems. As the weather enterprise strives to provide uncertainty information to users, it is our conclusion that addressing the forecaster needs identified in this survey will be a prerequisite to achieve this goal.

APPENDIX

Survey Questions

The survey questions discussed in the paper are presented below, using the question numbers and order from the full survey. The total number of respondents is noted for each question.

a. Q4 (N = 237)

In your office, which of the ensemble datasets from the following list are used in forecast preparation?

Ensemble datasets

a. Ensemble MOS
b. NCEP GFS ensemble (GEFS)
c. NCEP Short Range Ensemble Forecast System (SREF)
d. North American Ensemble Forecast System (NAEFS) (combined GFS and Canadian ensembles)
e. Climate Forecast Ensemble (CFS)
f. ECMWF ensemble
g. NCEP Wave Watch III Ensemble
h. Ensemble Steamflow Prediction (ESPADP)
i. Local ensemble systems (please specify)
j. other (please specify)

b. Q5 (N = 237)

Considering the datasets your forecasters use, what are the issues you would most like to see addressed? Please rank order them from 1–5, with 1 being "Address first" and 5 being "Address last."

a. The actual solution that falls too often outside the envelope of possible solutions generated by the ensemble (i.e., lack of dispersion among members)
b. The probabilities provided are not calibrated
c. The data format provided is not useful (e.g., need sensible weather elements on grid)
d. Data are not available in AWIPS/GFE/N-AWIPS
e. Other (please specify)

c. Q6 (N = 204)

What type of additional ensemble information do forecasters in your office need to prepare forecasts for high-impact events? Some examples of high-impact events include: hurricanes, tornadoes, hail storms, and damaging winds.

d. Q9 (N = 237)

Given a high-impact event, how frequently would your forecasters view individual members of an ensemble, if available?

a. All the time
b. Some of the time
c. Rarely
d. Never

e. Q9.2 (N = 237)

Given a high-impact event and a very large ensemble, how many individual members of an ensemble do you expect your forecasters would view (choose one)?

a. None
b. Up to 10
c. Up to 20
d. Up to 40
e. Up to 60
f. As many as are available

f. Q12 (N = 158)

What do you think the role of the forecaster should be in developing uncertainty forecasts?

g. Q13 (N = 237)

Has your office developed local training for uncertainty guidance?

h. Q13.1 (N = 53)

What did the training entail?

i. Q14 (N = 237)

On a scale of 1 to 10 where 1 is poor and 10 is excellent, rate your forecasters' knowledge in these areas as they apply to preparing forecasts.

a. Ensemble design and perturbations
b. Statistics
c. Decision support
d. Weather risk management
e. User requirements

j. Q14.1 (N = 112)

If the above is less than 6, then what additional training would your forecasters benefit from to better prepare themselves to produce uncertainty forecasts for high-impact events?

k. Q16 (N = 160)

What current deterministic forecast products, services, and processes would be most enhanced with additional uncertainty information?

l. Q17 (N = 164)

What requests for uncertainty information has your office received from end users?

m. Q18 (N = 146)

What local products has your office developed to specifically incorporate uncertainty information for end users?

REFERENCES

AMS, 2002: Enhancing weather information with probability forecasts. *Bull. Amer. Meteor. Soc.,* **83,** 450–452.

——, cited 2007: Ad hoc committee on uncertainty in forecasts charter. [Available online at http://www.ametsoc.org/boardpges/cwce/docs/BEC/ACUF/Charter.pdf.]

Appenzeller, C., and H. C. Davies, 1992: Structure of stratospheric intrusions into the troposphere. *Nature,* **358,** 570–572.

Arribas, A., K. B. Robertson, and K. R. Mylne, 2005: Test of a poor man's ensemble prediction system for short-range probability forecasting. *Mon. Wea. Rev.,* **133,** 1825–1839.

Baars, J. A., and C. F. Mass, 2005: Performance of National Weather Service forecasts compared to operational, consensus, and weighted model output statistics. *Wea. Forecasting,* **20,** 1034–1047.

Bosart, L. F., 2003: Whither the weather analysis and forecasting process? *Wea. Forecasting,* **18,** 520–529.

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.

Bright, D. R., and P. A. Nutter, 2004: On the challenges of identifying the "best" ensemble member in operational forecasting. Preprints, *16th Conf. on Numerical Weather Prediction,* Seattle, WA, Amer. Meteor. Soc., J11.3. [Available online at http://ams.confex.com/ams/pdfpapers/69092.pdf.]

Broad, K., A. Leiserowitz, J. Weinkle, and M. Steketee, 2007: Misinterpretations of the "cone of uncertainty" in Florida during the 2004 hurricane season. *Bull. Amer. Meteor. Soc.,* **88,** 651–667.

Brooks, H. E., M. S. Tracton, D. J. Stensrud, G. DiMego, and Z. Toth, 1995: Short-range ensemble forecasting: Report from a workshop, 25–27 July 1995. *Bull. Amer. Meteor. Soc.,* **76,** 1617–1624.

Bua, W. R., 2005: COMET ensemble prediction training. Preprints, *21st Conf. on Weather Analysis and Forecasting/17th Conf. on Numerical Weather Prediction,* Washington, DC, Amer. Meteor. Soc., J2.7. [Available online at http://ams.confex.com/ams/pdfpapers/94545.pdf.]

Buizza, R., 1997: Potential forecast skill of ensemble prediction and spread and skill distributions of the ECMWF ensemble prediction system. *Mon. Wea. Rev.,* **125,** 99–119.

——, P. L. Houtekamer, Z. Toth, G. Pellerin, M. Wei, and Y. Zhu, 2005: A comparison of the ECMWF, MSC, and NCEP global ensemble prediction systems. *Mon. Wea. Rev.,* **133,** 1076–1097.

Carroll, E. B., and T. D. Hewson, 2005: NWP grid editing at the Met Office. *Wea. Forecasting,* **20,** 1021–1033.

Cheng, W. Y. Y., and W. J. Steenburgh, 2007: Strengths and weaknesses of MOS, running-mean bias removal, and Kalman filter techniques for improving model forecasts over the western United States. *Wea. Forecasting,* **22,** 1304–1318.

Colle, B. A., and M. Charles, 2007: Verification of extratropical cyclones within NCEP forecast models using an automated tracking algorithm. Preprints, *22nd Conf. on Weather Analysis and Forecasting/18th Conf. on Numerical Weather Prediction,* Park City, UT, Amer. Meteor. Soc., 6B.3. [Available online at http://ams.confex.com/ams/pdfpapers/124612.pdf.]

Davis, C. A., B. Brown, and R. Bullock, 2006a: Object-based verification of precipitation forecasts. Part I: Methodology and application to mesoscale rain areas. *Mon. Wea. Rev.,* **134,** 1772–1784.

——, ——, and ——, 2006b: Object-based verification of precipitation forecasts. Part II: Application to convective rain systems. *Mon. Wea. Rev.,* **134,** 1785–1795.

Demirtas, M., and A. J. Thorpe, 1999: Sensitivity of short-range weather forecasts to local potential vorticity modifications. *Mon. Wea. Rev.,* **127,** 922–939.

Doswell, C. A., 2004: Weather forecasting by humans—heuristics and decision making. *Wea. Forecasting,* **19,** 1115–1126.

Du, J., G. DiMego, M. S. Tracton, and B. Zhou, 2003: NCEP short-range ensemble forecasting (SREF) system: Multi-IC, multi-model and multi-physics approach. Research activities in atmospheric and oceanic modeling, J. Cote, Ed., CAS/JSC Working Group Numerical Experimentation Rep. 23, WMO/TD 1161, 5.09–5.10.

Durante, A. V., R. E. Hart, A. Watson, R. H. Grumm, and W. Drag, 2005: The development of forecast confidence measures using NCEP ensembles and their real-time implementation within NWS web-based graphical forecasts. Preprints, *18th Conf. on Probability and Statistics in the Atmospheric Sciences,* Atlanta, GA, Amer. Meteor. Soc., 5.1. [Available online at http://ams.confex.com/ams/pdfpapers/100178.pdf.]

Ebert, E. E., 2001: Ability of a poor man's ensemble to predict the probability and distribution of precipitation. *Mon. Wea. Rev.,* **129,** 2461–2479.

——, and J. L. McBride, 2000: Verification of precipitation in

weather systems: Determination of systematic errors. *J. Hydrol.,* **239,** 179–202.

Eckel, F. A., and C. F. Mass, 2005: Aspects of effective short-range ensemble forecasting. *Wea. Forecasting,* **20,** 328–350.

Friday, E. W., Jr., 1994: The modernization and associated restructuring of the National Weather Service: An overview. *Bull. Amer. Meteor. Soc.,* **75,** 43–52.

Fowle, M. A., and P. J. Roebber, 2003: Short-range (0–48 h) numerical prediction of convective occurrence, mode, and location. *Wea. Forecasting,* **18,** 782–794.

Glahn, H. R., 2003: Comments on "IFPS and the future of the National Weather Service." *Wea. Forecasting,* **18,** 1299–1304.

——, and D. A. Lowry, 1972: The use of model output statistics (MOS) in objective weather forecasting. *J. Appl. Meteor.,* **11,** 1203–1211.

——, and D. P. Ruth, 2003: The new digital forecast database of the National Weather Service. *Bull. Amer. Meteor. Soc.,* **84,** 195–201.

Grumm, R. H., and R. Hart, 2001: Standardized anomalies applied to significant cold season weather events: Preliminary findings. *Wea. Forecasting,* **16,** 736–754.

Guerin, R., G. Desroziers, and P. Arbogast, 2006: 4D-Var analysis of potential vorticity pseudo-observations. *Quart. J. Roy. Meteor. Soc.,* **132,** 1283–1298.

Hamill, T. M., and S. J. Colucci, 1997: Verification of Eta–RSM short-range ensemble forecasts. *Mon. Wea. Rev.,* **125,** 1312–1327.

——, and J. S. Whitaker, 2006: Probabilistic quantitative precipitation forecasts based on reforecast analogs: Theory and application. *Mon. Wea. Rev.,* **134,** 3209–3229.

——, ——, and X. Wei, 2004: Ensemble reforecasting: Improving medium-range forecast skill using retrospective forecasts. *Mon. Wea. Rev.,* **132,** 1434–1447.

Hewson, T. D., 2002: A comparison of cyclone spectra in forecasts from the operational (G1) and new dynamics trial (NT) versions of the unified model—Aug to Dec 2001. Forecasting Research Tech. Rep. 376, Met Office, 15 pp. [Available online at http://www.metoffice.gov.uk/research/nwp/publications/papers/technical_reports/2002/FRTR376/FRTR376.pdf.]

Homar, V., D. J. Stensrud, J. J. Levit, and D. R. Bright, 2006: Value of human-generated perturbations in short-range ensemble forecasts of severe weather. *Wea. Forecasting,* **21,** 347–363.

Hoskins, B. J., M. E. McIntyre, and A. Q. Robertson, 1985: On the use and significance of isentropic potential vorticity maps. *Quart. J. Roy. Meteor. Soc.,* **111,** 877–946.

Houtekamer, P. L., and J. Derome, 1995: Methods for ensemble prediction. *Mon. Wea. Rev.,* **123,** 2181–2196.

——, ——, H. Ritchie, and H. L. Mitchell, 1996: A system simulation approach to ensemble prediction. *Mon. Wea. Rev.,* **124,** 1225–1242.

Jones, M. S., and B. A. Colle, 2007: Evaluation of a mesoscale short-range ensemble forecast system over the northeast United States. *Wea. Forecasting,* **22,** 36–55.

Joslyn, S., K. Pak, D. Jones, J. Pyles, and E. Hunt, 2007: The effect of probabilistic information on threshold forecasts. *Wea. Forecasting,* **22,** 804–812.

Kain, J. S., M. E. Baldwin, P. R. Janish, S. J. Weiss, M. P. Kay, and G. W. Carbin, 2003a: Subjective verification of numerical models as a component of a broader interaction between research and operations. *Wea. Forecasting,* **18,** 847–860.

——, P. R. Janish, S. J. Weiss, M. E. Baldwin, R. Schneider, and H. E. Brooks, 2003b: Collaboration between forecasters and

research scientists at the NSSL and SPC: The Spring Program. *Bull. Amer. Meteor. Soc.,* **84,** 1797–1806.

Kalnay, E., 2003: *Atmospheric Modeling, Data Assimilation, and Predictability.* Cambridge University Press, 341 pp.

Katz, R. W., and A. H. Murphy, 1997: *Economic Value of Weather and Climate Forecasts.* Cambridge University Press, 238 pp.

Keith, R., 2003: Optimization of value of aerodrome forecasts. *Wea. Forecasting,* **18,** 808–824.

Lawson, J. D., J. P. Tuell, F. P. Griffith, and R. K. Henry, 2007: Overview of the AWIPS SOA. Preprints, *23rd Int. Conf. on Interactive Information Processing Systems for Meteorology, Oceanography, and Hydrology,* San Antonio, TX, Amer. Meteor. Soc., P2.2.

Legg, T. P., and K. R. Mylne, 2004: Early warnings of severe weather from ensemble forecast information. *Wea. Forecasting,* **19,** 891–906; Corrigendum, **22,** 216–219.

Leith, C. E., 1974: Theoretical skill of Monte Carlo forecasts. *Mon. Wea. Rev.,* **102,** 409–418.

Leutbecher, M., 2005: On ensemble prediction using singular vectors started from forecasts. *Mon. Wea. Rev.,* **133,** 3038–3046.

Magsig, M. A., and E. M. Page, 2003: Weather event simulator implementation and future development. Preprints, *20th Int. Conf. on Interactive Information Processing Systems for Meteorology, Oceanography, and Hydrology,* Long Beach, CA, Amer. Meteor. Soc., 12.9.

Marsigli, C., F. Boccanera, A. Montani, and T. Paccagnella, 2005: The COSMO-LEPS mesoscale ensemble system: Validation of the methodology and verification. *Nonlinear Processes Geophys.,* **12,** 527–536.

Mass, C. F., 2003: IFPS and the future of the National Weather Service. *Wea. Forecasting,* **18,** 75–79.

——, and Coauthors, 2003: Regional environmental prediction over the Pacific Northwest. *Bull. Amer. Meteor. Soc.,* **84,** 1353–1366.

Matthieu, P., and P. Arbogast, 2005: A cyclogenesis evolving into two distinct scenarios and its implications for short-term ensemble forecasting. *Mon. Wea. Rev.,* **133,** 2016–2029.

McEnery, J., J. Ingram, Q. Duan, T. Adams, and L. Anderson, 2005: NOAA's advanced hydrologic prediction service: Building pathways for better science in water forecasting. *Bull. Amer. Meteor. Soc.,* **86,** 375–385.

Morss, R. E., and F. M. Ralph, 2007: Use of information by National Weather Service forecasters and emergency managers during CALJET and PACJET-2001. *Wea. Forecasting,* **22,** 539–555.

——, J. L. Demuth, and J. K. Lazo, 2008: Communicating uncertainty in weather forecasts: A survey of the U.S. public. *Wea. Forecasting,* **23,** 974–991.

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.

NRC, 2006: *Completing the Forecast: Characterizing and Communicating Uncertainty for Better Decisions Using Weather and Climate Forecasts.* National Academies Press, 124 pp.

Palmer, T., and R. Hagedorn, 2006: *Predictability of Weather and Climate.* Cambridge University Press, 718 pp.

Payne, K. A., R. L. Elsberry, and M. A. Boothe, 2007: Assessment of western North Pacific 96- and 120-h track guidance and present forecastability. *Wea. Forecasting,* **22,** 1003–1015.

Pielke, R. A., Jr., 1999: Who decides? Forecasts and responsibilities in the 1997 Red River floods. *Appl. Behav. Sci. Rev.,* **7,** 83–101.

Raftery, A. E., T. Gneiting, F. Balabdaoui, and M. Polakowski,

2005: Using Bayesian model averaging to calibrate forecast ensembles. *Mon. Wea. Rev.,* **133,** 1155–1174.

Roebber, P. J., D. M. Schultz, and R. Romero, 2002: Synoptic regulation of the 3 May 1999 tornado outbreak. *Wea. Forecasting,* **17,** 399–429.

——, ——, B. A. Colle, and D. J. Stensrud, 2004: Toward improved prediction: High-resolution and ensemble modeling systems in operations. *Wea. Forecasting,* **19,** 936–949.

Roulston, M. S., and T. R. Kaplan, 2008: A laboratory-based study of understanding of uncertainty in 5-day site-specific temperature forecasts. *Meteor. Appl.,* in press.

Rutledge, G. K., J. Alpert, and W. Ebisuzaki, 2006: NOMADS: A climate and weather model archive at the National Oceanic and Atmospheric Administration. *Bull. Amer. Meteor. Soc.,* **87,** 327–341.

Santurette, P., and C. G. Georgiev, 2005: *Weather Analysis and Forecasting.* Academic Press, 200 pp.

Sivillo, J. K., J. E. Ahlquist, and Z. Toth, 1997: An ensemble forecasting primer. *Wea. Forecasting,* **12,** 809–818.

Stensrud, D. J., and N. Yussouf, 2003: Short-range ensemble predictions of 2-m temperature and dewpoint temperature over New England. *Mon. Wea. Rev.,* **131,** 2510–2524.

Stuart, N. A., and R. H. Grumm, 2006: Using wind anomalies to forecast East Coast winter storms. *Wea. Forecasting,* **21,** 952–968.

——, and Coauthors, 2006: The future of humans in an increasingly automated forecast process. *Bull. Amer. Meteor. Soc.,* **87,** 1497–1502.

——, D. M. Schultz, and G. Klein, 2007: Maintaining the role of humans in the forecast process: Analyzing the psyche of expert forecasters. *Bull. Amer. Meteor. Soc.,* **88,** 1893–1898.

Toth, Z., and E. Kalnay, 1993: Ensemble forecasting at NMC: The generation of perturbations. *Bull. Amer. Meteor. Soc.,* **74,** 2317–2330.

——, and S. J. Majumdar, 2007: NOAA THORPEX science workshop. *Bull. Amer. Meteor. Soc.,* **88,** 717–719.

Walser, A., M. Arpagaus, C. Appenzeller, and M. Leutbecher, 2006: The impact of moist singular vectors and horizontal resolution on short-range limited-area ensemble forecasts for two European winter storms. *Mon. Wea. Rev.,* **134,** 2877–2887.

Watkin, H., and T. Hewson, 2006: The development of feature-based diagnostics to assess TIGGE handling of high-impact extratropical cyclones. Preprints, *Second THORPEX Int. Science Symp.,* Landshut, Germany, WMO, 258–259.

Weiss, S. J., D. R. Bright, J. S. Kain, J. J. Levit, M. E. Pyle, Z. I. Janjic, B. S. Ferrier, and J. Du, 2006: Complementary use of short-range ensemble and 4.5 km WRF-NMM model guidance for severe weather forecasting at the Storm Prediction Center. Preprints, *23rd Conf. on Severe Local Storms,* St. Louis, MO, Amer. Meteor. Soc., 8.5. [Available online at http://ams.confex.com/ams/pdfpapers/115386.pdf.]

WMO, 2003: 21st status report on the implementation of the World Weather Watch—2003. WMO 957, 61 pp.

——, cited 2007: The Observing System Research and Predictability Experiment (THORPEX) programme. [Available online at http://www.wmo.ch/pages/prog/arep/thorpex/.]

Woodcock, F., and C. Engel, 2005: Operational consensus forecasts. *Wea. Forecasting,* **20,** 101–111.

Yuan, H., X. Gao, S. Mullen, S. Sorooshian, J. Du, and H.-M. H. Juang, 2007: Calibration of probabilistic quantitative precipitation forecasts with an artificial neural network. *Wea. Forecasting,* **22,** 1287–1303.

Yussouf, N., and D. J. Stensrud, 2007: Bias-corrected short-range ensemble forecasts of near-surface variables during the 2005/06 cool season. *Wea. Forecasting,* **22,** 1274–1286.

——, ——, and S. Lakshmivarahan, 2004: Cluster analysis of multimodel ensemble data over New England. *Mon. Wea. Rev.,* **132,** 2452–2462.