NOAA'S FUTURE ENSEMBLE-BASED HURRICANE FORECAST PRODUCTS

BY THOMAS M. HAMILL, MICHAEL J. BRENNAN, BARBARA BROWN, MARK DEMARIA, EDWARD N. RAPPAPORT, AND ZOLTAN TOTH

The rapid maturation of ensemble prediction of tropical cyclones allows the possibility of new uncertainty products.

The chaotic growth of initial condition errors (Lorenz 1996) and model imperfections cause inevitable uncertainty in numerical weather forecasts. In the near future, for tropical cyclone (TC) forecast products, we expect to be able to estimate the uncertainty (i.e., the expected error) and convey how it varies from one storm to the next. For example, instead of providing emergency managers with a forecast that incorporates a standardized 225-km position uncertainty estimate[1] for a 3-day forecast, suppose that we could reliably estimate and convey and uncertainty estimates of 150 km in one circumstance and 300 km in another (Fig. 1). We could thus facilitate informed case-by-case decisions on the extent of coastline evacuation, saving unnecessary evacuations. Users generally prefer receiving uncertainty information rather than having that information hidden from them (Morss et al. 2008), and, when no such uncertainty information is provided, users tend to estimate the uncertainty themselves (Joslyn and Savelli 2010), perhaps inaccurately. Users also tend to make better decisions when reliable uncertainty information is provided (Joslyn et al. 2007; Nadav-Greenberg and Joslyn 2009) and are more willing to use marginally skillful forecasts and become more fault tolerant of forecast misses when the expected error is quantified (LeClerc and Joslyn 2009).

Forecasters and segments of the public use uncertainty products based on ensemble prediction techniques more and more, though there are significant

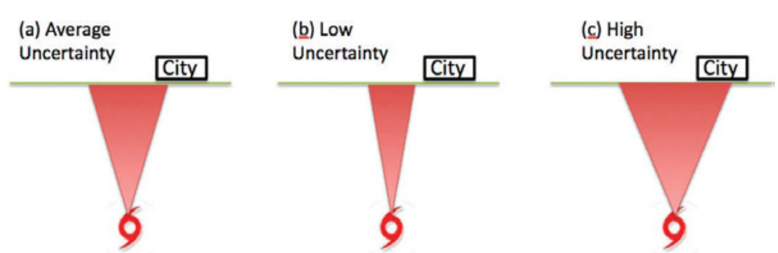

FIG. 1. Conceptual illustration of how situationally dependent uncertainty such as provided by ensembles in track forecasts can improve decision making. (a) The "cone of uncertainty" for a hurricane, estimated using the errors of prior forecasts, provides ambiguous information as to whether to evacuate a city. (b) In this case, a hypothetical calibrated ensemble is suggesting a narrower cone of uncertainty, indicating a decreased threat and implied commensurate actions for the city. (c) Conversely, the uncertainty is much larger and encompasses most of the city, suggesting a potential different course of actions for the city.

[1] An estimate chosen arbitrarily for purpose of illustration.

challenges to more widespread use (Novak et al. 2008; Hirschberg et al. 2011). With ensembles, multiple parallel forecast simulations are typically generated from slightly different initial conditions that reflect the analysis uncertainty. To also estimate the uncertainty contributed by model imperfections, different member forecasts sometimes incorporate stochastic dynamics, different parameterizations, or even different models. Modern ensemble systems demonstrate some ability to quantify the situational forecast uncertainty of TCs (Majumdar and Finocchio 2010; Hamill et al. 2011a,b).

Even before the advent of ensemble prediction, TC forecasts included some probabilistic information. For example, in 1983 the National Weather Service (NWS) implemented quantitative products that provided "strike" probabilities of TC tracks coming within about 60 nautical miles (111 km) of specified coastal locations (Sheets 1985). Beginning in 2006, these were replaced with more general surface wind probability products that included information about the uncertainty in the track, intensity, and wind structure forecasts (e.g., DeMaria et al. 2009). The original track probabilities and the more recent wind probability products are primarily statistically based, where the uncertainty information is estimated from error statistics of operational track, intensity, and structure forecasts from previous years. In 2010, a small step toward incorporating real-time ensemble model-based information was initiated. In particular, the track error distributions that are utilized to generate the probabilities are now being stratified based on the spread of the tracks from several deterministic forecast models (Kidder et al. 2009). These are used in the graphical and text products that provide the probability of 34-, 50-, and 64-kt (~17, 26, and 33 m s^{-1}) winds every 12–120 h (see section 2b and online appendix A for more information).

AFFILIATIONS: Hamill—NOAA/Earth System Research Laboratory/Physical Sciences Division, Boulder, Colorado; Brennan and Rappaport—National Hurricane Center, Miami, Florida; Brown—Research Applications Lab, NCAR, Boulder, Colorado; DeMaria—NOAA/NESDIS, Ft. Collins, Colorado; Toth—NOAA/Earth System Research Laboratory/Global Systems Division, Boulder, Colorado
CORRESPONDING AUTHOR: Thomas M. Hamill, NOAA/ Earth System Research Lab, R/PSD 1, 325 Broadway, Boulder, CO 80305-3328
E-mail: tom.hamill@noaa.gov

The abstract for this article can be found in this issue, following the table of contents.
DOI:10.1175/2011BAMS3106.1

A supplement to this article is available online (10.1175/2011BAMS3106.2)

In final form 12 September 2011

Current statistically based NWS TC probability products provide an accurate measure of the average expected uncertainty in the operational tropical cyclone forecasts. However, the uncertainty estimates do not change much from one forecast situation to the next, as desired. Conceptually, providing situation-dependent uncertainty forecasts based on ensembles appears to be straightforward. However, uncertainty estimates that are formulated directly from current-generation model-based ensembles are sometimes overly confident, offering an unrealistically narrow range of solutions and often biased uncertainty estimates, though the TC ensemble forecasts are improving (Hamill et al. 2011a,b).

Because these ensemble systems are improving and provide increasingly reliable TC uncertainty guidance, TC product developers should be ready to provide visually effective ensemble-based forecast products to forecasters, emergency managers, broadcasters, the media, and other users. What should these ensemble-based products look like? We start by providing some context for the development of ensemble products, discussing some of the scientific challenges related to ensemble prediction of TCs and the current suite of uncertainty products and tools (section 2). We then speculate on some possible ensemble-based products that could be developed in the next several years, especially products relevant to forecasters (section 3). Section 4 provides conclusions. We hope this article will stimulate discussion and will engage the community to think more broadly about how ensemble predictions can be leveraged in the forecast process. We welcome your contributions to this discussion.

THE CURRENT STATE OF PROBABILISTIC PREDICTION, PROBABILISTIC HURRICANE PRODUCTS, AND THEIR USE.

The evolving science of ensemble numerical weather prediction. As specific as we might wish the forecasts to be, the rapid growth of errors in numerical forecasts will inevitably introduce error. These errors grow at different rates, with small scales quicker than large. Hence, heavy rain associated with individual thunderstorms may have predictability[2]

[2] Predictability here means the ability to predict detail of a phenomenon with more specificity than the climatological distribution of that phenomenon.

in a TC for only a fraction of an hour, eye wall size and asymmetries may be predictable for less than a day, and the synoptic-scale environment may be predictable for perhaps a week or two. No matter what model and ensemble system improvements are implemented, the rapid growth of errors in an environment of 3D turbulence may constrain our ability to predict much detail beyond these approximate time ranges (Lorenz 1969; Tribbia and Baumhefner 2004). During the time period when forecast errors are growing but have not yet saturated, ensembles are expected to provide the case-dependent estimates of that uncertainty. What are the challenges to providing reliable estimates of uncertainty from ensemble predictions of TCs?

Defining an ensemble of accurate initial conditions around TCs is one such challenge, and it will require advanced data assimilation techniques. Four-dimensional variational data assimilation (4DVAR; Le Dimet and Talagrand 1986; Courtier et al. 1994; Rabier et al. 2000; Rawlins et al. 2007), the ensemble Kalman filter (EnKF), and hybridizations of EnKFs and variational systems represent the current evolving state of the art in data assimilation and TC initialization. The European Centre for Medium-Range Weather Forecasts (ECMWF) generates its control analysis using 4DVAR and perturbed analyses around the control using parallel cycles of 4DVAR that cycle reduced-resolution ensembles and assimilate distinct perturbed observations (Buizza et al. 2010). ECMWF also computes "diabatic singular vectors" (Puri et al. 2001) to generate perturbations around TCs that may grow rapidly. The U.S. Navy has recently updated their global data assimilation to a 4DVAR algorithm (Xu et al. 2005), and their operational ensemble initialization method was changed to a "banded" ensemble transform technique (McLay et al. 2010), which have both improved TC track forecasts. Other operational centers use a variety of ensemble initialization techniques, mostly not specific to TCs. For example, the National Centers for Environmental Protection (NCEP) currently uses an "ensemble transform with rescaling" (Wei et al. 2008), and the Met Office (UKMO) uses a local ensemble transform Kalman filter (Bowler et al. 2008, 2009).

The National Oceanic and Atmospheric Administration (NOAA) is experimenting with the EnKF (e.g., Houtekamer and Mitchell 1998; Hamill 2006; Hamill et al. 2011a) and its hybridization with the existing 3DVAR for global ensemble initialization (Buehner et al. 2010a,b; Kleist et al. 2009; Hamill et al. 2011b).[3] Nested high-resolution regional model EnKFs are also being developed for TCs (Torn and Hakim 2009; Zhang et al. 2009, 2011; Wu et al. 2010; Torn 2010), with the hope that the regional ensemble systems will, by virtue of their increased resolution, be able to provide more accurate physical depictions of the inner core and thereby storm intensity. The potentially more accurate inner-core estimates may improve the regional assimilation and may improve the subsequent regional ensemble forecasts (Zhang et al. 2009, 2011; Torn 2010).

An equally challenging problem is how to deal in scientifically sound ways with the significant uncertainty introduced by model imperfections. Large-domain, high-resolution ensembles are not currently computationally feasible on current-generation production computers. Hence, some errors are introduced by conducting the ensembles at limited resolution (Gentry and Lackmann 2010) and/or limiting the domain size for regional ensembles (Warner and Hsu 2000; Nutter et al. 2004). Deterministic and ensemble predictions of TCs are sensitive to parameterizations of sea spray (Andreas and Emanuel 2001; Gall et al. 2008), cloud microphysics and aerosols (Wang 2002; Li and Pu 2008), boundary layer processes (Drennan et al. 2007; French et al. 2007), parameters for horizontal mixing length (Bryan et al. 2010), and more. Ensemble predictions inherit deficiencies of deterministic models and the parameterizations they use, so that ensembles as well as deterministic forecasts commonly overforecast of TC genesis and do not predict rapid intensity changes well (Kaplan et al. 2010). NOAA's Hurricane Forecast Improvement Project (HFIP) is helping to address these problems through concerted development of improved parameterizations relevant to the inner-core physics.

Three general approaches can and are being combined to deal with systematic errors in ensemble prediction systems. The first is to improve the forecast systems, perhaps increasing the resolution (Fiorino 2009; Richardson et al. 2010, Fig. 27), improving the fidelity of parameterizations, generating the ensembles with larger or global domains, and appropriately coupling predictive models for the state components such as the ocean and atmosphere. For example, NCEP has recently upgraded the resolution of their global ensemble forecast system to T190

[3] NCEP expects to implement the global hybrid EnKF/3DVAR for purposes of improving the control initial condition during the first half of 2012. Pending more successful testing, the global EnKF perturbations may be used for ensemble initialization operationally somewhat later than that (B. Lapenta 2011, personal communication).

(~95 km grid spacing at 25°N),[4] and the deterministic model was upgraded in summer 2010 to T574 (~32 km grid spacing at 25°N), with upgrades to the many parameterizations. ECMWF recently upgraded their ensemble prediction system to T639 (~28 km at 25°N; Richardson 2010), which significantly reduced TC intensity biases. Several scientists are currently demonstrating the utility of regional ensemble prediction systems for hurricanes (Torn and Hakim 2009; Zhang et al. 2009, 2011; Wu et al. 2010; Torn 2010).

The second general approach is to introduce stochastic effects to represent model uncertainty in a physically realistic fashion. One underlying problem is that many of the parameterizations are formulated deterministically; given the same large-scale input, the same response of subgrid forcing upon the resolved scales is always predicted, even though a range of responses is plausible. Techniques that are being used in operations or are being tested include perturbing the parameterized tendencies with random numbers or structured noise (Buizza et al. 1999; Palmer et al. 2009), using an ensemble system with multiple parameterizations (Charron et al. 2010; Hacker et al. 2011; Berner et al. 2011), introducing stochastic aspects into parameterizations (Lin and Neelin 2002; Teixeira and Reynolds 2008; Plant and Craig 2008), and including "stochastic backscatter" (Shutts 2005; Berner et al. 2009, 2011; Charron et al. 2010). Collectively, this is an area where much more research is needed, because many of the parameterizations in models are still deterministic and many of the approaches taken currently are ad hoc (i.e., they increase spread but not necessarily for scientifically defensible reasons). At this point, most of the research has occurred with global models and is not specific to the problems of TCs.

The third approach is to postprocess. A simple approach may be to combine output from existing models. Goerss (2007) has shown the value of a multimodel consensus, if that is available. This technique has been used, at first subjectively, by forecasters at the National Hurricane Center (NHC) for more than a decade and currently provides among the best automated track guidances (Rappaport et al. 2009). The technique may be more suitable to variables where errors are more random (track) than systematic (intensity).

Another type of postprocessing consists of applying statistical corrections based on discrepancies between past forecasts and observations. Krishnamurti et al. (2006) discussed the "superensemble" concept, combining regression-corrected ensemble guidance. When making statistical corrections, especially for a phenomenon like a TC that may impact a given location only once every 5–10 yr, it is helpful to be able to examine the characteristics of numerical guidance from similar past events. Reforecasts, a database of past forecasts using the same model and assimilation system, can provide just that, compensating for model systematic biases (Hamill et al. 2006). ECMWF now operationally generates a five-member, real-time reforecast once weekly, creating forecasts for the past 18 yr (Hagedorn 2008; Hagedorn et al. 2012). For TCs, they use real-time reforecasts to calibrate their forecasts of TC genesis (Vitart et al. 2010). An even more extensive reforecast dataset, with forecasts more often than once weekly, may be helpful for TC calibration.

Current probabilistic forecast tools and products. Here, we will focus mostly on the probabilistic tools used by and products produced by NOAA's NHC (see online appendix B for more information). We will not consider tools and products from other centers such as the U.S. Navy's Joint Typhoon Warning Center (JTWC), though the spectrum of tools and products should be somewhat similar.

NHC forecasters routinely examine both conventional ensembles and "poor person's" ensembles, with the latter consisting of guidance from a range of deterministic models. Forecasters consider limited-area model output from the Geophysical Fluid Dynamics Lab (GFDL) model (Bender et al. 2007) and the Hurricane Weather Research and Forecast (HWRF) model (Bao et al. 2010). NHC forecasters also examine global deterministic forecasts from the Global Forecast System (GFS), the U.S. Navy's Operational Global Atmospheric Prediction System (NOGAPS; Peng et al. 2004); the Met Office global model (Bowler et al. 2008, 2009) and the ECMWF model (www.ecmwf.int/research/ifsdocs/). Forecasters consider guidance from some ensemble systems, including the NCEP Global Ensemble Forecast System (GEFS), and strike probabilities from the ECMWF ensemble system. NHC forecasters currently pay less attention to ensemble output from other international centers. Users from other organizations find the broader range of real-time ensemble TC information from operational centers of use. These are currently available, for example, through The Observing System Research

[4] Grid spacing calculations here uniformly assume 2N + 1 grid point per latitude circle, where N is the triangular truncation. In fact, some centers like NCEP transform to 3N + 1 grid points. See Hamill et al. (2011b) for more discussion on this.

and Predictability Experiment (THORPEX) International Grand Global Ensemble (TIGGE; Bougeault et al. 2010; for details, see online at http://cawcr.gov.au/projects/THORPEX/TC/index.html; http://tparc.mri-jma.go.jp/cyclone/login.php).

NHC forecasters routinely evaluate several other probabilistic or ensemble-based products. For track, the Goerss predicted consensus error (GPCE; Goerss 2007; see also online appendix B) is used. Given the errors associated with purely dynamical guidance, especially with intensity, NHC forecasters use statistical–dynamical intensity models such as the Statistical Hurricane Intensity Prediction Scheme (SHIPS; DeMaria et al. 2005) and the Logistic Growth Equation Model (LGEM; DeMaria 2009). Because none of the intensity forecast models reliably predicts rapid intensification, NHC forecasters also use a rapid intensity index (RII). The RII uses satellite observations and the NCEP global model forecast to provide a quantitative estimate of the probability of a rapid intensification in the next 24 h.

NHC forecasters produce real-time products to convey the current and forecast location, intensity (wind speed), and size of TCs and their precursors, as well as associated effects (e.g., storm surge). Textual products with uncertainty information (see online appendix A for examples) include a tropical weather outlook, which provides probabilities of TC formation in the next 48 h and a TC discussion providing forecaster reasoning and alternate scenarios based on model guidance diversity. Surface wind speed probabilities are provided in tabular format, based on a Monte Carlo approach that considers on the order of 1,000 realistic track, intensity, and size possibilities (DeMaria et al. 2009). Graphical forecast products include some with probabilistic elements: surface wind speeds, storm surge heights, a coastal watch/warning, and 3- and 5-day cones of uncertainty for TC center position (see online appendix A for examples). Local NWS weather forecast offices provide additional products (not shown).

Emergency management officials do use wind speed probabilities and the uncertainty cones for evacuation decisions, deployment of resources, and briefing elected officials (R. Jennings 2010, personal communication). How much the public uses these products is hard to quantify; all are posted on the NHC website and are used by some segment of the public and by some weather forecasters. We need to interact more with public, both to educate them on how to make better decisions with probabilistic information and to help guide our development of useful, intuitive probabilistic products (Hirschberg et al. 2011).

RECOMMENDATIONS FOR ENSEMBLE-RELATED PRODUCT DEVELOPMENT. We propose some ensemble-related products that could be helpful to emergency managers, the media, and especially forecasters.

Emergency managers and the media are likely to prefer products that will be visually intuitive and that will help them communicate uncertainty information to their audiences. Simple and easily explainable products are preferable because their customers will be more familiar with deterministic products. For example, one modest change might be to enhance the deterministic products with some confidence index (low/medium/high) based upon ensemble uncertainty estimates. For newer and more complex graphics, training may be necessary before users embrace them. Explanatory web pages could be created and associated with the new product web pages. Additionally, through conferences and training facilities such as the Cooperative Program for Operational Meteorology, Education, and Training (COMET), advanced users could be trained in how to interpret the new probabilistic guidance.

Below, we concentrate more on describing potential new uncertainty guidance for forecasters. As opposed to emergency managers, these products are more interpretive, designed to help forecasters understand the potential situational uncertainty rather than to convey it simply to end users. However, many beyond the forecast community may also be interested in such graphics. Should these experimental ensemble-based products be disseminated beyond the ranks of government forecasters, they should include appropriate disclaimers. They may note, for instance, that these products represent experimental guidance that may not produce reliable probabilities and that they should not be considered "official" forecasts.

Intensity products. Because track forecasts have made steady improvement but intensity forecasts have not, forecasters want improved intensity-related uncertainty products first and foremost, including the probability distribution of intensity change. This presents a considerable challenge given the current difficulties that ensemble systems and deterministic models have with predicting intensity change.

Ensemble information could be incorporated into existing statistical models of intensity changes such as the LGEM, providing a value-added ensemble product. LGEM predicts intensity and intensity change based on sea surface temperature (SST; DeMaria and Kaplan 1994; Whitney and Hobgood

1997), forecast vertical shear, convective instability, and translational speed. LGEM could easily be adapted to produce an ensemble of intensity change forecast estimates, one for each forecast track of an ensemble. The LGEM ensemble of intensity change estimates might differ as the ensemble produced a range of atmospheric environment forecast input and as the different forecast tracks positioned the TC over different SSTs. Unfortunately, LGEM ensemble guidance may be biased by ensemble systematic errors. For example, should the mean forecast position be biased, LGEM would use SSTs under the biased track positions, potentially degrading its accuracy.

Were a reforecast dataset available, forecast track statistics could be generated for a long period of time from a stable model. This approach may permit systematic track errors to be estimated and corrected from real-time forecasts. It may also be possible to train a statistical model like LGEM not using analysis data (a "perfect prog" approach) but rather using forecast data (a "model output statistics" approach) and in this way account for potential biases in forecasts of the environmental information.

Other products could also convey predictive information related to intensity to forecasters. Figure 2 provides a synthetic example of how a common spaghetti plot of tracks could be augmented with intensity-related information. Included on this plot is analyzed SST. The arrows associated with each track position represent the associated vector 850–200-hPa wind shear. Also displayed are the forecast central pressures rounded to the nearest hPa. A visual display like this permits a forecaster to see the interactions of several variables that may be related to intensity. In this case, for example, the wind shear is somewhat weaker and the SST warmer at 48 h to the southwest, suggesting that, if the storm's actual course is southwest of the predicted mean, it may be stronger than a storm moving northeast of the mean. In this case, the intensity difference is also reflected in the model–forecast central pressure estimates.

"Meteograms" of the ensemble distribution could also be produced for intensity-related storm parameters, such as shear, midlevel moisture, instability, and maximum wind speed, as illustrated in Fig. 3. Another possibility would be the creation of ensemble-derived probability maps of important variables, which might depict, for example, the probabilities of exceedance of critical variables, such as RH < 50% at 500 hPa or 850–200-hPa vertical shear > 15 kt (~7.7 m s^{-1}). Finally, were a reforecast dataset available, an "extreme forecast index" (LaLaurette 2003) could be created for intensity forecasts, representing how the ensemble of forecast intensities compare to the climatological distribution of forecast intensities.

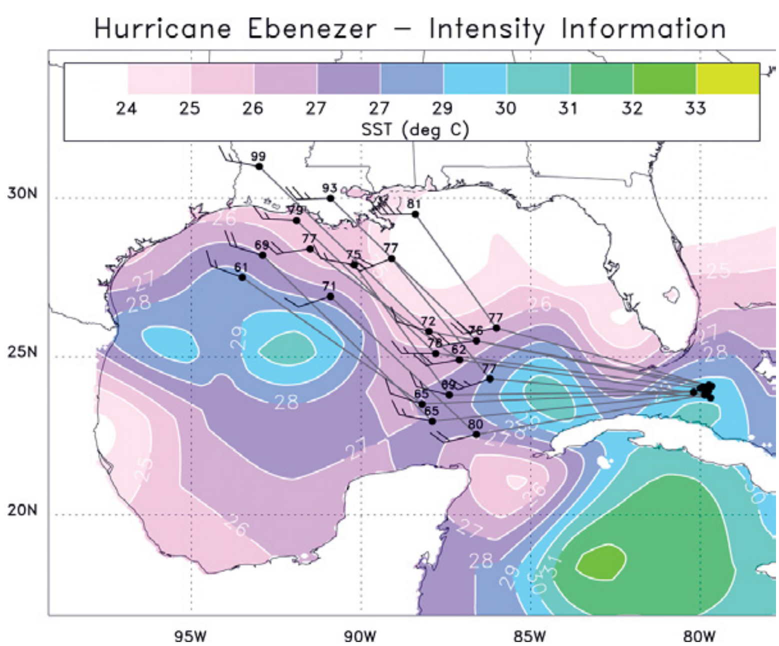

FIG. 2. Illustration of a potential display of intensity-related information. A (synthetic) 10-member ensemble forecast of tracks is shown, with 0-, 24-, and 48-hour positions denoted by the dots. The underlying contours denote the analyzed SST; the wind barbs denote the vertical (200–850 hPa) wind shear, measured in knots (1 kt = 0.5144 ms^{-1}). A full wind barb equals 10 knots, a half barb equals 5 knots. The two numbers plotted over top of the cyclone position denote the model-forecast mean sea level central pressure, measured in hPa minus 900.

Tropical cyclogenesis products. Forecasters also desire new products to help them predict the relative likelihood of TC genesis. However, given the common misestimation of genesis frequency in numerical models, the potential skill of such products may be marginal, at least until reforecasts are available to estimate the climatological frequency of genesis from past forecasts. Recently, ECMWF staff has developed experimental guidance for the probability of tropical cyclogenesis; these are calibrated using the forecast climatology derived from their 20-yr, once-weekly reforecast dataset (D. Richardson 2010, personal communication; Hagedorn

2008). Figure 4 illustrates how ECMWF represented the relative likelihood of TC genesis in several basins in their real-time forecast compared to the frequency of genesis determined from past forecasts.

Even without reforecasts, forecasters may still find ensemble-based guidance of TC genesis informative. Ensemble tracks of model-generated storms can be plotted, and genesis probabilities can be estimated from the ensemble within specific geographical regions. Another possibility is to examine whether a statistical model of genesis (Schumacher et al. 2009) incorporating ensemble genesis information might provide skillful guidance.

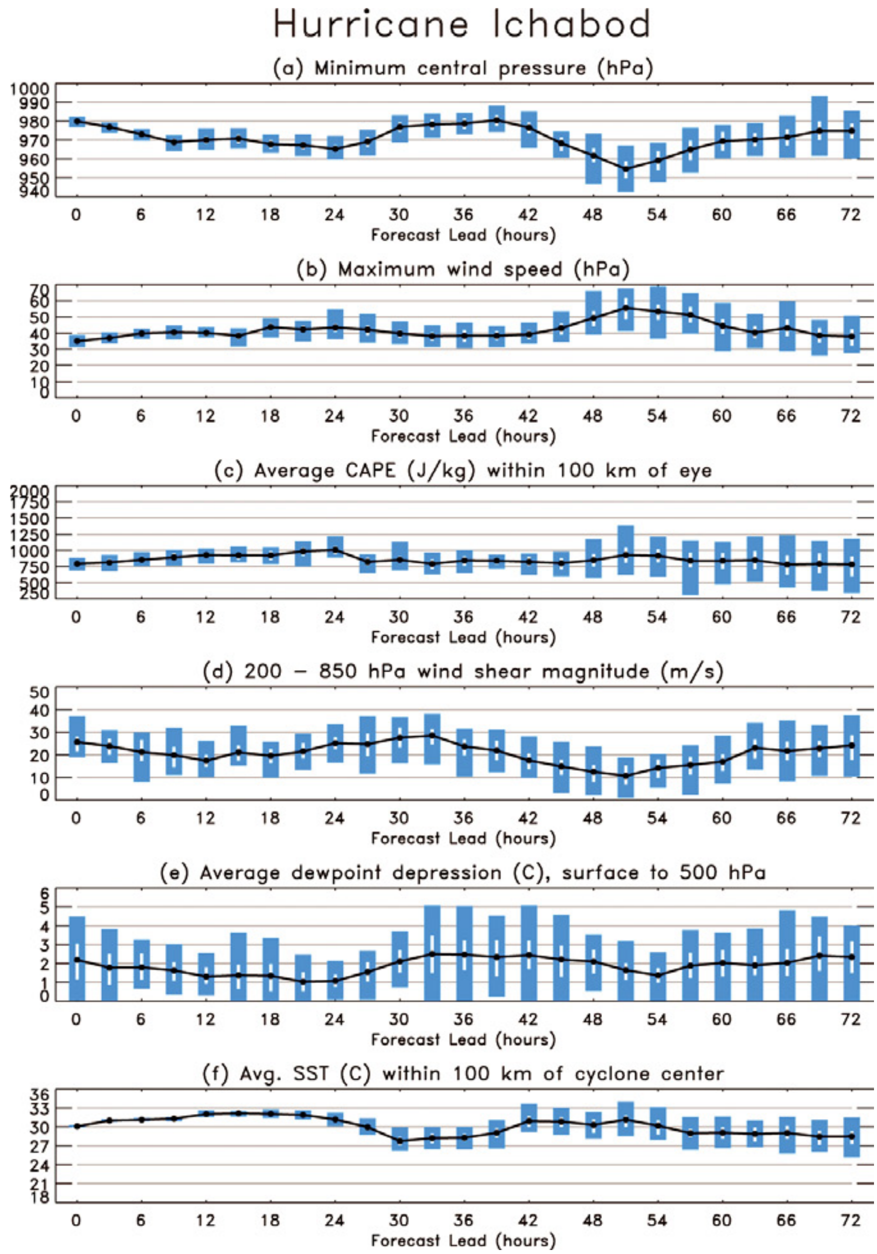

Fig. 3. Synthetic example of a potential meteogram-style plot of intensity-related information from an ensemble system. Each panel provides the ensemble-mean forecast (solid line and dots), whereas the blue bars denote the lowest and highest values from the ensemble. White lines denote the 20th and 80th percentiles of the distribution.

Structure products. Uncertainty information is desired for storm structure as well, including storm size and average wind radii at various thresholds in different storm quadrants. Possible products include 1) ensemble averages of the radius of the outermost closed isobar (OCI), which provides one possible measure of overall storm size, which is relevant for shipping and storm surge; 2) ensemble-mean predictions of 34-, 50-, and 64-kt (15.4, 25.7, and 32.9 m s^{-1}) wind radii in different quadrants, where the ensemble of storms is relocated to a common position; and 3) probability distributions of the OCI and wind radii.

Track products. In addition to the possibilities indicated in the "Intensity products" section above for augmenting track information with intensity-related variables, other track-related products may be useful. Figure 5 shows how track information could incorporate ensemble data from previous forecast cycles as well as cones of uncertainty that are no longer purely circular. Such a plot could be further enhanced if the older forecasts had been subjected to quality control to eliminate or deweight prior forecasts that did not match recent observed positions. In order to identify the dominant synoptic patterns associated with different track forecasts, a "cluster analysis" can be used to distinguish between dominant track paths, followed by the creation of a composite of the associated forecast fields in each cluster.

Associated phenomena: Storm surge, winds, rainfall, and tornadoes. As ensembles are improved and perhaps postprocessed, they may provide more reliable track, intensity, and storm size estimates. These in turn may permit probabilistic guidance to be generated for many important storm-related effects. For example, storm surge probabilities might be generated by driving surge models like the Sea, Lake, and Overland Surges from Hurricanes (SLOSH; Houston et al. 1999) using ensemble-based guidance. Postprocessed ensemble guidance can provide improved probabilistic estimates of precipitation (Hamill and Whitaker 2006) from landfalling TCs or from predecessor rain events (PREs; Galarneau et al. 2010). Statistical models can be formulated to estimate tornado probabilities based on the environmental characteristics. Additional probabilistic forecasts can be generated for winds above critical thresholds such as tropical storm or hurricane strength and can provide information on the timing (onset and duration).

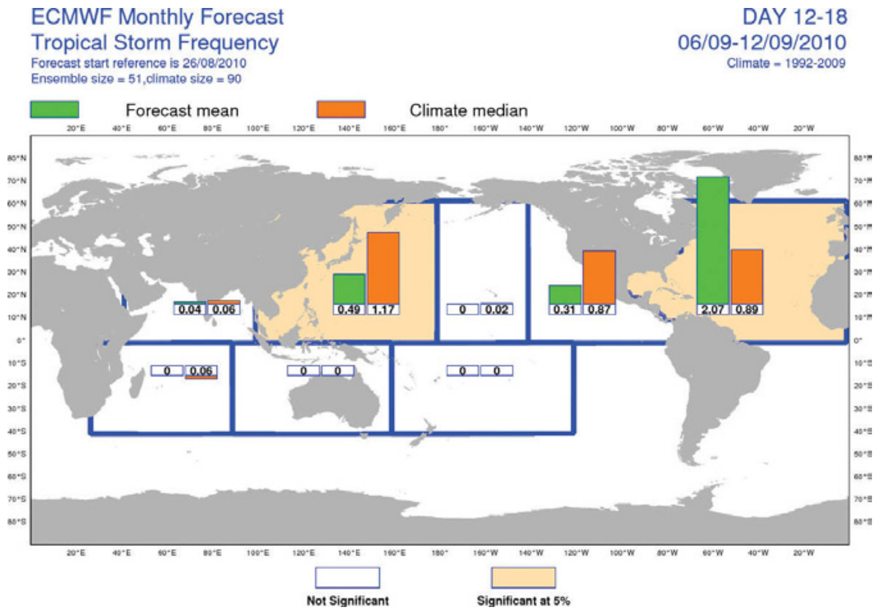

FIG. 4. An example of ECMWF's extended-range forecast of tropical cyclogenesis, here for the week of 6–12 Sep 2010. Activity is monitored separately in the various basins, separated by the blue lines. The green bars show the mean number of TCs predicted to develop during the week 6–12 Sep (i.e., in each basin count all new TC geneses in each ensemble members in that 7-day period and calculate the mean of that number over the ensemble; the actual number is shown below the green bar). The orange bar (and number) is determined from the reforecast climate at this time of the year.

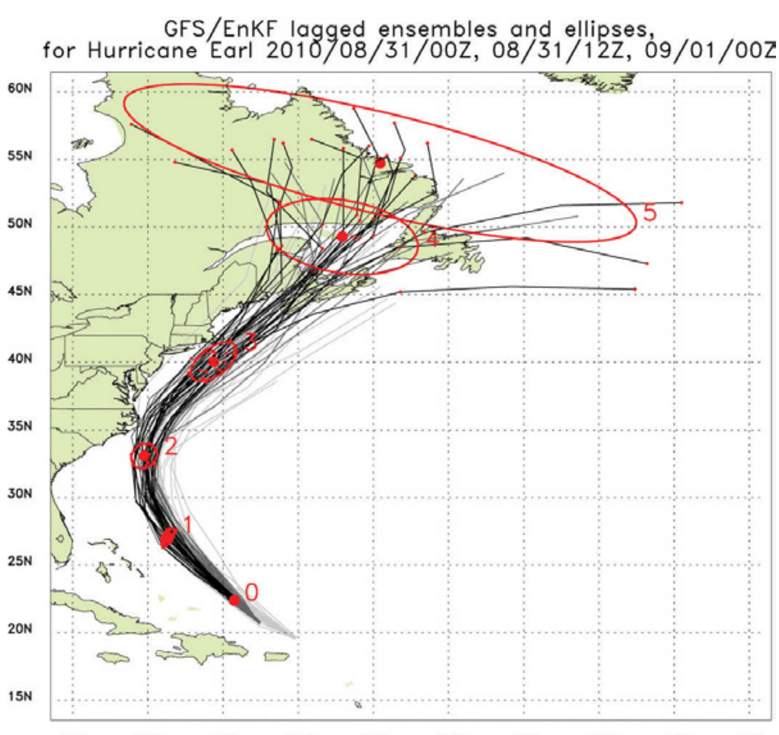

FIG. 5. Lagged ensemble track forecasts for Hurricane Earl (2010) from the GFS/EnKF (see Hamill et al. 2011a). Light grey lines denote track forecasts initialized at 0000 UTC 31 Aug 2010. Darker grey lines denote track forecasts initialized at 1200 UTC 31 Aug 2010. Black lines denote track forecasts initialized at 0000 UTC 1 Sep 2010. Red dots denote positions of ensemble members, with the large red dot the position of the ensemble-mean member. The ellipse is from a fitted bivariate normal distribution, with the contour enclosing 90% of the fitted probability (ibid.). Numbers indicate the forecast lead in days corresponding to the 1 Sep forecast.

Guidance for locations of supplemental observations. Ensemble-based techniques may be useful for determining the most useful locations for supplemental observations (Aberson 2010). The observations may be dropsondes from reconnaissance aircraft or, in the future, the processing of higher-density cloud-track winds or satellite radiances. Typically, an ensemble-based targeting algorithm considers how much the forecast uncertainty would be reduced as a result of the reduction in analysis-error variance in an ensemble due to the assimilation of extra data. These techniques are promising but are limited in their utility by the lack of calibration of the ensemble and the assumption of linear error growth (Majumdar et al. 2006; Reynolds et al. 2010).

CONCLUSIONS. Studies indicate that users are able to make better decisions when provided with relevant uncertainty information. Ensemble prediction techniques that generate such information for TCs are improving but are still affected by systematic errors, such as underforecasting intensity in global models. As ensemble systems improve, the weather forecast community is planning for how to incorporate more ensemble-based uncertainty information into the forecast process and how TC uncertainty can be conveyed to forecasters, emergency managers, the media, and the public. This article discussed some preliminary ideas for how ensemble information can be interpreted and used more extensively. Special attention was paid to new ensemble products for intensity forecasting, which has not improved as much as track forecasts over the past few decades. Our product recommendations assume that forecasters will be interested in products that are more diagnostic in nature and can provide additional information about the range of possible outcomes. Emergency managers and the media may be more interested in user-friendly graphics that could be understood readily by their customers. We invite you to join the discussion of how ensemble data can be leveraged to improve hurricane guidance.

ACKNOWLEDGMENTS. This article summarizes the HFIP-THORPEX Ensemble Product Development Workshop, which was held 20–21 April 2010 in Boulder, Colorado. About 50 ensemble prediction experts, meteorologists with expertise in statistics, hurricane forecasters, and representatives from the emergency management and social-science communities attended the workshop, in person or over the phone. The workshop was sponsored by 1) NOAA's The Observing System, Research, and Predictability Experiment (THORPEX) program, which has sponsored much of NOAA's ensemble prediction research; 2) the NOAA Hurricane Forecast Improvement Project (HFIP), which seeks to halve track and intensity errors in the next 10 yr through a coordinated research and development and transition to operations program on hurricanes; and 3) the workshop host, the National Center for Atmospheric Research (NCAR) Tropical Cyclone Modeling Testbed.

REFERENCES

Aberson, S. D., 2010: Ten years of hurricane synoptic surveillance (1997–2006). *Mon. Wea. Rev.,* **138,** 1536–1549.

Andreas, E. L, and K. A. Emanuel, 2001: Effects of sea spray on tropical cyclone intensity. *J. Atmos. Sci.,* **58,** 3741–3750.

Bao, S., R. Yablonsky, D. Stark, and L. Bernardet, 2010: HWRF user's guide. The Developmental Testbed Center Rep., 88 pp. [Available online at www.dtcenter.org/HurrWRF/users/docs/users_guide/HWRF_usersguide_2-9B-10_cm_lrb.pdf.]

Bender, M. A., I. Ginis, R. Tuleya, B. Thomas, and T. Marchok, 2007: The operational GFDL coupled hurricane–ocean prediction system and a summary of its performance. *Mon. Wea. Rev.,* **135,** 3965–3989.

Berner, J., G. Shutts, M. Leutbecher, and T. N. Palmer, 2009: A spectral stochastic kinetic energy backscatter scheme and its impact on flow-dependent predictability in the ECMWF ensemble prediction system. *J. Atmos. Sci.,* **66,** 603–626.

——, S.-Y. Ha, J. P. Hacker, A. Fournier, and C. Snyder, 2011: Model uncertainty in a mesoscale ensemble prediction system: Stochastic versus multiphysics representations. *Mon. Wea. Rev.,* **139,** 1972–1995.

Bougeault, P., and Coauthors, 2010: The THORPEX interactive grand global ensemble. *Bull. Amer. Meteor. Soc.,* **91,** 1059–1072.

Bowler, N., 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.

——, ——, S. E. Beare, K. R. Mylne, and G. J. Shutts, 2009: The local ETKF and SKEB: Upgrades to the MOGREPS short-range ensemble prediction system. *Quart. J. Roy. Meteor. Soc.,* **135,** 767–776.

Bryan, G. H., R. Rotunno, and Y. Chen, 2010: The effects of turbulence on hurricane intensity. Preprints, *29th Conf. on Hurricanes and Tropical Meteorology,* Tucson, AZ, Amer. Meteor. Soc., 8C.7.

Buehner, M., P. L. Houtekamer, C. Charette, H. L. Mitchell, and B. He, 2010a: Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part I: Description of configurations and results from idealized experiments. *Mon. Wea. Rev.,* **138,** 1550–1566.

——, ——, ——, ——, and ——, 2010b: Intercomparison of variational data assimilation and the ensemble Kalman filter for global deterministic NWP. Part II: Results from 1-month experiments with real observations. *Mon. Wea. Rev.,* **138,** 1567–1586.

Buizza, R., M. J. Miller, and T. N. Palmer, 1999: Stochastic simulation of model uncertainties in the ECMWF ensemble prediction system. *Quart. J. Roy. Meteor. Soc.,* **125,** 2887–2908.

——, M. Leutbecher, L. Isaksen, and J. Haseler, 2010: Combined use of EDA and SV-based perturbations in the EPS. *ECMWF Newsletter,* No. 123, ECMWF, Reading, United Kingdom, 22–28.

Charron, M., L. Spacek, P. L. Houtekamer, H. L. Mitchell, L. Michelin, G. Pellerin, and N. Gagnon, 2010: Toward random sampling of model error in the Canadian ensemble prediction system. *Mon. Wea. Rev.,* **138,** 1877–1901.

Courtier, P., J.-N. Thepaut, and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-Var, using an incremental approach. *Quart. J. Roy. Meteor. Soc.,* **120,** 1367–1387.

DeMaria, M., 2009: A simplified dynamical system for tropical cyclone intensity prediction. *Mon. Wea. Rev.,* **137,** 68–82.

——, and J. Kaplan, 1994: Sea surface temperature and the maximum intensity of Atlantic tropical cyclones. *J. Climate,* **7,** 1324–1334.

——, M. Mainelli, L. K. Shay, J. A. Knaff, and J. Kaplan, 2005: Further improvements to the statistical hurricane intensity prediction scheme (SHIPS). *Wea. Forecasting,* **20,** 531–543.

——, J. A. Knaff, R. Knabb, C. Lauer, C. R. Sampson, and R. T. DeMaria, 2009: A new method for estimating tropical cyclone wind speed probabilities. *Wea. Forecasting,* **24,** 1573–1591.

Drennan, W. M., J. A. Zhang, J. R. French, C. McCormick, and P. G. Black, 2007: Turbulent fluxes in the hurricane boundary layer. Part II: Latent heat flux. *J. Atmos. Sci.,* **64,** 1103–1115.

Fiorino, M., 2009: Record-setting performance of the ECMWF IFS in medium-range tropical cyclone track prediction. *ECMWF Newsletter,* No. 118, ECMWF, Reading, United Kingdom, 20–26.

French, J. R., W. M. Drennan, J. A. Zhang, and P. G. Black, 2007: Turbulent fluxes in the hurricane boundary layer. Part I: Momentum flux. *J. Atmos. Sci.,* **64,** 1089–1102.

Galarneau, T. J., Jr., L. F. Bosart, and R. S. Schumacher, 2010: Predecessor rain events ahead of tropical cyclones. *Mon. Wea. Rev.,* **138,** 3272–3297.

Gall, J. S., W. M. Frank, and Y. Kwon, 2008: Effects of sea spray on tropical cyclones simulated under idealized conditions. *Mon. Wea. Rev.,* **136,** 1686–1705.

Gentry, M. S., and G. M. Lackmann, 2010: Sensitivity of simulated tropical cyclone structure and intensity to horizontal resolution. *Mon. Wea. Rev.,* **138,** 688–704.

Goerss, J. S., 2007: Prediction of consensus tropical cyclone track forecast error. *Mon. Wea. Rev.,* **135,** 1985–1993.

Hacker, J. P., S.-Y. Ha, C. Snyder, J. Berner, F. A. Eckel, M. Pocernich, J. Schramm, and X. Wang, 2011: The U.S. Air Force Weather Agency's mesoscale ensemble: Scientific description and performance results. *Tellus,* **63A,** 625–641.

Hagedorn, R., 2008: Using the ECMWF reforecast data set to calibrate EPS forecasts. *ECMWF Newsletter,* No. 117, ECMWF, Reading, United Kingdom, 8–13.

——, R. Buizza, T. M. Hamill, M. Leutbecher, and T. N. Palmer, 2012: Comparing TIGGE multi-model forecasts with reforecast-calibrated ECMWF ensemble forecasts. *Quart. J. Roy. Meteor. Soc.,* in press.

Hamill, T. M., 2006: Ensemble-based atmospheric data assimilation. *Predictability of Weather and Climate,* Cambridge Press, 124–156.

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

——, ——, and S. L. Mullen, 2006: Reforecasts: An important dataset for improving weather predictions. *Bull. Amer. Meteor. Soc.,* **87,** 33–46.

——, ——, M. Fiorino, and S. J. Benjamin, 2011a: Global ensemble predictions of 2009's tropical cyclones initialized with an ensemble Kalman filter. *Mon. Wea. Rev.,* **139,** 668–688.

——, ——, D. T. Kleist, M. Fiorino, and S. G. Benjamin, 2011b: Predictions of 2010's tropical cyclones using the GFS and ensemble-based data assimilation methods. *Mon. Wea. Rev.,* **139,** 3243–3247.

Hirschberg, P., and Coauthors, 2011: A weather and climate enterprise strategic implementation plan for generating and communicating forecast uncertainty information. *Bull. Amer. Meteor. Soc.,* **92,** 1651–1666.

Houston, S. H., W. A. Shaffer, M. D. Powell, and J. Chen, 1999: Comparisons of HRD and SLOSH surface wind fields in hurricanes: Implications for storm surge modeling. *Wea. Forecasting,* **14,** 671–686.

Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.,* **126,** 796–811.

Joslyn, S., and S. Savelli, 2010: Communicating forecast uncertainty: Public perception of weather forecast uncertainty. *Meteor. Appl.,* **17,** 180–195.

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

Kaplan, J., M. DeMaria, and J. A. Knaff, 2010: A revised tropical cyclone rapid intensification index for the Atlantic and eastern North Pacific basins. *Wea. Forecasting,* **25,** 220–241.

Kidder, S. Q., M. DeMaria, and P. Harr, 2009: An improved wind probability estimation program. Joint Hurricane Testbed Project Final Rep., 10 pp. [Available online at www.nhc.noaa.gov/jht/07-09reports/final_Kidder_JHT09.pdf.]

Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, W.-S. Wu, and S. J. Lord, 2009: Introduction of the GSI into the NCEP global data assimilation system. *Wea. Forecasting,* **24,** 1691–1705.

Krishnamurti, T. N., T. S. V. Vijaya Kumar, W.-T. Yun, A. Chakraborty, and L. Stefanova, 2006: Weather and seasonal climate forecasts using the superensemble approach. *Predictability of Weather and Climate,* T. N. Palmer and R. Hagedorn, Eds., Cambridge University Press, 532–560.

LaLaurette, F., 2003: Early detection of abnormal weather conditions using a probabilistic extreme forecast index. *Quart. J. Roy. Meteor. Soc.,* **129,** 3037–3057.

LeClerc, J., and S. Joslyn, 2009: Role of uncertainty information and forecast error in weather-related decision making. *Proc. Annual Meeting of the Psychonomics Society,* Boston, MA, Psychonomics Society.

Le Dimet, F.-X., and O. Talagrand, 1986: Variational algorithms for analysis and assimilation of meteorological observations: Theoretical aspects. *Tellus,* **38A,** 97–110.

Li, X., and Z. Pu, 2008: Sensitivity of numerical simulation of early rapid intensification of Hurricane Emily (2005) to cloud microphysical and planetary boundary layer parameterizations. *Mon. Wea. Rev.,* **136,** 4819–4838.

Lin, J. W.-B., and J. D. Neelin, 2002: Considerations for stochastic convective parameterization. *J. Atmos. Sci.,* **59,** 959–975.

Lorenc, A. C., 1981: A global three-dimensional multivariate statistical interpolation system. *Mon. Wea. Rev.,* **109,** 701–721.

Lorenz, E. N., 1969: The predictability of a flow which possesses many scales of motion. *Tellus,* **21,** 289–307.

——, 1996: *The Essence of Chaos.* University of Washington Press, 227 pp.

Mainelli, M., M. DeMaria, L. K. Shay, and G. Goni, 2008: Application of oceanic heat content estimation to operational forecasting of recent Atlantic category-5 hurricanes. *Wea. Forecasting,* **23,** 3–16.

Majumdar, S. J., and P. M. Finocchio, 2010: On the ability of global ensemble prediction systems to predict tropical cyclone track probabilities. *Wea. Forecasting,* **25,** 659–680.

——, S. D. Aberson, C. H. Bishop, R. Buizza, M. S. Peng, and C. A. Reynolds, 2006: A comparison of adaptive observing guidance for Atlantic tropical cyclones. *Mon. Wea. Rev.,* **134,** 2354–2372.

McLay, J., C. H. Bishop, and C. A. Reynolds, 2010: A local formulation of the ensemble transform (ET) analysis perturbation scheme. *Wea. Forecasting,* **25,** 985–993.

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

Nadav-Greenberg, L., and S. Joslyn, 2009: Uncertainty forecasts improve decision making among nonexperts. *J. Cognit. Eng. Decis. Making,* **2,** 24–47.

Novak, D. R., D. R. Bright, and M. J. Brennan, 2008: Operational forecaster uncertainty needs and future roles. *Wea. Forecasting,* **23,** 1069–1084.

Nutter, P., D. J. Stensrud, and M. Xue, 2004: Effects of coarsely resolved and temporally interpolated lateral boundary conditions on the dispersion of limited-area ensemble forecasts. *Mon. Wea. Rev.,* **132,** 2358–2377.

Palmer, T. N., R. Buizza, F. Doblas-Reyes, T. Jung, M. Leutbecher, G. J. Shutts, M. Steinheimer, and A. Weisheimer, 2009: Stochastic parametrization and uncertainty. ECMWF Tech Memo. 598, 42 pp. [Available online at www.ecmwf.int/publications/library/ecpublications/_pdf/tm/501-600/tm598.pdf.]

Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Center spectral statistical interpolation analysis system. *Mon. Wea. Rev.,* **120,** 1747–1763.

Peng, M., J. A. Ridout, and T. F. Hogan, 2004: Recent modifications of the Emanuel convective scheme in the Naval Operational Global Atmospheric Prediction System. *Mon. Wea. Rev.,* **132,** 1254–1268.

Plant, R. S., and G. C. Craig, 2008: A stochastic parameterization for deep convection based on equilibrium statistics. *J. Atmos. Sci.,* **65,** 87–105.

Puri, K., J. Barkmeijer, and T. N. Palmer, 2001: Ensemble prediction of tropical cyclones using targeted diabatic singular vectors. *Quart. J. Roy. Meteor. Soc.,* **127,** 709–731.

Rabier, F., H. Jarvinen, E. Klinker, J.-F. Mahfouf, and A. Simmons, 2000: The ECMWF operational implementation of four-dimensional variational assimilation: I: Experimental results with simplified physics. *Quart. J. Roy. Meteor. Soc.,* **126,** 1143–1170.

Rappaport, E. N., and Coauthors, 2009: Advances and challenges at the National Hurricane Center. *Wea. Forecasting,* **24,** 395–419.

Rawlins, F., S. P. Ballard, K. J. Bovis, A. M. Clayton, D. Li, G. W. Inverarity, A. C. Lorenc, and T. J. Payne, 2007: The Met Office global four-dimensional variational data assimilation scheme. *Quart. J. Roy. Meteor. Soc.,* **133,** 347–362.

Reynolds, C. A., J. D. Doyle, R. M. Hodur, and H. Jin, 2010: Naval Research Laboratory multiscale targeting guidance for T-PARC and TCS-08. *Wea. Forecasting,* **25,** 526–544.

Richardson, D., 2010: Changes to the operational forecasting system. *ECMWF Newsletter,* No. 122, ECMWF, Reading, United Kingdom, 3.

——, J. Bidlot, L. Ferranti, A. Ghelli, T. Hewson, M. Janousek, F. Prates, and F. Vitart, 2010: Verification statistics and evaluations of ECMWF forecasts in 2009–2010. ECMWF Tech. Memo. 635, 45 pp.

Schumacher, A. B., M. DeMaria, and J. A. Knaff, 2009: Objective estimation of the 24-h probability of tropical cyclone formation. *Wea. Forecasting,* **24,** 456–471.

Sheets, R. C., 1985: The National Weather Service Hurricane Probability Program. *Bull. Amer. Meteor. Soc.,* **66,** 4–13.

Shutts, G., 2005: A kinetic energy backscatter algorithm for use in ensemble prediction systems. *Quart. J. Roy. Meteor. Soc.,* **131,** 3079–3102.

Teixeira, J., and C. A. Reynolds, 2008: Stochastic nature of physical parameterizations in ensemble prediction: A stochastic convection approach. *Mon. Wea. Rev.,* **136,** 483–496.

Torn, R. D., 2010: Performance of a mesoscale ensemble Kalman filter (ENKF) during the NOAA high-resolution hurricane test. *Mon. Wea. Rev.,* **138,** 4375–4392.

——, and G. J. Hakim, 2009: Ensemble data assimilation applied to RAINEX observations of Hurricane Katrina (2005). *Mon. Wea. Rev.,* **137,** 2817–2829.

Tribbia, J. J., and D. P. Baumhefner, 2004: Scale interactions and atmospheric predictability: An updated perspective. *Mon. Wea. Rev.,* **132,** 703–713.

Vitart, F., A. Leroy, and M. C. Wheeler, 2010: A comparison of dynamical and statistical predictions of weekly tropical cyclone activity in the Southern Hemisphere. *Mon. Wea. Rev.,* **138,** 3671–3682.

Wang, Y., 2002: An explicit simulation of tropical cyclones with a triply nested movable mesh primitive equation model: TCM3. Part II: Model refinements and sensitivity to cloud microphysics parameterization. *Mon. Wea. Rev.,* **130,** 3022–3036.

Warner, T. T., and H.-M. Hsu, 2000: Nested-model simulation of moist convection: The impact of coarse-grid parameterized convection on fine-grid resolved convection. *Mon. Wea. Rev.,* **128,** 2211–2231.

Wei, M., Z. Toth, R. Wobus, and Y. Zhu, 2008: Initial perturbations based on the ensemble transform (ET) technique in the NCEP global operational forecast system. *Tellus,* **60A,** 62–79.

Wu, C.-C., G.-Y. Lien, J.-H. Chen, and F. Zhang, 2010: Assimilation of tropical cyclone track and structure based on the ensemble Kalman filter (EnKF). *J. Atmos. Sci.,* **67,** 3806–3822.

Xu, L., T. Rosmond, and R. Daley, 2005: Development of NAVDAS-AR: Formulation and initial tests of the linear problem. *Tellus,* **57A,** 546–559.

Zhang, F., Y. Weng, J. A. Sippel, Z. Meng, and C. H. Bishop, 2009: Cloud-resolving hurricane initialization and prediction through assimilation of Doppler radar observations with an ensemble Kalman filter: Humberto (2007). *Mon. Wea. Rev.,* **137,** 2105–2125.

——, ——, J. F. Gamache, and F. D. Marks, 2011: Performance of cloud-resolving hurricane initialization and prediction during 2008–2010 with ensemble data assimilation of inner-core airborne Doppler radar observations. *Geophys. Res. Lett.,* **38,** L15810, doi:10.1029/2011GL048469.