Digital Forecast Process

ER Predictability Team

Team Members: Dan Cobb, Rich Grumm, Josh Korotky, David Novak, and Paul Sisson

Statement of Question

Within the limits of predictability, what are the optimal spatial and temporal resolutions needed to provide a useful and versatile digital service, while maintaining scientific validity?

Introduction

Feedback on the question was solicited by the **I**FPS **S**cience **S**teering **T**eam (ISST). The question was considered by a group of Eastern Region SOOs and Regional Scientific Services Division members over the course of four months. This report addresses the central question of whether current NDFD horizontal and temporal resolutions are justifiable in terms of intrinsic atmospheric predictability limits. The team made several findings and recommendations that would partially validate current spatial and temporal resolutions.

Additionally, the team engaged in several healthy discussions concerning alternative approaches which centered on addressing the validity of National Digital Forecast Database (NDFD) forecast process from a broader perspective. These discussions were important in that they framed the context of our recommendations. Brief outlines of the following points of context are included in this report: whether population of NDFD could be less labor intensive, whether graphical editing procedures add value, a broader discussion of predictability and scales, and the need to address the status of NDFD in terms of one of its primary goals – the ability of the user to "fit the forecast to their problem".

Current NDFD Scales

NDFD is currently at a 5 km horizontal resolution nationwide at all forecast projections out to 7 days. The temporal resolution is 3 hourly out to forecast projections of 72 h, and 6 hourly beyond 72 h. Several local offices use spatial resolution to 2.5 km (some to 1.25 km) and their temporal resolution as low as 1 h. The resolutions have been primarily driven internally and are primarily the result of a need to resolve topographic differences in temperature and wind.

Fine resolution is critical in the western U.S. where the majority of the population live in narrow valleys and canyons. Similar logic applies to coastal cities where land sea differences are considerable. If these topographic features are not resolvable in the NDFD then providing useful information for such locations in a gridded format is not possible, resulting in a degraded forecast from those manually produced text and automated Model Output Statistics (MOS) point forecasts used previously.

However, spatial resolutions of the NDFD are significantly finer than currently available Numerical Weather Prediction (NWP) guidance (Eta/DGEX 12 km, GFS 60 km, and RUC 20 km.) In order to compensate for the disparity between available and needed resolutions, various downscaling methods have been developed to synthesize such data. Again, these tools focus on producing topographic representations of temperature, wind, and to lesser degrees Probability of Precipitation (POP), Quantitative Precipitation Forecast (QPF), and precipitation type/character.

Temporal scales are apparently arbitrary. It is not known why the Graphical Forecast Editor (GFE) baseline was set to hourly resolution out to seven days or on what basis NDFD resolution varies from 3 hourly from F00 to F72 and 6 hourly from F78 to F168. Here again, only limited amounts of NWP guidance are delivered at 3 hourly resolution to support the first 72 hours.

Predictability

The atmosphere is a chaotic dynamic system in which two nearly identical initial states, each evolving according to the same physical laws, will develop into final states that bear no resemblance at all to one another (Lorenz, 1993). Thus our ability to make detailed forecasts is often severely restricted by even small amounts of initial condition uncertainty, and is further limited by the inability of NWP to precisely model many of the physical laws governing the atmosphere.

In terms of NWP, divergence due to initial condition uncertainty (e.g., sensitivity to initial conditions) refers to the growth of forecasts errors with time. NWP errors begin with small scale features and propagate in an upscale fashion. As a result, predictability varies with the scale of the feature being forecast, decreasing as scale decreases (Dalcher and Kalnay 1987 and Droegmeier 1997). In general, predictive skill falls away quickly for small scale features like thunderstorms and more gradually for synoptic scale features like mid latitude storms. Table 1 presents a subjective temporal scale of predictability for some basic atmospheric scales and NDFD related variables.

Table 1: Subjective skill of a forecast for different scales of phenomena versus time. (adapted from http://www.ecmwf.int/products/forecasts/guide/Scale_and_predictability.html ECMWF 2005)

Feature/Variable	***< Day 1***	***Day 1 - 2***	***Day 3-5***	***Day 6-7***	***Day 7-10***
Hemispheric flow transitions	Excellent	Excellent	Very Good	Good	Fair
Cyclone life cycle	Excellent	Very Good	Good	Fair	----
Fronts	Excellent	Good	Fair	----	----
Mesoscale: Banded structures/ Convective Clusters	Good	Fair	----	----	----
Temp / wind	Excellent	Very good	Skill with max/min		Skill w/ Departures from Normal
QPF/ mean clouds	Very Good	Good	Fair	Some skill with 5-10 day accumulated precip.	

The rate at which forecast features deviate from the actual atmospheric state is not constant, but varies from day to day. In broad terms, the predictability of the flow will hinge on the scale at which key atmospheric process are occurring. For example, at synoptic scales the divergence of forecasts starting from two slightly different initial states will generally be small over the first 1 to 3 days and then grow rapidly beyond that point. There are occasions (e.g., when mesoscale processes such as convection dominate) when rapid divergence of forecasts begins in the first hours of a synoptic scale forecast and conversely times when such divergence may not begin for 5 to 7 days into the forecast. Overall, for any given year and at any location, 4-6 days are expected when a 9-day forecast can be made with a skill that is equitable to a 12 hour forecast and vice versa (Toth, 2001).

As an example, consider Fig. 1, which shows the anomaly correlation between NDFD forecasts and gridded analysis (ADAS) over the western United States during the winter of 2003-4. The separate lines represent forecast projections, warmer (cooler) shaded colors being shorter (longer) forecast projections. Note that in general the warm shaded lines are above the cool shaded lines. This implies the short-range forecasts were more accurate. However, at other times the warm and cool shaded colors are very near each other, suggesting the long range forecast had as much skill as the short-range forecast. At other times the short-range forecasts exhibited anomaly correlation values twice the magnitude as the long-range forecasts.

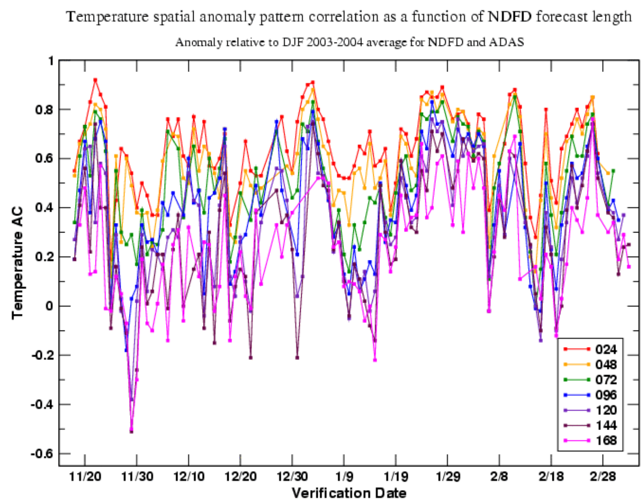

Figure 1

The east coast *Superstorm* of March 1993 and the infamous *Surprise Snowstorm* that struck Washington D.C. in January of 2000 present memorable and extreme examples of the variability of predictability itself. The March 1993 storm was consistently and accurately forecast by the GFS (then AVN) 5 days prior to the storm's occurrence. However, a newer and more sophisticated version of the GFS in January 2000 failed to accurately predict the evolution of the Surprise Snowstorm until after a record snowfall had begun in the Carolinas and just hours before heavy snow began to fall in Washington, D.C. Both the respective success and failure of the GFS were heavily influenced by the inherent atmospheric predictability of the flow regimes of which each storm was embedded.

Figure 2 represents a conceptual model depicting the growth of a distribution of possible synoptic forecasts (e.g., weather maps) given some slight differences (uncertainty) about the initial conditions. The purple shaded areas represent a growing number of possible future weather maps. Note how the area expands in an orderly way through about day 3 (arbitrarily chosen for this example) and then changes chaotically thereafter. This represents the change from slow to rapid forecast divergence and the complimentary increase of possible future weather outcomes. The red line depicts a forecast of one possible atmospheric evolution that an NWP model might forecast. The blue line represents the actual evolution of the atmosphere (e.g., weather map from the true initial state). Note how the NWP forecast follows a similar path to that of the actual atmosphere during the linear (stable) regime through day 3, and then varies widely thereafter in the chaotic (unstable) regime.

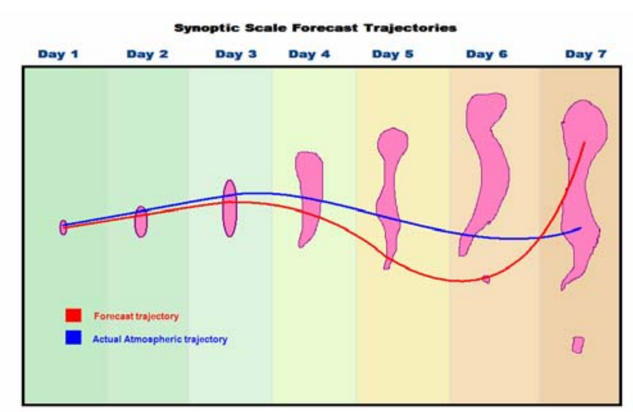

Figure 2: Adapted from Buizza, 2002: http://www.gi.alaska.edu/~bhatt/Teaching/ATM693.Climate.JC/climate.papers/Chaos.pdf

Numerical Weather Prediction: Approaches to Predictability:

Referring again to Fig. 2, consider the single forecast in the context that the actual future weather (blue line) is not yet known. How could one know ahead of time that they should not use the forecast after the third day of a particular forecast cycle, and after the second day of another forecast cycle? Or would some information on the likelihood of alternative outcomes be useful? *An Ensemble Prediction System (EPS)* was developed to address such questions. Currently, NCEP supports two operational EPSs: SREF and the GFS Long-Range Ensemble.

In contrast to a single or deterministic forecast, a well constructed EPS is composed of several, perhaps numerous model integrations (ensemble members), each evolving from slightly different initial conditions, model physics formulations, and/or both. Ideally, each ensemble member has an equally likely chance to represent the actual evolution of the atmosphere. Through such an approach, many forecast trajectories are realized allowing one to estimate the size and complexity of the distribution of possible future weather maps (i.e. the purple areas in figure 2).

Collectively, the many estimates of future weather from an ensemble represent a probability distribution. The complexity of the distribution can then be described statistically. For example, as suggested by Lorenz (1993), the rate at which individual forecasts diverge from each other over time can be used to estimate when the transition from stable to chaotic forecast divergence will occur and hence, provide some knowledge of the point beyond which the forecast should not be used. Alternatively, one would derive a sense of confidence in the forecast by looking for periods when individual forecasts tended to cluster around a common outcome. At other times, there may be more than one cluster of solutions (multimodal probability distribution). Such distributions help one understand possible alternative weather scenarios.

Taking an arithmetic mean of the forecast ensemble represents the application of a dynamic filter. By averaging all the weather maps of a given set of ensemble forecasts, features that are present in all forecasts are coherently represented in the mean while features that are present in only a few of the members are averaged or filtered out. This has important implications considering the predictability characteristics of the atmosphere. The first is that minimally predictable features are filtered out of the EPS. Secondly, as consequence of atmospheric predictability, the ensemble mean will act to maintain forecast detail as a function of the stability of the current atmospheric flow regime. That is the ensemble mean will dynamically provide high detail for clustered or confident forecasts and little detail for highly dispersive forecasts.

In a well constructed EPS the verifying weather map is always represented by one of the ensemble forecast members. However, we already know that model physics are imperfect and so instances arise when the actual solution was not predicted by the EPS. One way to overcome this restriction is to apply statistical techniques to correct (*calibrate)* the EPS, thereby removing biases arising due to model deficiencies. The idea of calibration can be extended, similar to single model NWP, by using regression techniques to derive sensible weather elements via for example MOS techniques for each ensemble member. This approach in sum total would provide a calibrated ensemble MOS.

Current NWP approach to NDFD

Currently, NDFD inputs are primarily based on a collage of single NWP inputs spanning various forecast projections (ie. RUC 0-12 h, Eta 12-72 h, GFS or DGEX 72-168 h). These forecasts are further downscaled to yield the topographic detail previously noted. This deterministic NWP approach to populating the NDFD has several critical deficiencies:

- Current approaches provide as much gridded detail in the day one forecast as the day seven forecast. The forecaster lacks intelligent methods for consistently identifying and filtering out potentially unpredictable features.
- Limited information as to whether the forecast of any given element represents a likely outcome. Users would logically presume that the values given in NDFD (max temperature grid) represent the most likely to occur. Forecasters attempt to compensate by using NWP trends – a type of crude ensemble.
- No confidence information – how good was today's NWP guidance? There are always potentially significant storms on the horizon. Forecasters are limited in their ability to distinguish between *false alarms* and real storms, and hence reliably convey potential impacts to the customer.
- No information is provided on possible alternative outcomes – forecasters are unable to provide specific information as to other possibilities thereby limiting the ability of the user to develop meaningful contingency plans.

Collectively these deficiencies invalidate the current approach to populating NDFD. The process fails the scientific integrity test in that the output does not clearly communicate known predictability limits of weather features over time. More importantly, the process fails to provide the information needed to enable the user to "fit the forecast to their problem" as stated on the NWS NDFD homepage (http://www.nws.noaa.gov/ndfd/).

Discussion: A Valid Approach to NDFD

The state of NWP has advanced to the point where real-time calibrated probabilistic forecasts of sensible weather elements are possible (Mass 2003). Such an approach would address the issues of scientific integrity by:

- Providing multiple inputs to the forecast process. An EPS acts as dynamic filter (Fritsch, 2000) determining the resolvability of scales for any given forecast. In this way, forecast grid detail is proportional to the predictability of the current flow regime;

And the needs of the user by:

- Providing multiple or probabilistic outputs. The ability to express additional dimensions of forecast data such as the most likely outcome, an expression of forecast confidence of significant features, and realistic alternative outcomes to name a few.

Limitation of computing power has previously been a substantial obstacle in the ability to run a high-resolution EPS. Either one could run a single high-resolution deterministic forecast or a relatively low resolution EPS. However, recently strides have been made in the drive towards high-resolution EPS. Such has been the focus of current CSTAR and COMET Coop-funded projects at the University of Washington [Seattle, WA; Mass and Coauthors (2003)] and Stony Brook University (Stony Brook, NY; Colle 2004), respectively.

The advent of such high-resolution ensemble systems such as the SUNY Stony Brook MM5/WRF system provides promise that reliably predictable features like land-sea differences and terrain induced wind regimes will be preserved while filtering out uncertain features as discussed above. For instance, Fig. 3 shows the 48-h forecast ensemble mean 2 m temperature from the SUNY SB 18 member 12 km ensemble, valid 0000 28 January 2005. Note that critical topographic features (such as land-sea interfaces, mountains, and valleys) are evident in the ensemble mean field, while smaller-scale free atmosphere features have been filtered out by the ensemble.

Given the possibilities of a high-resolution EPS, one may ask what role the forecaster should play in the forecast production loop? As Mass (2003) points out, perhaps the greatest failure of the weather forecasting enterprise is the lack of detailed information concerning the next few hours of the forecast we provide. Humans are particularly adept at integrating disparate observations and comparing with short-term model forecasts. Therefore, we agree with Mass (2003) that instead of altering elements in the medium and extended range, forecaster effort should be focused on: providing detailed short-term forecasts; evaluating model predictions and trends; and most importantly, interpreting and explaining the forecast to public and other users.

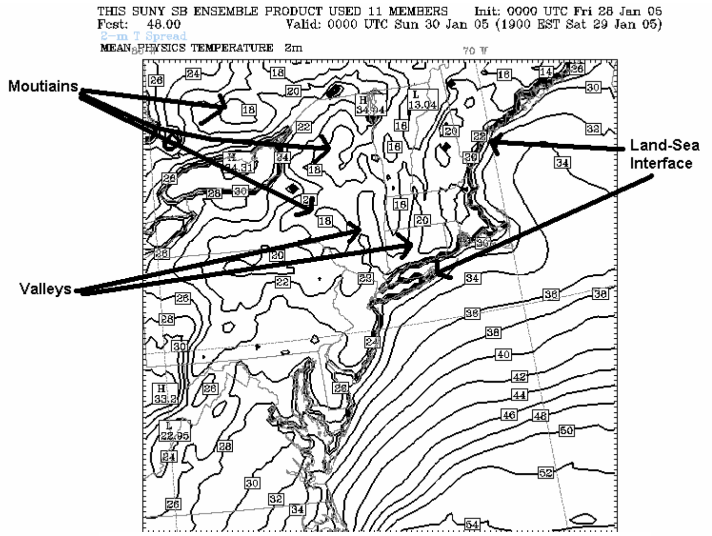

Figure 3: 48-h forecast ensemble mean 2 m temperature from the SUNY SB 18 member 12 km ensemble, valid 0000 28 January 2005.

A Versatile and Useful NDFD

The implicit vision that drives NDFD is the idea that the decision process is more important than the individual decisions or in the NWS case – forecast products. This view recognizes that the value of weather information lies essentially in its ability to support a decision. Results from a number of studies have demonstrated that a probabilistic approach to weather prediction can provide more economic value than a deterministic approach based on a single, deterministic forecast (Mylne 1999; Richardson 2000; and Zhu et al. 2001).

Ensemble prediction systems are particularly useful, if not necessary, to provide reliable early warnings of extreme weather events. The damaging flooding in 1997 on the Red River of the North (Pielke, 1999) was the outcome of an extreme hydrological event and many decisions made over many years. In the aftermath of the disaster, attention quickly focused on the role that official forecast river stages played in flood fighting decision making. The following four conclusions drawn from the Grand Forks experience have broader implications on effective warning strategies:

- The NWS needs to better understand the uncertainty inherent in its own outlooks and forecasts. Information about uncertainty and predictability has potential value to decision makers.
- The NWS needs to explore how to better communicate uncertainty to decision makers. Misuse of predictions can lead to greater costs than if no prediction were provided.
- Responsibility for flood…decision making belongs at the local level. The NWS should not place itself in the position of determining how much risk a community should face. Rather it should primarily be a provider of information needed by decision makers.
- The policy research community needs to focus more attention on understanding the actual use and misuse of predictions. As the forecast community develops a greater range of more sophisticated products, more attention will have to be paid to their

appropriate use. Misuse of predictions can result in large costs and loss of support for NWS activities.

There are many other examples that support conclusion similar to those found above. An evolved NDFD that uses probabilistic inputs and provides probabilistic outputs is an essential step in addressing the most salient findings of both the Red River Flood as well as many other assessments surrounding extreme weather events.

Recommendations

The use of probabilistic weather, water and climate forecasts was formerly adopted as one of the strategic goals to be achieved by 2005 (NWS 1999) in the NWS. This report realizes the necessity and reiterates the importance of such a goal for NDFD. In order to accomplish this goal, the NWS forecasting culture must evolve from a product-centric database populated with discrete forecasts, to a probabilistic system that provides decision support data. An EPS provides a scientifically valid approach to populating and maintaining such a database. Probabilistic forecasts will allow the NWS to convey information in a form that allows customers to manage their risks according to their unique standards.

Table 2 summarizes our response to the stated question. It represents a simple concept-of-operations (CONOPS) for the integration of probabilistic forecast process into the IFPS process. The temporal resolutions are based on natural grouping of various users' needs and the intrinsic predictability at certain scales. The team attempted to provide NWP/EPS inputs that will likely be available in the near future. Lastly, the CONOPS table builds on existing IFPS – NDFD infrastructures.

Forecast Projection	0-6 hour	6-24 hour	Days 1-3	Days 4-7
Forecast Method	Extrapolation of analysis using remote sensing inputs w/ expert system processing; RUC/LAPS tools.	NCEP SREF, local mesoscale EPS.	NCEP SREF, local mesoscale EPS. Calibrated GFS based EPS	Calibrated GFS based EPS
NDFD Time Scale	**Minutes/Hourly**	**Hourly**	**3 Hourly**	**6/12 Hourly**
Smallest features likely depicted	Lake/river breezes, squall-line gust fronts, mesoscale precipitation bands	All terrain forced-features (sea-breezes, local upslope and downslope flows	Frontal precipitation bands / gradients; Major terrain-forced features (ocean-land interfaces, Mountain Ranges)	Major cyclone flows
Forecaster Involvement	**High** Focus on evolution of mesoscale features	**High** Focus on evolution of mesoscale features	**Moderate** Intervention when necessary	**Low** EPS production failure

Table 2: Proposed NDFD Inputs

In terms of spatial predictability (Table 1), there are no limits (other than technological) to the spatial resolution of the grids if an EPS is used as input to the NDFD. The predictability of features generally decreases with the size of the feature, but the rate of decrease of predictability itself varies considerably from forecast to forecast. An EPS acts as a dynamic filter ensuring that the detail (scale of features depicted in NDFD) are consistent with the estimate of predictability of those features for a given forecast.

To that end, this team recommends the following:

1) Apply an EPS to ensure that the spatial details of any given element are proportional to the predictability of that element given the stability of the current flow regime. Additionally, configure the temporal resolution of IFPS – GFE to match the temporal resolution of the NDFD.

2) Accelerate existing operationally produced ensemble data such as the NCEP MREF and SREF into AWIPS following the table 2 CONOPS. This ensemble data should be a complete set of central tendency measures (means), measures of probability (confidence), and alternative scenarios.

3) Develop EPS post-processing methods that will enable the delivery of highly detailed grids of sensible weather elements to the forecaster via AWIPS. The scales of forecasts inputs to NDFD should be comparable to the spatial resolution of data provided as outputs (currently 5 km).

4) Continue to invest in local high-resolution EPS projects and develop procedures to move successful systems (those that produce reliable *and well resolved* -distinguishable from climatology - probabilistic forecasts) to national operational production at NCEP.

5) Task the ISST or other appropriate team with discovering and understanding NDFD enhancements needed to robustly support risk-based management systems. Produce a comprehensive CONOPS detailing a shared vision (with common commitments and clear roles and responsibilities) between all elements of the NWS, including NWS Headquarters, the NWS Regional Headquarters, WFOs, and NCEP (EMC, HPC, SPC, TPC, AWC, etc.).

Final Remarks:

As noted in the introduction to this report – the team engaged in many discussions centering on the context in which to answer the question. The need for probabilistic inputs and outputs in the IFPS forecast process became a reoccurring theme and ultimately framed the discussion above. However, there were three other secondary themes regarding the validity of the current NDFD forecast process:

- Current division of labor and the collaboration process
- A question of the utility of the GFE approach
- A fundamental neglect of the customer in terms of providing and developing risk-management systems – a primary reason for developing an NDFD.

Our motivation below is to simply outline these points.

Given the state of the science and technology, should so many forecasters across 120+ WFOs collectively be involved in the production of NDFD grids? Unlike numerical guidance or EPS data, human collaborated forecasts are inconsistent across space and time. This includes both the production and methods used to arrive at these forecasts which will consistently produce spatially and temporally incoherent forecasts.

A fundamental problem of GFE is the ability of a forecaster to manually adjust one element in such a way as to remain meteorologically consistent with related elements. This has proved to be a daunting task and lead to the development of many GFE “consistency check” tools. Currently, such tools can only address obvious inconsistencies such as a Td > T but do not address more complex relationships such as temp trends with sea breeze fronts and distributions of precipitation given wind and hence quadrant of a given low pressure system. Instead of manipulating outputs of NDFD – the team feels it might be worth attempting to manipulate the inputs of NDFD which is conceivable in an EPS framework.

Lastly, much work has been published on the topic of risk-management of mitigation via various cost-loss assessments. The stated NDFD goal of “tailoring the forecast to the user” speaks to this very application. However, little effort is seen in understanding or building such systems. Also there is a growing frustration in the private sector concerning the roles of public and private sectors in producing and distributing information in graphical format. A renewal or reinvention of NWS partnerships with outside entities is needed. Such an effort would entail an investigation of the interrelationship clarifying the roles and division of labor between public, private, and university sectors in building operational risk assessment systems based substantially on NDFD inputs.

Acknowledgement. The team would like to thank ISST member, Steve Keighton; NWS Eastern Region HQ – SSD Chief, Ken Johnson; and all others who contributed to the development of this report.

References

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

Colle, B. A., M. Jones, J. S. Tongue, and J. B. Olson, 2004: An investigation of the mesoscale predictability over the Northeast U.S. Preprints, *20th Conf. on Weather Analysis and Forecasting,* Seattle, WA, Amer. Meteor. Soc., CD-ROM, 23.2.

Dalcher, A. and E. Kalnay, 1987: Error Growth and predictability in operational ECMWF forecasts. *Tellus*, **39A**, 474-491.

Droegemeier, K.K., 1997: The numerical predictions of thunderstorms: Challenges, potential benefits, and results from real-time operational tests. WMO Bulletin, **46**, 324-336.

Fritsch J. M., J. Hilliker, J. Ross and R. L. Vislocky, 2000: Model Consensus. *Wea. Forecasting*, **15**, 571–582.

Glahn, B., 2003: Comments on “IFPS and the future of the National Weather Service.” *Wea. Forecasting*, **18**, 1299–1304.

Keith, R., 2003: Optimization of Value of Aerodrome Forecasts. *Wea. Forecasting,* **18**, 808-824.

Klazura, G.E. and D.A. Imy, 1993: A description of the initial set of analysis products available from the NEXRAD WSR-88D system. Bull. Amer. Meteor. Soc.**, 74**, 1293-1311.

Levit, N.L, K. Droegemeier, and F.Kong, 2004: High-resolution storm-scale ensemble forecasts of the 28 March 2000 Fort Worth tornadic storms. *Preprints 20th Conference on Weather Analysis and Forecasting*, Seattle, WA.

Lorenz, Edward N, The essence of chaos, University of Washington Press, 1993.

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. Soc., 84,1353-1366.

National Weather Service Improvement Plan, 2005: http://www.nws.noaa.gov/os/nsip/05nsip/Public05.pdf

Nicosia, D.J. and R.H. Grumm, 1999: Mesoscale band formation in three major northeastern United States snowstorms. *Wea. Forecatin*g,**14**, 346-368.

Orlanski, I., 1975: A rational subdivision of scales for atmospheric processes. Bulletin of the American Meteorological Society, **56**(5), 527-530.

Pielke, R. A., 1999: Who decides? Forecast and responsibilities in the 1997 Red River flood. Applied Behavioral Science Review, **7**(2), 83-101.

Ruth, D. P., 2000: Models, forecasters, and Interactive Forecast Preparation in the new millennium. *Preprints 16th International Conference on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology,* Long Beach, Amer. Meteor. Soc., 82-85.

Toth, Z., Y. Zhu and T. Marchok, 2001: The use of ensembles to identify forecasts with small and large uncertainty. *Wea. Forecasting*, **16**, 463–477.

Zhang, F., Snyder, C, and R. Rotunno, 2003: Effects of moist convection on mesoscale predictability. *JAS*, **60**,1173-1184.