Comments on "Implementation and Refinement of Digital Forecasting Databases"

—Bob Glahn
NOAA, National Weather Service,
Meteorological Development Laboratory,
Silver Spring, Maryland

Abrams (2004) presents a very interesting discussion concerning the status and challenges that we all face in implementing and refining the production and dissemination of digital data. The most important nugget from the article was pulled from the conclusions and highlighted in the capsule summary in the *BAMS* article as follows: ". . . many people and processes must work together . . . to create accurate, timely, and consistent forecasts." I could not agree more.

Planning for the National Digital Forecast Database (NDFD) started in early 2000, although its roots in the Interactive Forecast Preparation System, and its predecessors, reach much farther back than that. The Glahn and Ruth (2003) paper gave a snapshot of it in late 2002. After a period of refinement, forecast grids of some weather elements were declared "operational" in late 2004; others will follow in 2005, and still others later. A third (vertical) dimension is being developed for aviation purposes and uncertainty information will be added.

During the "experimental" stage, grids were available for viewing and downloading. Customer response was overwhelmingly positive, even though the schedule for making them operational was not certain. Declaring them operational certified, the National Oceanic and Atmospheric Administration's (NOAA's) National Weather Service (NWS) will provide them as scheduled to the same degree of reliability as it does its other data and products. The NWS's digital services operations concept is contained in Austin (2004). This responds positively to the National Research Council's (NRC's; NRC 2003) following recommendation:

> The NWS should make its data and products available in Internet-accessible digital form. Information held in digital databases should be based on widely recognized standards, formats, and metadata descriptions to insure that data from different observing platforms, databases, and models can be integrated and used by all interested parties in the weather and climate enterprise.

The NWS determines the priority of enhancements to the NDFD based on feedback from customers and partners that was solicited and received in a number of ways. For instance, the makers of wireless handheld/auto/boat devices urgently want NWS watch and warning information on high-resolution grids with rapid refresh.

Abrams discusses many aspects of developing and providing digital data and products. For me to comment on them all would be redundant; I agree with most of what he presents. He asks questions—questions we have asked ourselves before—to which no one knows the complete answers yet. However, there are a few aspects that I will discuss.

HOW THE NDFD CONCEPT DIFFERS FROM OTHER DIGITAL DATABASES.

Abrams, in a footnote, takes issue with the Glahn and Ruth (2003) statement that the NDFD was the "first product of its kind in the United States, and, to our knowledge, in the world." There are two aspects of the NDFD leading to that statement. First, the forecasts are prepared by forecasters at local NWS offices who are familiar with local terrain and other factors that allow them to expertly provide a level of detail that a completely automated system or a centralized system cannot today provide. Second, the NDFD is not just a database, but the word "product" in the statement says it, itself, in its entirety, is a product provided to any and all who have a need and can handle data from the Internet. The use of the word "product" was probably too subtle for those outside the NWS to appreciate, and I apologize for any ambiguity or misunderstanding. To my knowledge, still, no one else has a *national* database, produced by forecasters on site, so to speak, that is made available wholly as data to all interested parties.

Customers are encouraged to use our free downloading and display software. This is not an enhanced product, just data with the facility to use it. Users can download whole grids or smaller sectors. More recently, an Extensible Markup Language (XML; a popular standard for the transmission of small amounts of data) experimental service has been provided for those users needing only a small portion of the NDFD data. The response has, again, been voluminous and very positive; hits as of this writing were averaging 14,000 per day and were increasing at the rate of 25% per month.

PRODUCTION OF A LOCAL GRIDDED DATABASE. All grids in the NDFD are produced locally. This is, at present, time consuming, as Abrams discusses. The software design for the production of grids will undoubtedly undergo evolution. Concepts and prototypes exist for a different method (Ruth 1998, 2000) that has a high potential for making better use of numerical and statistical model output. Abrams' point about careful analysis of data being important is very true; forecasters must have time to do this, with the appropriate tools, and not spend all of their time manipulating a software package. Advanced methods are being extended in the vertical direction for possible use at the Aviation Weather Center, and could likely be used effectively at the Hydrometeorological Prediction Center (HPC). For some uses, a more object-oriented approach will be better. Putting such ideas and prototypes into production takes time and resources, but it is critical to implement what we can when we can. The current production process is where we are today; we will be at another place at another time.

EFFECTIVE USE OF GUIDANCE. Abrams discusses the challenges associated with numerical model outputs and their products—models, ensembles, blends, etc. This is, of course, not unique to the production of a digital database. However, the work that is involved in producing the database is causing forecasters and management alike to assess how the production process might be streamlined. In the past, model output statistics (MOS) have been provided at up to 1500 sites nationwide. By collecting and quality-controlling data from a variety of sources, including co-op stations and mesonets, we are now producing guidance for some weather elements for over 3200 sites over the western mountainous states alone, and 8700 nationwide; starting with the West, these will soon be available in gridded form. Undoubtedly, the additional sites will be useful; however, will the *gridded* MOS be useful, and how can it best be used? That remains to be determined. But, quite likely, more automation will be used for the longer projections by local forecasters through some combination of direct model guidance, MOS, and HPC forecasts.

FREQUENCY OF UPDATING THE NDFD. Updating the NDFD at hourly intervals was an ambitious goal in 2000. This, too, is evolving. Likely, the updating of the forecasts for days 4 through 7 will be less frequent and tied to the four per day model and MOS issuance times. On the other hand, updating in the shorter ranges will be more frequent.

Local databases are updated as the need arises, and software is being developed to tie the warning function to the local digital database. However, the *National* Digital Forecast Database was not designed to be the medium for warning the public of life-threatening events; the grids in the NDFD will need to be updated on a more scheduled basis to make the mosaic palatable. Access to the NDFD through the Internet allows quick probing down to the local database level in situations where this is warranted and a "large scale" picture is not as important as local detail. Putting the NDFD on a 2.5-km grid rather than the current 5-km grid will be a step in its evolution in the not too distant future. In addition, watches and warnings will be added to the NDFD on an unscheduled basis.

PRODUCTION OF WORDED MESSAGES. Abrams is right concerning the challenges of turning digital data into meaningful and pleasing text. I encountered these challenges in developing the concept in the late 1960s (Glahn 1970, 1979); it is really no easier today. It is all in the software and its design. Yes, there are different software languages that may, or may not, be better for that specific task, but the success is in the software design and its controlled flexibility. The NWS still has many customers who do not deal with digital data and require voiced or textual products. This capability must be maintained. Its degree of success will be in refining a design and not redesigning the software every few years, because each implementation must solve the same knotty problems anew.

MATING WITH GIS. As Abrams states, integration with geographic information systems (GISs; which are themselves evolving, is not easy. However, that is the future. Meteorological data are more useful when merged with other data, such as locations of highways, hospitals, storm-surge-prone beaches, and mountain ridges and valleys, than when used without such reference information. We have found that getting an exact match of the same features from different sources is even a challenge. The reference locations of some entities are in error in some accessible databases. The mating with GIS is young ground. But, it is part of our future. We have to embrace it and learn as we go, and not sit back and wait for something perfect to come along—we have to help perfect the system of which we are a part.

QUALITY OF THE PRODUCT. Not mentioned by Abrams is the overall quality and usefulness of the digital database. Not only does this encompass frequently updated information and, to many users, a pleasing picture, but the accuracy and/or skill of the individual forecasts. This is a very knotty issue in itself, because we do not, in general, know the weather at grid points. To assess skill and accuracy, forecasts have to be matched with eventual reality. There are a couple of directions in which this can go with a gridded database.[1] One can either interpolate into the grid to find values at observation locations and compare them in standard ways to define accuracy, skill, and/or usefulness in decision making; or, one can take existing observations and fashion a grid that is consistent with them according to some criteria, and then compare the forecasts at the grid points with the pseudo-observations on the grid.[2]

The first process—interpolating from a regularly spaced grid to random points—is straightforward. Interpolation from random points, if we call it that, to grid points is generally an iterative process, and the quality of the result depends highly on the process itself and the nature and scales of the data involved. But, the question is asked, "How good are the gridded forecasts?" There is work ongoing to determine an "analysis of record," which would be, essentially, the "NWS official" verifying analysis. The success of this endeavor, itself, will be hard to evaluate, because it is trying to produce the unknowable. There may well be as much "error" in the analysis, if we could know what it is, as there is in the short-range forecasts. Dynamic consistency, which plays such a large role in the free atmosphere when initializing for numerical models, is of limited use when analyzing, for instance, maximum temperature at a resolution of 2.5 km in mountainous terrain. For longer-range forecasts, say 5 days, the longwave patterns that are forecast by the models will largely determine the skill, or lack of it, and verification at observation points ought to suffice.

PROBABILITY FORECASTS. I am energetically in favor of probability forecasts, in agreement with the American Meteorological Society's (AMS's) statement in 2002. At the last annual AMS meeting in San Diego, California, the Probability and Statistics Committee, which I chair, held a 1-day course in probability forecasting. This is an area of forecasting and the use of forecasts too long neglected. Probability forecasts can be produced—reliable probability of precipitation (PoP) forecasts have been with us, with skill, since 1966. Dissemination of probabilistic forecasts and education of the user community is probably a bigger challenge than is producing them. As Abrams says, "The issue of deterministic versus probabilistic forecasts is worthy of extensive consideration and debate, but not here."[3]

CONCLUSIONS. In short, Abrams provides a very learned discussion on a topic of our mutual interest. I agree with most of what he says, and it gave me an opportunity to amplify a few things and report on the current status of NDFD. And I will conclude with his statement: "Many people must work together . . . to create sets of forecasts that are accurate, timely, and consistent." One thing this statement implies to me is

[1] It is recognized that a digital database does not necessarily imply a gridded database. Each has its uses.

[2] Mathematicians would call the second process interpolation, while meteorologist would be inclined to call it objective analysis or data assimilation.

[3] I believe "deterministic" is a poor choice of a word meaning; in this context, only "nonprobabilistic". Deterministic implies to me that one can determine the forecasts without error, whereas where there is error, probabilistic is used. "Categorical" can apply to any nonprobabilistic forecast or forecasts without an air of accuracy. There seems to be a misconception that a "categorical forecast" necessarily applies to a range of values; however, "categorical" can be used for its primary definition as "unqualified" or "absolute." If one insists on the "range of values" definition, just use a small range. For most purposes, a temperature forecast of 80°F can mean a range of 79.5°–80.5°F.

that we must eliminate duplication wherever possible and each organization contribute where it best can.

ACKNOWLEDGMENTS. The views expressed are those of the author and do not necessarily reflect those of any element of the U.S. Government.

REFERENCES

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

Abrams, E., 2004: Implementation and refinement of digital forecasting databases. *Bull. Amer. Meteor. Soc.,* **85,** 1667–1672.

Austin, G. S., 2004: NWS digital services operations concept. U.S. Department of Commerce, NOAA National Weather Service, 9 pp.

Glahn, H. R., 1970: Computer-produced worded forecasts. *Bull. Amer. Meteor. Soc.,* **51,** 1126–1131.

——, 1979: Computer worded forecasts. *Bull. Amer. Meteor. Soc.,* **60,** 4–11.

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

NRC, 2003: *Fair Weather—Effective Partnerships in Weather and Climate Services.* National Research Council, The National Academic Press, 220 pp.

Ruth, D. P., 1998: Model interpretation using slider bars in interactive forecast preparation. Preprints, *14th Int. Conf. on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology,* Phoenix, AZ, Amer. Meteor. Soc., 464–468.

——, 2000: Models, forecasters, and interactive forecast preparation in the new millennium. Preprints, *16th Int. Conf. on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology,* Long Beach, CA, Amer. Meteor. Soc., 82–85.