1.2 REAL TIME FORECAST VERIFICATION TOOLS AT THE WEATHER CHANNEL, ATLANTA GA

Jeral G. Estupiñán *[1,2], Joseph Koval [1], Royston Lucksinger [1], Christopher Peters[1], M. Patrick Rouse [1], M. Long[1], and B. Rose[1]
[1] The Weather Channel, Atlanta, GA; [2] School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, GA

1. INTRODUCTION

Synoptic-scale forecasts have increased in skill over the last 30 years (Landis 1994; Kalnay et al. 1990; Shuman 1989). Improvements in numerical and statistical guidance have led to strong upward accuracy trends in all types of precipitation forecasts Charba et al. (1980). Charba et al. (1980) found that the largest upward trends in skill in probability of precipitation (PoP) forecasts occurred for the periods 24-36 h and 36-48 h after 0000 and 1200 UTC since the late 1960's. They showed that the skill of these forecasts improved at an average rate of 7.2% during each 10-year interval. Over the same period there has been improvement in PoP forecasts of 2% for the period 12-24 h per decade as well as in the accuracy of "yes/no" forecasts of measurable precipitation. Improvements began in the early 1970's and became more pronounced in the 8-year period prior to 1980.

The Hydrometeorological Prediction Center (HPC) has provided verification of quantitative precipitation forecasts (QPF) since 1960 (Olsen et al. 1995). For the first 24 h period, threat scores have increased from 0.16 in 1965 to 0.24 in 1998 for the 1.0 inch bin (Fristch et al., 1998). Threat score values for different bins are now available via the HPC verification website (NVP). For the 1.0 inch bin the values have improved to 0.26 in 2004 (McDonald et al., 2000). The accuracy of temperature forecasts has also increased since the early 1970's. For minimum temperatures the mean absolute error (MAE) has decreased from about 5.0° to 3.5°F for day 3 forecasts from 1972 to 2004. For the same period the day 3 MAE for the maximum temperatures has decreased from about 6° to 4°F.

The improvement in skill and accuracy of the results described above are mainly due to vast improvements in model physics and resolution coupled with statistical analysis of the model via more numerous observations. Doswell (2004) describes how the gap between human produced weather forecasts (not warnings), and guidance has been steadily shrinking. Baars et al. (2005) showed that the average model output statistics (MOS) from two or more models is competitive or superior to human forecasts at nearly all locations. Further, they found that weighting component MOS predictions by their past performance is superior to the average of MOS results from two or more models. Human forecasters were found to be skillful compared to MOS during the first forecast day and for periods when temperatures differ greatly from climatology.

Mass (2003) indicates that forecasters can make a real contribution modifying forecasts for phenomena that are poorly simulated by the models. The challenge consists of maximizing resources to focus on these events. Nelson et al. (1999) indicate that the demands of the modernization of the NWS have brought improvements and challenges. One of the challenges of the forecaster has been to assimilate a large amount of numerical model data in limited time. In order to assist with these tasks, the NWS Weather Forecast Office (WFO) in Tulsa Oklahoma implemented "Coach" in July of 1998 to enhance its operational forecasting program (Nelson et. al, 1999). As explained by Nelson et al. (1999), Coach was created to enhance current weather pattern learning. Coach stores meteorological conditions for different dates for future availability to forecasters. Then these data are made available in three different ways: (a) through a list of hyperlinks with information about the most recent cases; (b) allowing for the retrieval of a single case for any date; and (c) keying off meteorological data and forecaster entries to recall similar events stored in the database. Coach is designed to be used at a specific WFO. The period of interest is the first 24 hours of the forecast. As described by Nelson et al. (1999) the power of Coach lies in its ability to collect, organize, and produce a wide variety of relevant

* Corresponding author address: Jeral G. Estupiñán, 300 Interstate N. Parkway, The Weather Channel, Atlanta, GA 30339 (jestupinan@weather.com) – 770-226-2390.

data for a forecaster at the very time the support is needed. As described by Nelson et al. (1999) performance support systems (PPS) in general have been developed in the military and medical community to enable a high level of job performance with a minimum of support from other people.

Inspired by the Coach system, Ryan et al. (2000) created Mentor, another performance support system, to assist the forecast operations of the Australian Bureau of Meteorology. As described by Ryan et al. (2001) Mentor's goal is to support forecasters' decision making by recording their opinions and concerns about a large number of weather situations, and by providing easy access to statistical analyses and individual case data.

2. BENEFITS AND ANTICIPATED RESULTS OF ADVISOR

We will present the results of the operational use of a real time verification system at The Weather Channel (TWC). A real time verification system is needed at TWC in order to focus our attention to forecasting tasks that are handled relatively poorly by computers at a time when numerical guidance drives almost every aspect of the job of the forecaster. The verification system is referred as ADVISOR (Advanced Verification Information for Smart Operational Results). Advanced verification information refers to near real time verification information that can provide quick results to operational weather forecasters. This is different from the more conventional verification statistics (threat scores, bias, MAE, etc) that are normally produced to support longer term analyses for bookkeeping purposes. Smart operational results refer to the intelligent decisions that forecasters will be able to perform (e.g. focus efforts to situations where humans can add most value) after examining the verification results from ADVISOR. The system is designed to provide support for the entire continental United States since TWC is responsible for providing weather forecasts for the entire country. ADVISOR possesses characteristics of both a performance support system and a real time verification system.

Before the era of automation, Sverre Petterssen explained that it is important that the forecaster be able to identify the "abnormal" situations when the models are likely to be inadequate (Doswell, 2004). The unexplained residual from the machine-made forecasts invites study as indicated by Petterssen. ADVISOR allows us to investigate "this residual" with the aid of automated analyses in order to minimize the time spent by forecasters trying to decide between disagreeing solutions from the guidance. ADVISOR is also designed to investigate the temporal aspects of forecast accuracy (human added value as a function of forecast day). ADVISOR will provide feedback about the tasks that are done well and what aspects of the forecasts need some effort. As described earlier, it is often a challenge for human forecasters to improve upon guidance in many cases. ADVISOR provides meteorological and technological solutions which provide the flexibility to add additional verification and forecast tools. These tools are needed to tune human intervention to the continuous improvement in meteorological guidance.

It is expected that ADVISOR will produce useful feedback in these areas:

- The ability to compare present weather situations with past events in context of forecasters' discussions (subjective judging) and verification results (objective judging) for the whole country. The analysis of this data should help to focus the forecaster efforts to situations where humans can add most value.
- ADVISOR will be a learning tool for the forecasters containing a database of verification statistics for the entire country. The classification of the data by importance of meteorological events (extreme events vs. fair weather) will allow forecasters to learn in context of weather conditions.
- The capability to determine after which day of the forecast it becomes difficult for human forecasters to add value in reference to the guidance used at TWC. This will allow us to focus our efforts to short term forecasts and extreme events.
- Motivate an open dialog between forecasters, managers, scientists, and engineers of the end-to-end forecast system with the goal of maximizing efficiency at TWC.
- The ability to identify needed engineering improvements in the end-to-end forecast system.

Results in these areas will be presented. Detailed analyses of individual weather situations will be presented to show that the results of ADVISOR focus forecasters to make better decisions in an operational setting. The large amount of information that forecasters are faced with every day, conflicting sources of guidance, or simply routine forecast decisions which are buried within the vast amount of data, make it difficult to be fully efficient when decisions need to made under pressure. There will be missed opportunity if we do not tune and optimize the job of the forecaster to attain maximum efficiency. Some of this optimization can be achieved without improvements in the numerical guidance but instead by redirecting the efforts of the human forecaster. ADVISOR will facilitate the complex decisions that today's forecasters are faced with.

With a better understanding of these situations, more time can be spent on nowcasting and high impact weather events (where models are currently less accurate). ADVISOR will provide us with the objective data to be able to achieve these results.

3. GENERAL DESCRIPTION OF ADVISOR

ADVISOR is a web-based system that runs on a Linux operating system. The system displays and archives temperature and precipitation verification results in the form of maps for daytime and nighttime forecasts. All verification results are calculated for the unedited forecasts (before human intervention) and the edited forecasts (human intervened forecasts). The system logs automatically all verification statistics to data files for future analysis.

The temperature maps display mean absolute errors (MAE) (Figure 1-2) and information about the magnitude and success/failure of each temperature edit (Figure 3). If the temperature edit was performed in the right direction it is displayed as a green number indicating the amount by which the temperature was moved in the right direction.

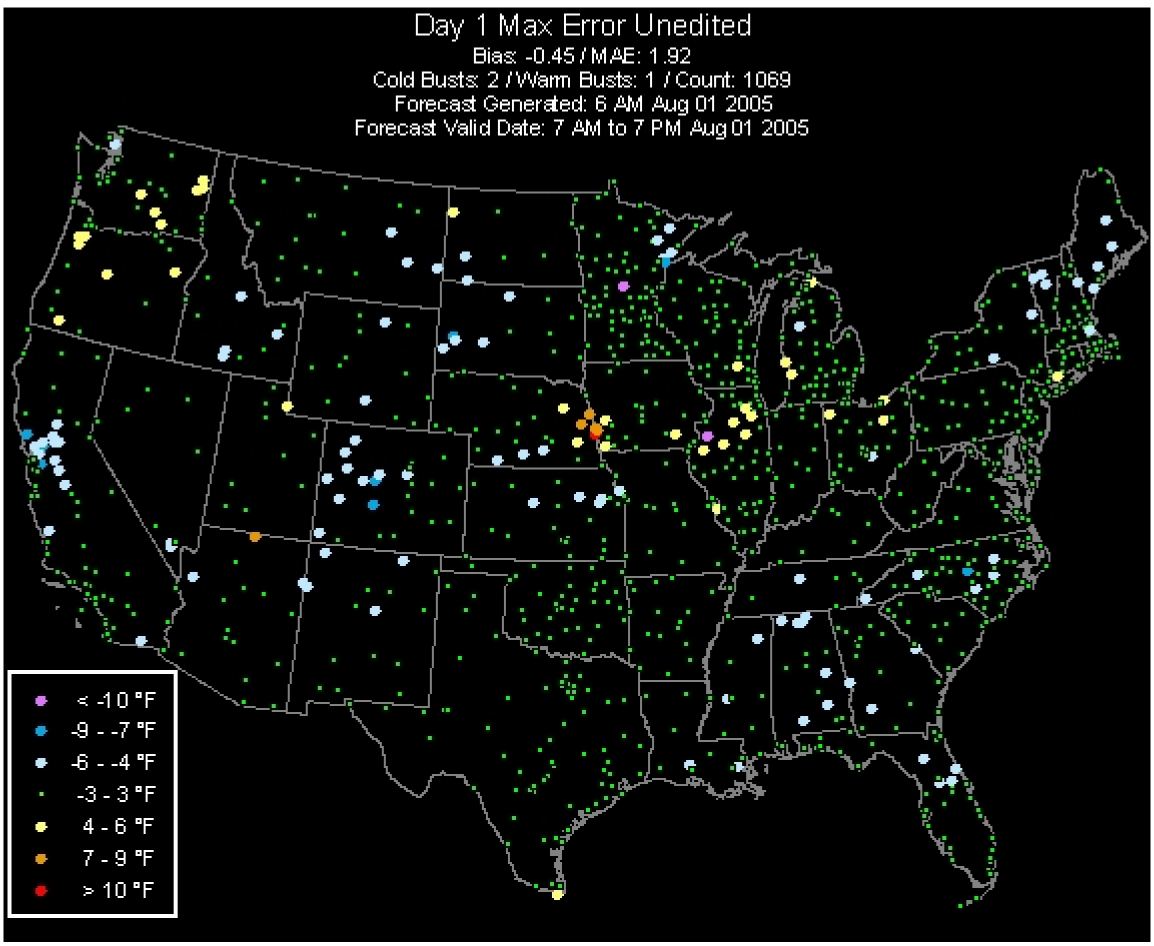

Figure 1. Example map showing the maximum temperature error for the raw guidance (unedited data). Red dots indicate places where the temperature forecasts were overestimated by more than 10F.

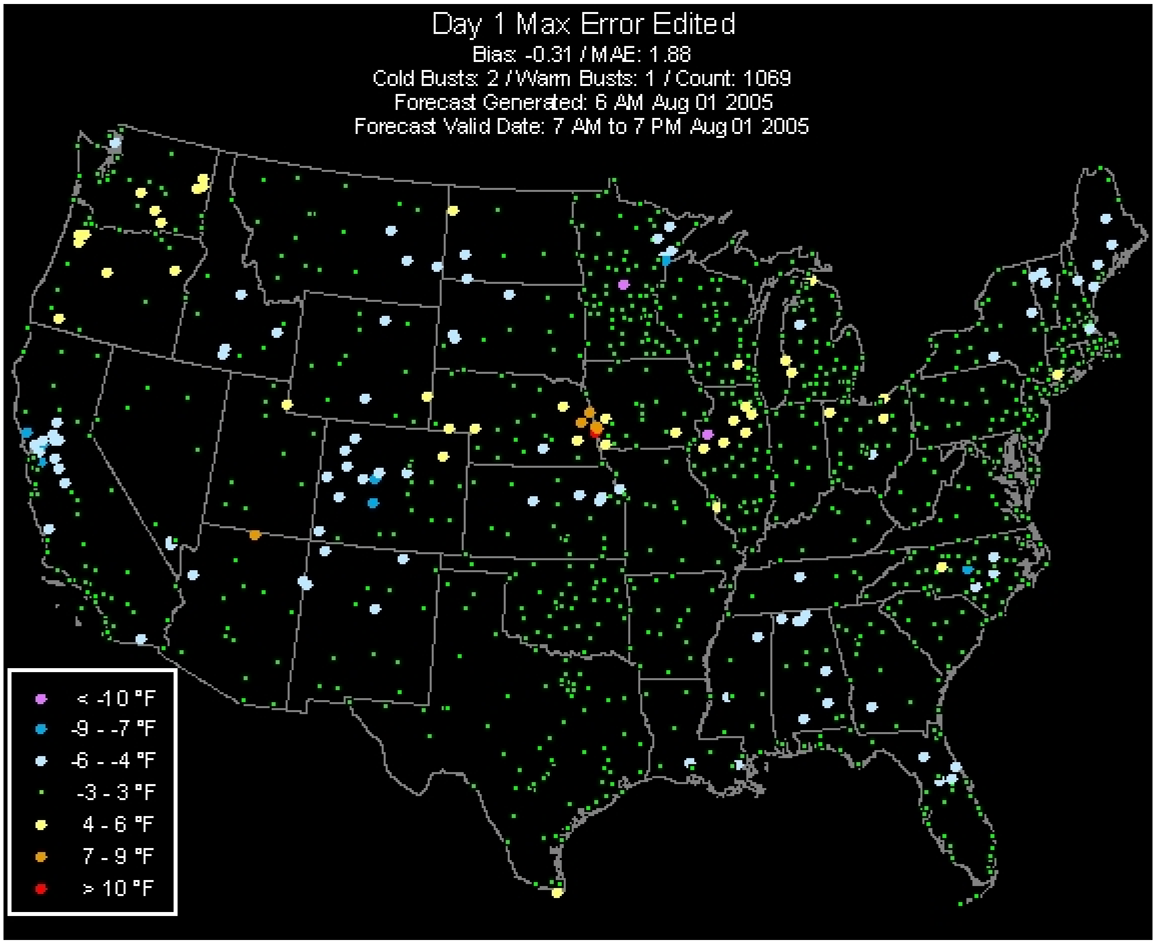

Figure 2. Example map showing the maximum temperature error for the human-intervened data. Red dots indicate places where the temperature forecasts were overestimated by more than 10F.

If the temperature edit was performed in the wrong direction it is displayed as a red number indicating the amount by which the temperature was moved in the wrong direction (Figure 3). Summary verification statistics are displayed at the top of each map (MAE, bias, total number of points displayed, number of cold/warm busts, and number of good/bad edits). These daily data are logged to archive files for future analysis.

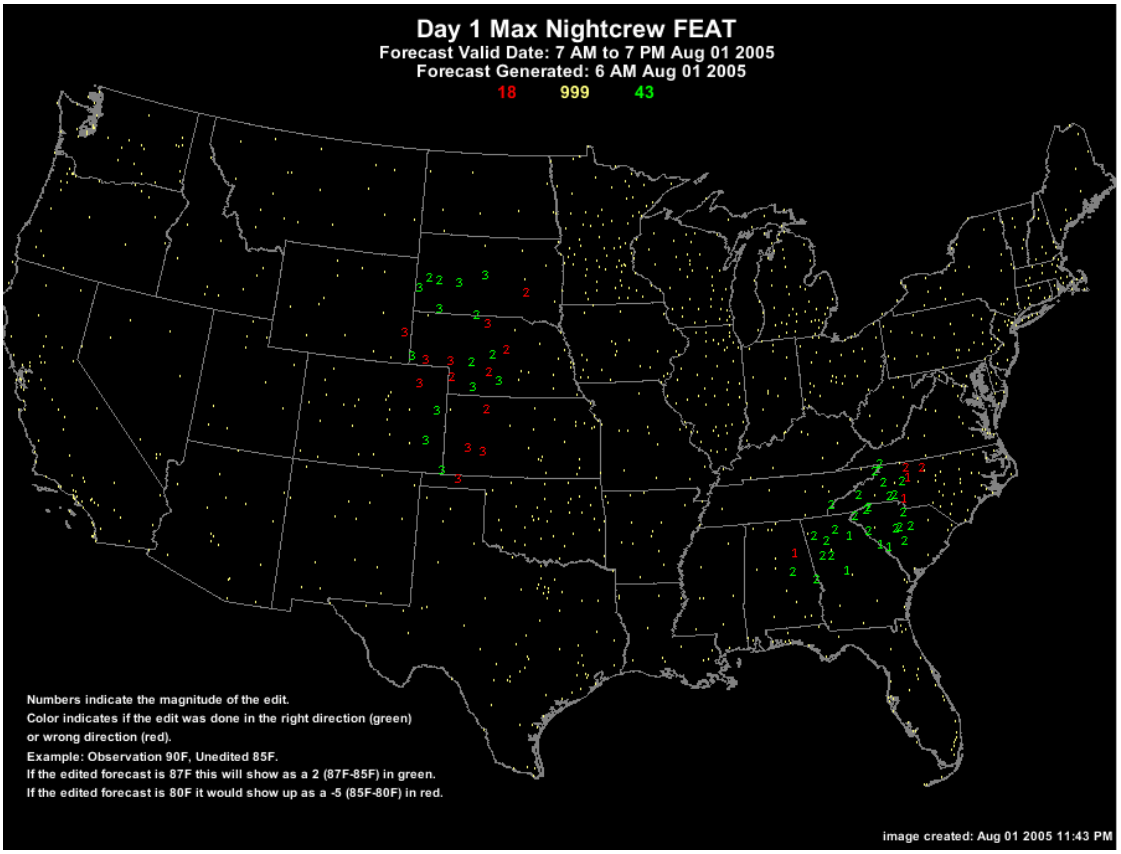

Figure 3. Map showing magnitude and direction (good/bad) of the temperature edit (see map legend)

The precipitation verification maps compare forecast values of quantitative precipitation forecast (QPF) with actual precipitation amounts measured at ground stations. The observations are displayed as colored dots over a shield of forecast QPF amounts. All precipitation statistics are logged to daily files for future analysis. As in the case of temperature, all precipitation statistics are calculated for the edited and unedited data. The 06z radar/satellite image and the 12 hour accumulated precipitation prior to 00z are also displayed and archived next to the precipitation verification maps.

ADVISOR allows the forecasters to enter forecast discussions using a web interface. These discussions are then classified, displayed, and archived next to the verification for each of the forecast days.

The system also displays and archives surface and 500 hPa maps of the GFS, NAM, and ECMWF models which were used by forecasters as a guide at the time the forecasts were generated. These maps are displayed next to the verification maps and forecast discussions. The actual model initialization map for the time of the verification is also displayed. This allows the forecasters to compare the map of the model they used when preparing the forecasts with the actual conditions that verified.

4. REFERENCES

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

Charba, J. P., W. H. Klein, 1980: Skill in precipitation forecasting in the National

Weather Service. *Bull. Amer. Meteor. Soc.,* **61**, 1546-1555.

Doswell, C, 2003: Is there a role for humans in the NWS of the future? (Is there an NWS in our future at all?), http://webserv.chatsystems.com/~doswell/forecasting/human_role/future_forecasters.html

Fritsch, J. M., R. A. Houze Jr., R. Adler, H. Bluestein, L. Bosart, J. Brown, F. Carr, C. Davis, R. H. Johnson, N. Junker, Y.-H. Kuo, S. Rutledge, J. Smith, Z. Toth, J. W. Wilson, E. Zipser, and D. Zrnic, 1998: Quantitative precipitation forecasting: Report of the eighth prospectus development team, U.S. Weather Research Program. *Bull. Amer. Meteor. Soc.,* **79**, 285-299.

Kalnay, E., M. Kanamitsu, and W.E. Baker, 1990: Global numerical weather prediction at the National Meteorological Center. *Bull. Amer. Meteor. Soc.* **71**, 1410-1428.

Landis, R., 1994: Comments on "Forecasting in Meteorology." *Bull. Amer. Meteor. Soc*., **75**, 823-827.

Mass, C. F., 2003: Forecaster's forum IFPS and the Future of the National Weather Service, *Wea. Forecasting,* **18**, 75-79.

McDonald, B. E., T. M. Graziano, and C. K. Kluepfel, 2000: The NWS National QPF Verification Program. Preprints, 15th Conference on Hydrology, Long Beach, CA, January 9-14, American Meteorological Society, p. 247-250.

National Weather Service HPC National Verification Program (NVP), 2003: http://www.hpc.ncep.noaa.gov/html/hpcverif.shtml, 2003.

Nelson, Steven E., and L. P. Rothfusz, C. Jacobson, 1999: Coach – Design and Implementation of an operational support system at the Tulsa WFO. Proc. 15th Conf. On IIPS for Met., Ocean., and Hyd.,(Dallas, TX), AMS, volume.

Olsen, D. A., N. W. Junker, and B. Korty, 1995: Evaluation of 33 years of quantitative precipitation forecasting at the NMC. *Wea. Forecasting,* **10**, 498-511.

Ryan, J. Christopher, Ankur Jha, and Swarup Joshi, 2003: Mentor – A performance support system for forecasters, 19th Conference on IIPS, American Meteorological Society, Long Beach, CA., USA.

Shumman, F. G., 1989: History of numerical weather prediction at the National Meteorological Center. *Wea. Forecasting*, **4**, 286-296.