3.12 THE NWS NATIONAL QPF VERIFICATION PROGRAM

Brett E. McDonald # *
NOAA/NWS/OM & UCAR/COMET
Thomas M. Graziano & Charles K. Kluepfel
NOAA/NCEP/OM

1. INTRODUCTION

The United States Weather Research Program (USWRP) and the National Weather Service (NWS) have recently identified quantitative precipitation estimates (QPEs) and forecasts (QPFs) as a priority for improvement in the research and operational communities (Fritsch et al. 1998; Office of Meteorology 1999). Objective assessment and quantification of the skill of QPFs in the NWS end-to-end (ETE) forecast process are necessary to: (1) identify the value added at each step of the ETE forecast process; (2) assist in improving the forecasts; and (3) insure that the ETE forecast process represents the most efficient use of resources to produce quality QPF information for hydrologic services. Currently, no such program exists within the NWS. Thus, the NWS recently outlined a uniform national QPF verification program and plans to establish the National Precipitation Verification Unit (NPVU) to fulfill these requirements. Verification statistics from the NPVU will serve to support NWS programmatic decisions and numerical weather predication (NWP) model changes, provide feedback to individual forecasters and forecast offices, and ultimately improve QPFs and associated products for outside users. The success of the program is dependent upon the timely availability of all QPEs and QPFs.

The purpose of this paper is to present the basic components of the national QPF verification program as described in Office of Meteorology (1999) with adjustments according to recommendations from the NWS QPF Process Assessment Team (NWS 1999). The NPVU will be established at the NOAA Science Center in Camp Springs, MD. Preliminary efforts within the prototype NPVU have been in assisting the Hydrometeorlogical Prediction Center (HPC) in updating their QPF verification activities. QPF verification methods and results for HPC's 06-h, 24-h, and 5-day QPFs will be described to illustrate the current quasi-real-time HPC QPF verification process. Research on verification methodologies are and will continue to be important in establishing and improving the national QPF verification program. The NPVU participated with the QPF Process Assessment Team, and some methods and results from this work will be discussed. Finally, future plans for the national QPF verification program will be presented.

This research was supported in part by an appointment to the COMET Postdoctoral Fellowship Program and UCAR Visiting Scientist Program sponsored by the National Weather Service and administered by the University Corporation for Atmospheric Research pursuant to the National Oceanic and Atmospheric Administration award No. NA97WD0082.

* *Corresponding author address:* Dr. Brett E. McDonald, NOAA/NWS/NCEP/HPC, 5200 Auth Road Room 410, Camp Springs, Maryland 20746; E-mail: brett.mcdonald@noaa.gov.

2. CONCEPTUAL OUTLINE

The premise of a QPF verification program is dependent on the availability of both observed and forecast precipitation data. If the data are in the same format, comparisons can be made with informative performance measures. If not, then care must be taken to manipulate the data so that the observed and forecast products have similar formats. Even still, some products are incompatible for verification. For example, much debate exists over the treatment and comparison of gauge precipitation observations with areally-averaged gridded model output (Gaudet and Cotton 1998).

Ideally, QPF verification would be performed with a perfect gridded multi-sensor analysis which incorporates gauge, radar, and satellite data. Quality control of the observed data is necessary since limitations still exist in the observation and transmission of precipitation data. However, the best "ground truth" currently available for most of the country is the RFC Stage III analysis (Breidenbach et al. 1998). Precipitation in the western U.S. is best represented by point gauge data or grids rendered via Mountain Mapper (MM, Henkel and Peterson 1996). Thus, the NPVU will utilize the aforementioned analyses as the observed data while constantly assessing, along with other appropriate NWS personnel, the quality and coverage of the observed data. Most observed precipitation data utilized by the NPVU will be gathered and sent by the NWS River Forecast Centers (RFCs).

Initially, the proposal was made to verify all precipitation forecasts from the Environmental Modeling Center (EMC), Techniques Development Laboratory (TDL), HPC, Weather Forecast Offices (WFOs), and RFCs. The NWS QPF Process Assessment Team recently recommended that QPFs from the WFOs east of the continental divide do not need to be generated for input into the NWS River Forecast System (NWSRFS). However, all WFOs may, if they have a local requirement, produce QPFs for internal and external local use. Although QPFs may vary in type, format, resolution, etc., appropriate methods will be taken to translate the QPFs to common formats for fair and accurate comparisons.

The use of climatological precipitation data in QPF verification is necessary to provide a baseline of skill with which to compare the QPFs. Appropriate precipitation climatologies, such as those developed via PRISM (Daly et al. 1994) and TDL (Charba et al. 1998), will be utilized.

Verification statistics will be computed from NWP and forecaster-generated QPFs for all possible combinations of the following *as appropriate*: (1) forecast increments of 3-, 6-, 12-, 24-h, etc. for all possible forecast projections; (2) spatial domains - nation, NWS region, geographical and climatological regions, RFC domain, WFO area, river forecast group, and MAP area; (3) temporal domains ranging from individual forecast periods and model runs to a day, week, month, season, year, etc.; and (4) spatial resolutions beginning with ~32km and including multiples thereof (4, 8, 16, 64, 128, 256 km). Measures of performance will include: threat and equitable threat score; bias score; errors (mean, mean absolute, root-mean-squared); bias; Bayesian informativeness score; correlation coefficient; Nash-Sutcliffe sufficiency score; Brier score; ranked probability score; etc. These statistics can be derived for specified precipitation thresholds and discrete intervals or for the full range depending upon the verification measure being computed.

Integral to the QPF verification program is the display and feedback method employed to relate verification information to the forecasters, model developers, researchers, and management. Statistics will be computed for offices as a whole as well as for individual forecasters, where privacy will be maintained. Although details have not been finalized, utilization of AWIPS and the World Wide Web is planned.

Software developed for the QPF verification system will initially be located and run at the NPVU. The NPVU will perform QPF verification for the entire nation as well as for smaller domains. Archival of all data pertaining to QPF verification at the NPVU is essential to the program.

Further details concerning the national QPF verification program are given in Section 11.3 of Office of Meteorology (1999).

3. HPC VERIFICATION

In January 1999, the HPC transferred QPF verification from the Intergraph system to a Unix-based HP workstation at the prototype NPVU. Several errors in the prior QPF verification system were corrected, but the basic characteristics were continued so that results can be compared with the 30+ years of QPF verification data (Olsen et al. 1995). Additionally, the 5-day QPF verification system has been established whereas prior efforts were performed manually.

3.1 *06-h QPF Verification*

At the present time, 06-h QPF verification is performed at 600+ METAR locations throughout the conterminous U.S. Each of these stations has been evaluated for reliability and consistency. However, HPC forecasters still quality control the precipitation reports before verification is performed. Both EMC model and HPC forecasts are bilinearly interpolated to the station locations. Because HPC forecasts are in a threshold format starting at 0.25", only threshold statistics (threat score, bias score) are computed.

Beginning shortly, HPC will issue fully continuous QPF grids. More informative and hydrologically meaningful verification measures will then be computed and intercomparisons with other QPFs will be made. Also, over the course of the following year, the aforementioned RFC Stage III (soon to be RFC-wide) and MM analyses will be utilized as observational grids such that the 06-h QPF verification system methodology will change from a point to a grid structure.

Monthly 0.25" threat scores from Jan.-Aug. 1999 are shown in Figure 1. The HPC 91E QPF (F00-F06) is compared with the Eta, AVN, and RUC2 F06-F12 QPFs. Threat scores indicate that HPC forecasters are adding value to the NWP guidance for this forecast period. The QPF performance is much better during the cool season than during the convective warm season.

3.2 *24-h QPF Verification*

The HPC has been issuing 24-h QPFs since September 1960 and has always utilized a verification system to measure progress and monitor forecast quality (Olsen et al. 1995). An areal-verification scheme has been consistently maintained resulting in nearly 40 years of threat and bias scores. Much of the effort has involved manual intervention; however, more and more individual components are becoming automated to maximize resources without degrading the quality of the QPF verification.

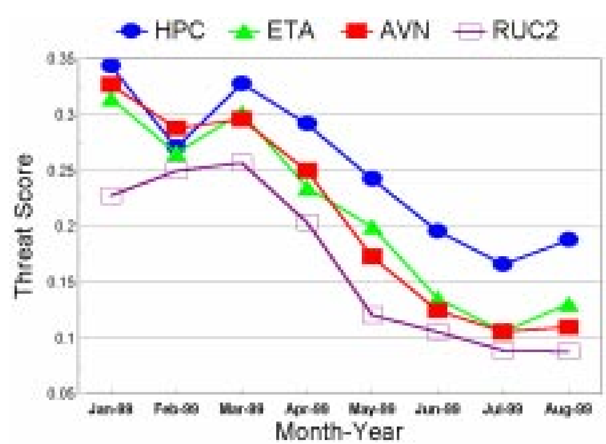

Figure 1. National 0.25" threat scores at 600+ sites for 06-h QPFs from Jan. to Aug. 1999. HPC QPFs are for F00-F06. Eta, AVN, and RUC2 QPFs are for F06-F12.

At present, the 24-h QPF verification system is characterized by manual gridded analyses of 24-h gauge data from the RFCs, translation of the HPC graphical product to a grid, and remapping of NWP model output to a common grid. The QPF verification grid used prior to 1999 has a spacing of 1/6 of the LFM grid (~30 km). Since January 1999, the AWIPS #221 grid (~32 km) has been used. Only those grid points over the conterminous U.S. are verified, and threat and bias scores for the day 1, update, and day 2 QPFs are produced.

Figure 1 in Hoke et al. (2000) shows the yearly 1" threat scores from 1965 to 1998 for each HPC 24-h QPF product. Trends indicate that improvements in 24-h QPF have been made. As expected, the day 1 QPF is better than the update and day 2 QPF.

Figure 2 shows the 0.50" 24-h threat scores for Jan. to Aug. 1999. HPC is consistently better than the best NWP QPF with extremely high scores during Jan. 1999. The AVN proved to be the better NWP QPF during winter while the Eta is better during the summer. The NGM, frozen since 1991, has the lowest threat scores, but still shows modest skill in QPF.

3.3 *5-day QPF Verification*

Part of the medium range (3-7 day) forecast product suite at the HPC includes a day 1 to day 5 cumulative QPF for the conterminous U.S. This QPF is also produced as a graphic and translated to the verification grid. Its valid time corresponds with the F24-F144 5-day QPF from the 0000 UTC run of the Medium Range Forecast (MRF) model. The best observed data available for this forecast period however incorporates 1200 UTC to 1200 UTC 24-h accumulations. Thus, the verification and forecast data are time lagged by 12 h with the observed data ending time occurring prior to the forecast data ending time. This time lag is not deemed critical given the length of the period, although there may be some unknown implications/errors. Future HPC plans will allow the QPF product to better match the accumulation period of the observed data. The observed data is grid averaged to the verification grid without HPC quality control but with some CPC automated quality control.

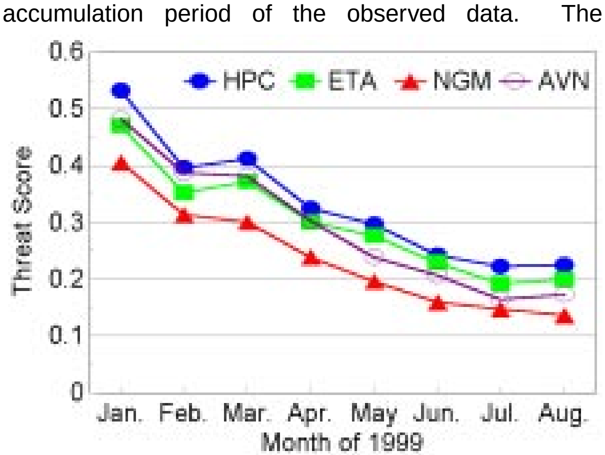

Figure 2. National 0.50" threat scores for Jan.-Aug. 1999 via grid methodology. HPC QPF is from F00-F24 and NWP QPFs are from F12-F36.

4. NWS QPF PROCESS ASSESSMENT TEAM METHODS AND RESULTS

Details concerning the charge given to and the efforts of the QPF Process Assessment Team are found in NWS (1999). This section will describe the objective comparative QPF verification study that was key to the findings and recommendations of the team. The study included QPF products from EMC, TDL, HPC, WFOs, and RFCs for a 6-month cool-season period (Oct. 1998 to Mar. 1999). QPF verification was conducted for three geographically and climatologically diverse RFC areas: ABRFC; OHRFC; and CNRFC. The study was limited due to the nature of data archival at each of the various centers and offices and the time allowed for the team to complete its assessment.

NWP 12-36-h QPFs were obtained for the 0000 UTC models runs of the Nested Grid Model (80 km), Eta model (32 km), and the Aviation (AVN) run of the MRF (1° x 1°). Model grids were remapped to a 30-km grid using an area-preservation technique (Mesinger 1996). The TDL QPF product comes from the Local AWIPS MOS Program (LAMP) QPF model (20 km). LAMP guidance comes in the form of probabilities, best category, and expected value. Four 06-h QPFs were merged together from three different LAMP runs to create a 24-h QPF. HPC 24-h and 06-h graphic QPFs were translated to the verification grid in a semi-continuous manner since the HPC does not include a zeroline. WFO QPFs on the HRAP grid (4 km) were mosaicked together over each RFC domain and grid-averaged to the verification grid. RFC QPFs were also grid-averaged from the HRAP grid to the verification grid.

For the ABRFC and OHRFC, observed data were obtained from the RFC Stage III analyses in NetCDF or xmrg format (both on the HRAP grid). The observed data was also grid-averaged to the 30-km verification grid so that all of the observed and forecast data were on the same resolution grid. Observed data for the CNRFC were obtained from Mountain Mapper, which is a program that renders observed point data to the HRAP grid via climatology (PRISM).

A modest suite of 6- and 24-h verification measures were computed for a variety of temporal domains (1, 3, & 6 mo and single days). Forecast projections were limited to the 1200 to 1200 UTC period, which corresponded with HPC's day 1 QPF and the WFO and RFC 1200 UTC QPFs, and included 06-h forecast increments at standard synoptic times.

Evaluation of the results by the QPF Process Assessment Team indicated that, for both the ABRFC and the OHRFC over the 6-mo period, 24-h mean absolute errors were better for the HPC than for the WFOs and the RFCs (Fig. 3). Results also indicated that

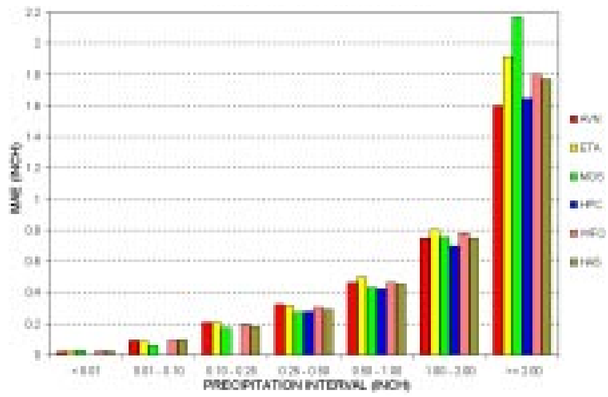

Figure 3. Mean absolute errors for 24-h QPFs over both the ABRFC and OHRFC for the 6-mo period Oct. 1998 to Mar. 1999.

the AVN performed the best of the NWP and statistical model QPFs for this cool season.

Results of the QPF process assessment over the CNRFC were found to be inconclusive. The WFO scores were better than all other QPFs for higher amounts, but several factors may have biased the results. These factors include common grid rendering of the observed and forecast point data via Mountain Mapper and inclusion of updated forecasts during wet events up to 18 h into the forecast period. Thus, the verification study will be continued during the 1999-2000 wet season to better ascertain the relative performance of QPF process components in the western U.S.

5. FUTURE PLANS

The NPVU plans to become fully operational in fall of 2000 to include verification of QPFs from the EMC, TDL, HPC, and RFCs over the conterminous U.S. Initial development will concentrate with selected RFCs. Most likely, this will include a display and feedback method using the World Wide Web.

Eventually, if possible, the software for the national QPF verification program will be incorporated into AWIPS for greater accessibility in the NWS. Since the NWS is tending toward probabilistic QPF in the next few years (Office of Meteorology 1999), the verification program will be modified to reflect these changes.

The conference presentation accompanying this paper will be available at the following URL address: http://www.hpc.ncep.noaa.gov/npvu/hydro15.

6. REFERENCES

Breidenbach, J. P., D.-J. Seo, and R. Fulton, 1998: Stage II and III post processing of NEXRAD precipitation estimates in the modernized Weather Service. Preprints, *14th International Conference on Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology*, Phoenix, AZ, January 11-16, American Meteorological Society, 263-266.

Charba, J. P., Y. Liu, M. H. Hollar, B. E. Exley, and A. Belayachi, 1998: Gridded monthly climatic frequencies of precipitation for 1-, 3-, and 6-h periods over the conterminous United States. *Wea. Forecasting*, **13**, 25-57.

Daly, C., P Neilson, and D. L. Phillips, 1994: A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *J. Appl. Meteor.*, **33**, 140-158.

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.

Gaudet, B., and W. R. Cotton, 1998: Statistical characteristics of a real-time precipitation forecasting model. *Wea. Forecasting*, **13**, 966-982.

Henkel, A. and C. Peterson, 1996: Can deterministic quantitative precipitation forecasts in mountainous regions be specified in a rapid, climatologically consistent, and RFC-compatible manner with Mountain Mapper functioning as the tool for mechanical specification, quality control, and verification? Abstracts, *Fifth National Heavy Precipitation Workshop*, State College, PA, September 9-13, p. 31.

Hoke, J. E., D. W. Reynolds, E. J. Danaher, and K. C. McCarthy, 2000: The Hydrometeorological Prediction Center -- Its future role in quantitative precipitation forecasting. Preprints, *15th Conference on Hydrology*, Long Beach, CA, January 9-14, American Meteorological Society.

Mesinger, F., 1996: Improvements in quantitative precipitation forecasts with the Eta regional model at the National Centers for Environmental Prediction: The 48-km upgrade. *Bull. Amer. Soc.*, **77**, 2637-2649.

National Weather Service, 1999: Quantitative Precipitation Forecast Assessment. National Weather Service, NOAA, U.S. Department of Commerce, 62 pp. See also: http://www.nws.noaa.gov/er/hq/QPF/.

Office of Meteorology, 1999: The Modernized End-to-End Forecast Process for Quantitative Precipitation Information: Hydrologic Requirements, Scientific Issues, and Service Concepts. National Weather Service, NOAA, U.S. Department of Commerce, 187 pp. See also: http://tgsv5.nws.noaa.gov/om/qpi_all_final.pdf.

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.