P14.4 **EVALUATION OF RADAR PRECIPITATION ESTIMATES FROM NMQ AND WSR-88D DPA PRODUCTS OVER CONUS: PRELIMINARY RESULTS**

Wanru Wu* and David Kitzmiller
Office of Hydrologic Development, NOAA National Weather Service, Silver Spring, MD

1. INTRODUCTION

Radar and multi-sensor precipitation estimates are used extensively in NOAA National Weather Service (NWS) operations. The National Mosaic and Multi-sensor Quantitative Precipitation Estimation system (NMQ; Vasiloff et al., 2007; Zhang et al., 2005, 2009) is a joint effort among several NOAA offices, intended to test and demonstrate quantitative precipitation estimation (QPE) algorithms that are not currently implemented in the NEXRAD Precipitation Processing System (PPS; Fulton et al., 1998). Following favorable test results and positive feedback from field users, the NWS is investigating options for implementing NMQ operationally. Toward that end, we have undertaken a broad-scale evaluation and comparison of NMQ and currently-operational PPS QPE products, by using real-time data from most of the conterminous United States (CONUS). This effort will provide valuable guidance to field personnel, who will have access to both NMQ and PPS products and who must select and blend input from the two sources in operations. It will also serve to guide future algorithm improvement and implementation efforts.

The data used in this study are collected from the real-time NMQ prototype system, developed and maintained by the National Severe Storms Laboratory, and from mosaicked WSR-88D Digital Precipitation Array (DPA) products, in the form of Stage2 gridded fields prepared operationally by the National Centers for Environmental Prediction (NCEP) (Lin and Mitchell, 2005). The goals are to explore and better understand when and where NMQ and DPA products differ; how to advise field staff to use NMQ; where improvements to NMQ and the PPS are needed; and what are the current geographic limitations of radar QPE coverage. This paper presents the preliminary results.

In Section 2, both NMQ and Stage2 radar QPEs are compared with 24-hour total rain gauge reports from the Automated Surface Observing System (ASOS), which are of generally high quality and utilize a weighing mechanism less subject to mechanical error than the tipping-bucket mechanism commonly used at automated reporting sites. We evaluate the two radar QPE systems against ASOS rain gauge observations in terms of correlation, root mean squared error and mean error. Statistics are also stratified by geography and rainfall regime, in particular: (1) all cases with valid data, (2) all cases with either radar or rain gauge ≥0.25 mm, (3) all cases with radar or gauge ≥10 mm, and (4) all cases with either radar or gauge ≥25 mm. For the stratifications listed above, the initial results based on daily accumulations indicate that NMQ estimates generally have smaller errors and less statistical bias than do the DPA products. In the future, this analysis will be extended to 6-h and 1-h rainfall.

Another analysis, for effective areal coverage of radar QPE, is presented in Section 3, to identify the areas where spatial discontinuities in radar coverage are suspected. Information on radar coverage is essential to optimum blending of radar information and data from other sources including rain gauge, satellite, and numerical model output. In Section 4 we summarize the preliminary results obtained at present and discuss our ongoing further research work for this study.

* *Corresponding author address:* Dr. Wanru Wu, NOAA/NWS/OHD, Silver Spring, MD 20910;
e-mail: Wanru.Wu@noaa.gov.

2. STATISTICAL ANALYSIS FOR ACCURACY OF RADAR QPEs

Our analyses for DPA QPEs are based on NCEP Stage2 radar-only data. The DPA products contain hourly precipitation accumulated over elements of a 4-km polar stereographic grid (Fulton et al., 1998, Lin and Mitchell, 2005). The values were accumulated for 24-hour periods ending at 1200 UTC, and then collocated with the ASOS sites by extracting the Stage2 value of the grid box in which the site lies.

While the NMQ data are produced for a grid with approximately 1 km mesh length (Zhang et al., 2009), most current NWS hydrologic applications are based on the 4-km DPA grid, and therefore we averaged the NMQ QPEs over those same grid elements to insure consistency with future operational use. The NMQ precipitation processing contains quality-control and dynamic Z-R selection features which are not included in the current PPS. Early comparisons indicate these features have substantial positive impact on the overall quality of the radar QPE.

Our intent is to examine “radar-only” products, to avoid complications in interpreting results and to isolate any PPS-NMQ differences to those arising from differences in radar data processing. During the preparation of this analysis, we discovered that the Stage2 mosaics were not, in fact, based purely on radar input, but had been corrected for the local forecast offices’ automated estimates of the mean field gauge/radar bias (Ying Lin, personal communication). However, we expect that the bias correction should improve performance of the DPA products; therefore our findings likely represent an upper bound on the performance of the DPA products without bias correction applied. These tests will be repeated with the original DPA products, as originally intended.

Radar and rain gauge estimates at approximately 375 ASOS sites across CONUS were examined to determine NMQ-to-gauge and Stage2-to-gauge linear correlation and root mean squared error for all valid 24-hour precipitation accumulations and for the amount larger than 0.25, 10, and 25 mm detection. The overall procedure is illustrated in Fig. 1. The temporal domain is from 1 December 2008 to 31 August 2009, covering both cool season (December-March) and warm season (May-August).

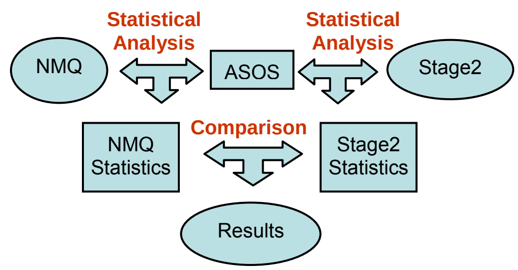

Figure 1. Conceptual diagram for statistical evaluation of radar-only QPEs from NMQ and Stage2 against ASOS.

Statistics for the entire CONUS are listed in Tables 1-3, for the four rain amount stratifications described above. The sample size is shown in the “Count” column, “R” represents the linear correlation coefficient and “RMSE” the root mean squared error in mm.

Table 1. Correlation coefficient (R) and root mean squared error (RMSE, in mm) of 24-h precipitation accumulation from 1 December 2008 to 31 August 2009, between NMQ and Stage2 radar QPEs versus ASOS rain gauge observations (NMQ/ASOS and ST2/ASOS), respectively. **NOTE: results may change following re-analysis with DPA (Stage2) data without gauge/radar bias correction.**

Case Category	Count	R NMQ/ ASOS	R ST2/ ASOS	RMSE NMQ/ ASOS	RMSE ST2/ ASOS
all valid	96,328	0.87	0.64	4	8
≥0.25mm	44,180	0.85	0.60	5	12
≥10mm	11,112	0.78	0.37	11	24
≥25mm	4,087	0.72	0.15	15	37

Table 2. Same as Table 1 but for cool season from 1 December 2008 to 31 March 2009.

Case Category	Count	R NMQ/ ASOS	R ST2/ ASOS	RMSE NMQ/ ASOS	RMSE ST2/ ASOS
all valid	42,896	0.86	0.67	3	5
≥0.25mm	16,221	0.84	0.63	5	9
≥10mm	3,053	0.73	0.40	11	19
≥25mm	885	0.63	0.14	17	33

Table 3. Same as Table 1 but for warm season from 1 May 2009 to 31 August 2009.

Case Category	Count	R NMQ/ ASOS	R ST2/ ASOS	RMSE NMQ/ ASOS	RMSE ST2/ ASOS
all valid	42,693	0.87	0.61	4	10
≥0.25mm	22,679	0.86	0.57	6	14
≥10mm	6,754	0.79	0.34	11	26
≥25mm	2,712	0.74	0.13	15	38

The results from Tables 1-3 show that, when aggregated over all 24-hour accumulation cases, NMQ data have better scores (higher correlation coefficient and lower root mean squared error) with ASOS compared to Stage2, especially for larger precipitation amount (≥10 mm and ≥25 mm). The absolute radar-to-gauge mean errors (not listed) are generally smaller in the cool season and larger in warm season for both NMQ and Stage2 in all four precipitation amount categories. As displayed in Fig.2, individually NMQ have better scores in warm season and Stage2 have better scores in cool season.

Note that a statistical relationship within two samples can be strong and yet not statistically significant; conversely, a relationship can be weak but significant. The key factor is the size of the sample. Since the sample numbers (N) are large in our study (N > 500), the z score can be estimated on the upper bond by $z \approx r\sqrt{N}$ for correlation coefficient r. By putting into critical values of z 1.64, 1.96 and 2.58 for 90%, 95% and 99% significance levels, respectively, the corresponding correlation coefficients are estimated as in Table 4. We see that the correlation coefficients listed in Tables 1-3 are all above the 99% significance level. Similarly, when the differences of correlation coefficients are larger than these values, they are considered statistically significant. Thus, the individual differences between cold season and warm season are statistically significant for NMQ/ASOS cases ≥10 mm and ≥25 mm and for Stage2/ASOS all cases but large rainfall detection ≥25 mm.

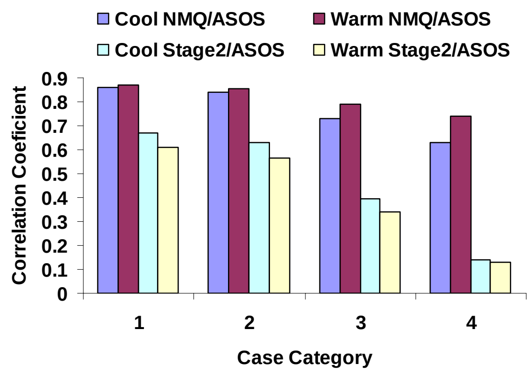

Figure 2. Correlation coefficients of NMQ vs. ASOS and Stage2 versus ASOS 24-h precipitation in cool season and warm season, respectively. Case categories 1-4 denote the same four as listed in Tables 1-3.

Table 4. The 90%, 95%, and 99% two-tailed significance levels of the correlation coefficient for typical sample numbers.

N	90%	95%	99%
42,000	0.008	0.010	0.013
16,000	0.013	0.015	0.020
3,000	0.030	0.036	0.047
1,000	0.052	0.062	0.082
500	0.073	0.088	0.115

To identify any geographical dependency in the verification statistics, the correlation between gridded QPE and ASOS observations at individual sites was calculated and plotted (Fig. 3). The major differences are located in the western mountainous United States, with lower correlations shown in the Stage2/ASOS map (Fig. 3a). Similar results (not shown) are also obtained from spatial maps with the cool and warm seasons considered separately.

The NMQ and DPA quality differences likely arise from several factors. In this analysis, we did not attempt to limit the comparison to cases with rainfall at the surface. The NMQ Z-R selection algorithm specifically seeks to identify cases with snow at the surface, based on reflectivity profile and Rapid Update Cycle model (Black et al., 2005) temperature input, and provides a water-equivalent estimate, while PPS does not. This difference might account for some of the differences particularly over the Western United States. Some differences might be also attributed to advanced automated quality control features within NMQ that are not incorporated with the DPA products. We will test specifically for radar areal coverage by light anomalous rainfall amounts which radar sometimes produces when quality control does not properly identify insects and migrating birds.

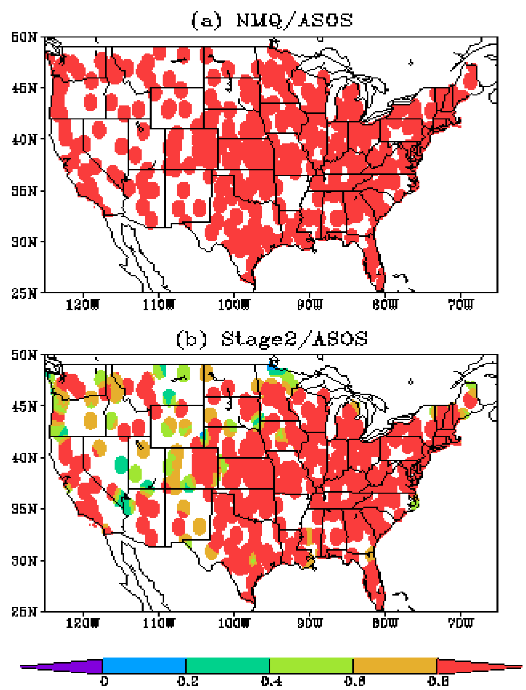

Figure 3. Linear correlation of 24-h precipitation accumulations between ASOS rain gauge observations and (a) NMQ radar QPE; and (b) Stage2 radar QPE, respectively, from 1 December 2008 to 31 August 2009.

Regional statistics for the Eastern-only United States (East of 90°W) are calculated to exam further details of the two radar QPE systems and the results are listed in Tables 5-7, in which the items are same as described above for Tables 1-3. While the statistical results for NMQ versus ASOS are relatively stable, the Stage2 versus ASOS scores are largely improved, with higher correlation coefficients and lower root mean squared errors compared to the results for CONUS, especially in the cool season (displayed also in Fig. 4). In addition, the correlation coefficients of Stage2 versus ASOS are significantly higher in the cool season than in the warm season.

Table 5. As in Table 1, but only for sites in the eastern United States (east of 90°W).

Case Category	Count	R NMQ/ ASOS	R ST2/ ASOS	RMSE NMQ/ ASOS	RMSE ST2/ ASOS
all valid	41,400	0.87	0.74	4	6
≥0.25mm	23,081	0.85	0.70	6	8
≥10mm	6,377	0.74	0.48	11	15
≥25mm	2,092	0.64	0.24	16	24

Table 6. Same as Table 5 but for cool season from 1 Dec 2008 to 31 March 2009.

Case Category	Count	R NMQ/ ASOS	R ST2/ ASOS	RMSE NMQ/ ASOS	RMSE ST2/ ASOS
all valid	18,430	0.86	0.79	4	5
≥0.25mm	9,168	0.84	0.76	6	7
≥10mm	2,088	0.72	0.57	11	14
≥25mm	563	0.60	0.36	18	23

Table 7. Same as Table 5 but for warm season from 1 May 2009 to 31 August 2009.

Case Category	Count	R NMQ/ ASOS	R ST2/ ASOS	RMSE NMQ/ ASOS	RMSE ST2/ ASOS
all valid	18,364	0.87	0.71	5	8
≥0.25mm	11,316	0.85	0.67	6	10
≥10mm	3,590	0.74	0.43	11	17
≥25mm	1,304	0.65	0.17	16	25

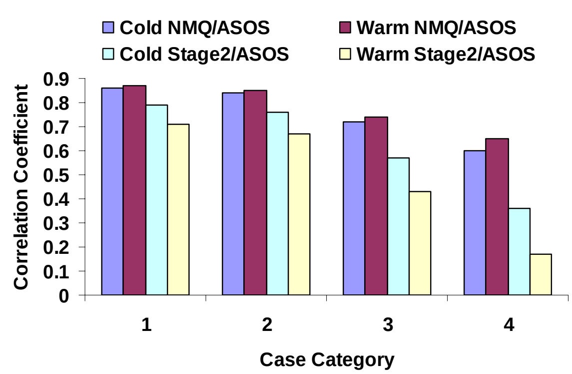

Figure 4. Same as Fig. 2 but for Eastern United States only (East of 90°W).

3. STATISTICAL ANALYSIS FOR EFFECTIVE AREAL COVERAGE

Multi-sensor merging, such as gauge-radar or satellite-radar, depends on accurate knowledge of coverage limitations, particularly for radar. Radar estimates are affected not only by radar site elevation and precipitation climatology, whose effects on precipitation detection are generally spatially continuous, but also by terrain and other beam-blocking objects, which can cause large horizontal gradients in the radar QPE field. This property of radar QPE means that its quality can range from excellent to very poor within the space of a few kilometers. Over mountainous regions, it is impossible to extrapolate radar QPE quality based on point data from sites with rain gauge reports.

In the NWS Multi-sensor Precipitation Estimator (MPE; Seo 1998; Seo et al. 1999; Fulton 2005; Glaudemans et al., 2008), radar coverage boundaries are accounted for through a set of subjectively-derived binary grids depicting long-term radar detection efficiency around each radar (the "misbin" grid, see Fulton, 2005). In multi-radar or multi-sensor QPE analysis, areas outside the coverage zone of one radar are either covered by another radar, or by rain gauge or satellite estimates.

In an attempt to define the quality of radar coverage objectively, and in terms of a continuous measure that might enable application of compromised but useable radar input, we correlated the radar QPE to a gridded continuous reference rainfall field. This reference field is daily 24-hour accumulations from the North American Mesoscale (NAM) model (Black et al. 2005), which provides a spatially-continuous estimate of rainfall that, in the mean, accurately simulates terrain-dependent features of local precipitation climatology. The 24-hour NAM precipitation was taken from the sum of the 0-12 h forecasts from the 0000 and 1200 UTC runs. An initial experiment was conducted for the northwestern CONUS, which features sharp gradients in climatic precipitation and numerous terrain features causing radar beam blockages.

We determined that the radar-NAM correlation is an effective proxy for radar-rain gauge correlation by comparing the two sets of statistics at several hundred individual daily-reporting sites in this region (Fig. 5). Each point in the figure represents the radar-NAM and radar-gauge correlation over the cool seasons 2007-2009 for one gauge site. The radar-NAM correlation explains almost 80% of the variance in the radar-gauge correlation; therefore we can infer the overall quality of radar QPE data from its correlation to NAM precipitation simulations, at least in the cool season in mid-latitudes.

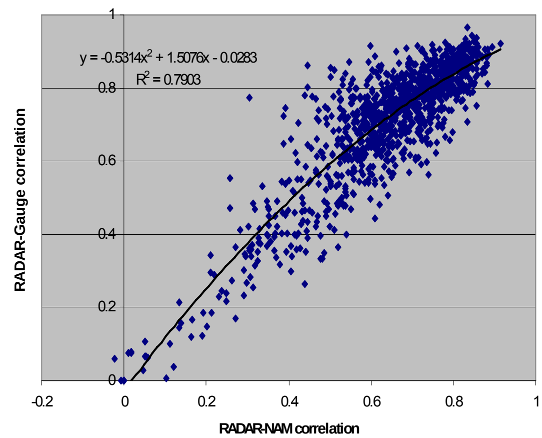

Figure 5. Correlations of 24-h precipitation between Stage2 radar QPEs and NAM forecasts (x-axis) vs. that between Stage2 radar QPEs and daily climatic network rain gauge reports in cool seasons of 2006-2009. Each point represents the Stage2-NAM and radar-gauge correlations at an individual gauge site over the time period.

By extending this analysis to the continuous 4-km Stage2 grid, we obtained the correlation field for the northwestern CONUS shown in Fig. 6. The field is highly consistent with known blockage and coverage limitation feature; these include the general decrease in radar's precipitation detection with range (the area of low correlations over west-central Oregon), and terrain blockages (radial spike patterns in northeastern Washington and western Montana). Areas in eastern Oregon and central Idaho which are beyond the PPS coverage limit of 230 km are shown with zero correlation to NAM precipitation.

In the future, this analysis will be repeated with NMQ QPEs, and extended to the remainder of the CONUS and to the warm season. We anticipate that during the warm season a combination of NAM simulations and satellite estimates will provide the best reference precipitation field, rather than NAM forecasts alone. The correlation grids themselves will be applied in multi-sensor merging, to provide estimates of relative weights for radar and other QPE sources such as satellites.

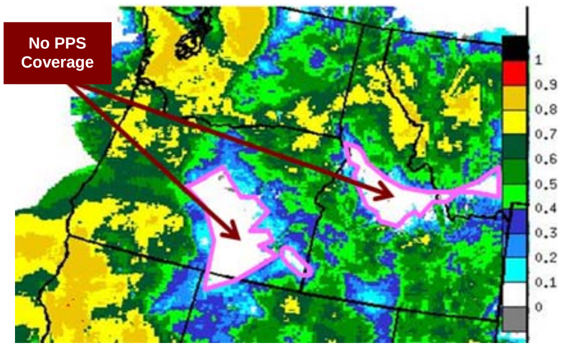

Figure 6. Linear correlation of 24-h precipitation between NAM simulations and Stage2 radar estimates in cool seasons of 2006-2009. Each datum in the map represents the Stage2-NAM correlation at that geographic location.

4. SUMMARY

This study evaluates the radar QPEs from NMQ and from Stage2 products, with verification from ASOS rain gauge reports and from a reference rainfall field from NAM simulations. The preliminary results indicate that NMQ radar precipitation estimates have better statistical scores relative to ASOS observations. The Stage2 product scores are appreciably lower for higher-rainfall categories (≥10mm and ≥25mm) than for lower amount categories. NMQ radar QPEs have higher scores in warm season and lower scores in cool season, contrary to Stage2 radar QPEs, which have higher scores in cool season and lower scores in warm season. The Stage2 and NMQ correlations are expected to have differences in some areas due to NMQ's more advanced automated quality control and adaptive Z-R selection features.

We will later verify NMQ and Stage2 QPEs against Stage4 precipitation estimates, which are generated at River Forecast Centers (RFCs) with manual quality control and rain gauge input, and are considered the best gridded rainfall estimates available in near-real time.

The Stage2-to-NAM precipitation correlation (Fig. 5) shows limits for radar coverage from the DPA products. We will extend this analysis to the warm season and apply it to NMQ products, In the future, such correlation field information could be used in multi-sensor merging in MPE and NMQ multi-sensor products.

The results presented in this paper are obtained from 24-hour precipitation accumulation. Further evaluation of the radar estimates at different time scales will be carried out in order to assess the QPE accuracy and statistical distribution for different operational purposes. For instance, the 1-hour time scale is presently applied in some distributed and lumped operational hydrologic models, and we can anticipate that such uses will expand in the near future. The 6-hour interval matches the time step currently used in most operational lumped river models, and is a standard for gauge-only analysis in much of the western United States. The 24-hour interval corresponds to daily observations, which are collected at many more points than are sub-daily observations, and daily precipitation reports are routinely used in some hydrologic operations.

ACKNOWLEDGEMENTS

This research is supported by the NWS Advanced Hydrologic Prediction Service (AHPS) program. Hollings Scholar Robert Setzenfand (Pennsylvania State University) contributed extensively to the analysis of precipitation estimates in mountainous terrain.

REFERENCES

Benjamin S. G., and coauthors, 2004: An hourly assimilation/forecast cycle: The RUC. *Mon. Wea. Rev.*, **132**, 495–518.

Black T. L., M. Pyle, and H.-Y. Huang, 2005: The operational WRF-NMM at NCEP. *Preprints, 21st Conf. on Weather Analysis and Forecasting/17th Conf. on Numerical Weather Prediction,* Washington, DC, Amer. Meteor. Soc., 4B.4. [Available online at http://ams.confex.com/ams/pdfpapers/94734.pdf]

Fulton, R., J. Breidenbach, D.-J. Seo, D. Miller, and T. O'Bannon, 1998: The WSR-88D Rainfall Algorithm. *Wea. Forecasting*, **13**, p377–395.

_____, 2005: Multisensor Precipitation Estimator (MPE) Workshop. [Available from http://www.nws.noaa.gov/oh/hrl/papers/wsr88d/MPE_workshop_NWSTC_lecture2_121305.pdf]

Glaudemans, M., P. Tilles, and B. Lawrence, 2008: Interactive quality control and operational product generation of hourly multi-sensor precipitation estimates in the NWS. Preprints *23rd Conference on Hydrology*, New Orleans, Amer. Meteor. Soc., **5B.3**.

Lin, Y., Mitchell, K.E., 2005. The NCEP Stage II/IV hourly precipitation analyses: development and applications. *Preprints, 19th Conference on Hydrology*. San Diego, Amer. Meteor. Soc., **1.2**.

Seo, D.-J., 1998: Real-time estimation of rainfall fields using radar rainfall and rain gauge data. *J. Hydrol.,* **208**, 37-52.

_____, D.-J., J. Breidenbach, and E. Johnson, 1999: Real-time estimation of mean field bias in radar rainfall data. *J. Hydrol*., **223**, 131-147.

Vasiloff, S., D.J., Seo, K. Howard, J. Zhang, D. H. Kitzmiller, M. G. Mullusky, W. F. Krajewski, E. A. Brandes, R. M. Rabin, D. S. Berkowitz, H. E. Brooks, J. A. McGinley, R. J. Kuligowski, and B. G. Brown., 2007: Q2: Next generation QPE and very short-term QPF. *Bull. Amer. Met. Soc.* **88**, 1899-1911.

Zhang, J., K. Howard, and J.J. Gourley, 2005: Constructing three-dimensional multiple-radar reflectivity mosaics: Examples of convective storms and stratiform rain echoes. *J. Atmos. Oceanic Technol.*, **22**, 30–42.

______, and coauthors, 2009: National Mosaic and QPE (NMQ) System – Description, results and future plans. *Preprints 34th Conference on Radar Meteorology*, Williamsburg, Amer. Meteor. Soc., **7A.1.**