

Evaluating NEXRAD Multisensor Precipitation Estimates for Operational Hydrologic Forecasting

C. BRYAN YOUNG, A. ALLEN BRADLEY, WITOLD F. KRAJEWSKI, AND ANTON KRUGER

*Iowa Institute of Hydraulic Research and Department of Civil and Environmental Engineering,
The University of Iowa, Iowa City, Iowa*

MARK L. MORRISSEY

Environmental Verification and Analysis Center, University of Oklahoma, Norman, Oklahoma

(Manuscript received 9 August 1999, in final form 10 January 2000)

ABSTRACT

Next-Generation Weather Radar (NEXRAD) multisensor precipitation estimates will be used for a host of applications that include operational streamflow forecasting at the National Weather Service River Forecast Centers (RFCs) and nonoperational purposes such as studies of weather, climate, and hydrology. Given these expanding applications, it is important to understand the quality and error characteristics of NEXRAD multisensor products. In this paper, the issues involved in evaluating these products are examined through an assessment of a 5.5-yr record of multisensor estimates from the Arkansas–Red Basin RFC. The objectives were to examine how known radar biases manifest themselves in the multisensor product and to quantify precipitation estimation errors. Analyses included comparisons of multisensor estimates based on different processing algorithms, comparisons with gauge observations from the Oklahoma Mesonet and the Agricultural Research Service Micronet, and the application of a validation framework to quantify error characteristics. This study reveals several complications to such an analysis, including a paucity of independent gauge data. These obstacles are discussed and recommendations are made to help to facilitate routine verification of NEXRAD products.

1. Introduction

In recent years, the National Weather Service (NWS) has installed the Next-Generation Weather Radar (NEXRAD) system at forecast offices across the country. The NEXRAD system consists of a network of Weather Surveillance Radar-1988 Doppler (WSR-88D) radars (Crum et al. 1993). Reflectivity observations from each WSR-88D are used to generate many operational products, including estimates of precipitation developed with the NEXRAD precipitation processing system (Klazura and Imy 1993; Fulton et al. 1998). These radar-based precipitation estimates are used in NWS Forecast Offices by meteorologists and hydrologists for guidance in forecasts and warnings.

In addition, these radar-based estimates are combined with gauge data to make *multisensor* precipitation estimates (Krajewski 1987; Seo 1998). NWS River Forecast Centers (RFCs) produce regional multisensor precipitation products using information from numerous

WSR-88D radars and a network of gauges that report observations to the RFC in near-real time (Fulton et al. 1998). Multisensor precipitation estimates were developed to replace the gauge estimates used for operational hydrologic forecasting at the RFCs. NEXRAD multisensor estimates also will become widely used by scientists and engineers in nonoperational applications (Hudlow 1988; Johnson et al. 1999). For example, NEXRAD precipitation products will support many research components of the Global Energy and Water Cycle Experiment Continental-Scale International Project. The products will be used in studies of precipitation variability, land-atmosphere water budgets, and coupled modeling of the land surface and atmosphere for climate prediction.

This paper examines issues associated with the evaluation of NEXRAD multisensor estimates through a case study involving a 5.5-yr record of products from the Arkansas–Red Basin RFC (ABRFC). Two needs motivated this evaluation. The first was to examine the effect of known biases in radar estimates of hourly precipitation on the multisensor products (Smith et al. 1996; Young et al. 1999). The second was to evaluate the uncertainty in multisensor precipitation estimates. Combining radar and gauge information should produce

Corresponding author address: Dr. Allen Bradley, The University of Iowa, College of Engineering, Iowa Institute of Hydraulic Research, 301B Hydraulics Laboratory, Iowa City, IA 52242-1528.
E-mail: allen-bradley@uiowa.edu

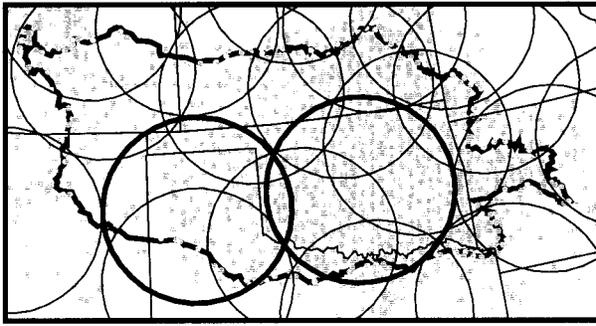


FIG. 1. WSR-88D weather radars and precipitation gauges used in producing the NEXRAD multisensor products at the Arkansas-Red Basin River Forecast Center.

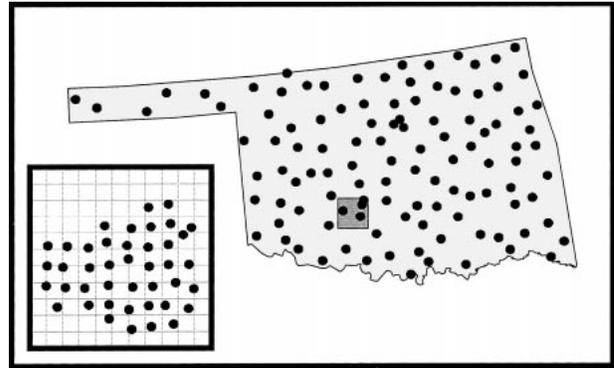


FIG. 2. Locations of Oklahoma Mesonet and ARS Micronet gauges. The dark gray box indicates the location of the ARS Micronet, which is shown with the 4 km \times 4 km multisensor grid in the inset.

improved precipitation estimates, in terms of both quality and spatial resolution, in comparison with either radar or gauge estimates alone. Evaluating the error characteristics of estimates is an essential step to quantify these improvements, validate processing algorithms, and compare competing algorithms.

Still, rigorous evaluation of errors in multisensor precipitation products is more difficult than evaluation of errors in either radar or gauge products. First, using point estimates of precipitation from gauges to verify areal estimates of precipitation from the multisensor products requires careful consideration of the large difference in sampling area represented by both estimates. To address this issue, the error-separation approach proposed by Ciach and Krajewski (1999) was used. NEXRAD multisensor products tend to use all available gauge data for precipitation estimation, however, so few independent gauges remain for error estimation. Another difficulty in evaluation is that the input data and processing algorithms used to make operational precipitation estimates have changed over the years. In addition, operational methodologies involve human interaction, which adds subjectivity to the final estimates. Often, there is little or no record of the data, methods, and manual changes used in producing the final estimates. Such problems are discussed in detail, and recommendations to help to facilitate routine evaluation and verification of NEXRAD products are made.

2. Study area and data resources

For this paper, NEXRAD multisensor hourly precipitation estimates from ABRFC were evaluated. This study area was selected because of the history of NEXRAD research in the southern plains (Smith et al. 1996; Johnson et al. 1999), the availability of gauge observations, and the long period of record for NEXRAD multisensor estimates (May 1993 through September 1998). The NEXRAD record over this period is nearly continuous except for a large gap in 1994 (27 January through 7 March). These multisensor estimates are routinely downloaded from the ABRFC archive and pro-

cessed at the Iowa Institute of Hydraulic Research. To facilitate data analyses, the multisensor estimates were converted to a compact run length encoded format (Krujer and Krajewski 1997).

During the study period, the ABRFC employed two different precipitation processing methodologies to produce multisensor precipitation estimates. The two methods are described in section 3. Both methods combine information from 22 WSR-88Ds with available precipitation data from up to 1586 gauges to produce estimates of hourly accumulation (resolution of 0.01 mm). These estimates are produced over a 335 \times 159 grid for which each grid cell is approximately 4 km \times 4 km. Figure 1 displays the radar coverage (230-km radius) and the gauge locations. Note the high gauge density in the eastern portion of the river basin.

For this study, multisensor estimates were compared with precipitation data from two precipitation gauge networks. Figure 2 shows the locations of gauges from the Oklahoma Mesonet (Brock et al. 1995) and the Agricultural Research Service (ARS) Micronet. The precipitation record available for the 111 mesonet gauges begins in 1994. Data from the mesonet have been included by ABRFC as part of the NEXRAD multisensor precipitation processing. Therefore, the mesonet gauges are not an independent data source for evaluating the NEXRAD products. Additional precipitation data are available from the ARS Micronet stations in the Little Washita River watershed. Precipitation data from the 42 micronet gauges were obtained for June through September 1997 from the ARS archive created for the Southern Great Plains Experiment 1997 (SGP97). This network is separate from the NWS gauge network used in the NEXRAD precipitation processing and thus provides a data source for evaluating the uncertainty of the multisensor estimates. Long-term precipitation records are not readily available for the micronet, however, so the period for evaluation was limited to less than 4 months.

Both the Oklahoma Mesonet and the ARS Micronet use tipping-bucket gauges that record 5-min precipitation accumulations at a 0.254-mm (0.01-in.) resolution.

These data have been aggregated to 1-h accumulations for comparison with NEXRAD products. Basic quality control for the mesonet gauges is carried out by the Oklahoma Climate Survey to identify and to correct problems in the data. Quality control for the micronet gauges is carried out by the ARS, and a data quality flag is indicated with each set of observations. Although gauges often are considered to be ground truth for evaluation of remotely sensed precipitation, they also are subject to measurement error. Studies conducted by Habib et al. (1999, 2000), however, indicate that the random error for hourly gauge observations is negligible and that the systematic error (bias) due to wind effect is small (less than 5%).

Additional hourly precipitation gauge data are available from the National Climatic Data Center (NCDC). Data for many of the gauges archived at NCDC are not available in real time and therefore are not used in NEXRAD precipitation processing. In the preliminary assessment of these data, however, it was found that there are too few gauges for verification throughout the study region, the gauge resolution usually is coarse [2.54 mm (0.1 in.)], and the quality of these data are unknown. Thus, we chose not to include these gauges in the evaluation of hourly precipitation products. NCDC data from both hourly and daily precipitation gauges could be valuable for evaluating NEXRAD products at longer timescales (daily or monthly), though quality issues and the difficulty of identifying which gauges have been included in the multisensor estimates remain.

3. NEXRAD precipitation processing

NEXRAD multisensor precipitation products have been produced at ABRFC using two methods. Both use the same radar-only precipitation product. Their differences lie in the way in which gauge precipitation data are combined with the radar precipitation products. The following sections briefly describe the radar-only product, the two methods (Stage III and P1) for producing multisensor products, and the approximate time line of operational changes in NEXRAD multisensor precipitation products at ABRFC.

a. NEXRAD hourly digital products

At each WSR-88D site, measurements of radar reflectivity factor Z are used to produce a radar-only precipitation product. The algorithms used are described in detail by Fulton et al. (1998). First, a radar-reflectivity field is constructed for each complete radar scan with data from multiple radar tilts. A power-law Z - R relationship then is used to estimate rain rate R . Next, the rain-rate intensity maps are integrated over time to produce hourly accumulations. Quality-control algorithms attempt to identify and to remove outliers and ground returns from the radar precipitation estimates. The results then are transformed to the Hydrologic Rainfall

Analysis Project (HRAP) grid of approximately 4 km \times 4 km. The HRAP grid is a polar stereographic projection that conforms to a 1/40th limited fine mesh grid used by NWS in numerical weather prediction (Reed and Maidment 1999). The resulting product is known operationally at NWS as hourly digital precipitation (HDP). An evaluation of HDP products has been carried out by Smith et al. (1996) for the southern plains and by Young et al. (1999) for the complex mountainous terrain of the northern Appalachians.

b. Stage-III products

The Stage-III methodology was developed at the NWS Hydrologic Research Laboratory (HRL) and was the first approach developed for NEXRAD multisensor precipitation estimation. Stage I of this process is the HDP product described above. In Stage II of the processing, the HDP product for an individual radar is merged with gauge observations to produce a multisensor product (NWS Office of Hydrology 1997a). This step uses the gauge values and the radar estimates at corresponding HRAP cells to estimate the mean field bias in the radar-based precipitation field using a Kalman filter algorithm (Smith and Krajewski 1991; Seo 1998; Anagnostou et al. 1998). The resulting multiplicative bias factor is used to adjust the radar field to produce a bias-corrected radar precipitation field. A gauge-only-based precipitation field is also produced on the HRAP grid by interpolating available gauge values using a distance-weighting scheme. Radar data are used to define the rain/no rain area for this analysis. Last, the gauge-only and the bias-corrected radar fields are merged to produce a multisensor field using an optimal estimation procedure that assigns weights to the two fields based on the proximity of gauges to the HRAP cell.

Although Stage-II processing originally was intended to run at each WSR-88D site, it now is run at RFCs along with Stage-III processing. In Stage III, the stage-II estimates from multiple radars are combined into a single product that covers the entire RFC region (NWS Office of Hydrology 1997b). This combined field is constructed using the average of all Stage-II estimates available for each HRAP cell. Stage III can involve a considerable degree of human interaction. Hydrometeorological analysis forecasters (HAS forecasters) at RFCs are responsible for producing the final Stage-III estimates. They may decide to alter gauge reports that are suspect or to insert "pseudo gauges." The HAS forecaster then reruns the analysis to produce the final Stage-III product.

c. P1 products

The P1 (process 1) methodology was developed at ABRFC and uses processing algorithms developed by the U.S. Army Corps of Engineers (Tulsa district). Un-

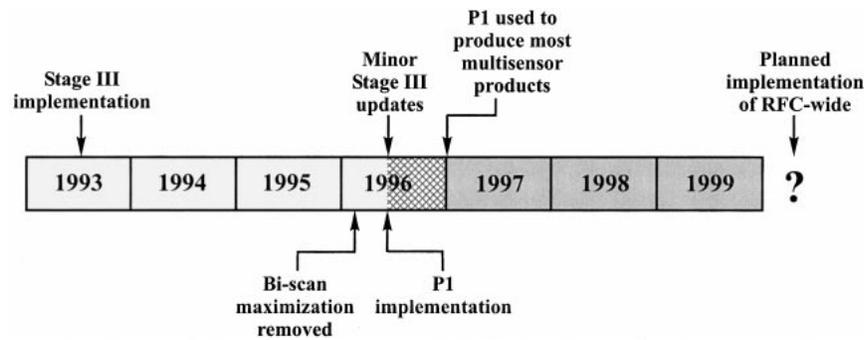


FIG. 3. Time line of major implementations and updates in NEXRAD multisensor processing at the Arkansas–Red Basin River Forecast Center.

like Stage III, in which precipitation products for individual radars are combined near the end of the processing, with P1 the first step is to merge the HDP products to produce a radar-only product for the entire RFC area. The next step computes the ratio between the gauge and radar-only precipitation estimates and assigns the ratio to the HRAP cell that contains the gauge. For HRAP cells that do not contain a gauge, a ratio is computed from nearby cells by interpolation using a distance-weighting scheme. In the final step, the radar precipitation estimates are multiplied by the ratio field to produce the P1 multisensor precipitation product. As with Stage-III processing, there can be considerable interaction by the HAS forecaster to handle quality control problems and to adjust the final product. The P1 also has options that are not available in Stage III, including a “make-snow” option to assign precipitation to an area for which the radar fails to detect winter precipitation.

d. Comparison of Stage-III and P1 methods

The Stage-III and P1 methodologies take vastly different approaches to multisensor precipitation estimation, each with its own strengths and weaknesses. Stage III is designed to merge two quasi-independent precipitation estimates, one based on gauges and the other based on the radar. For two independent, unbiased estimates for which the uncertainty is known, optimal estimation theory can be used to find the best combined estimate. Stage III may have problems if biases are not completely removed before merging. For example, known range-dependent biases in HDP products can affect multisensor estimates adversely. Other problems can occur in situations in which there are inconsistencies in the two estimates, which can arise with light precipitation and snow, because of the inherent difficulties in observing this type of precipitation with radar and gauges. A strength of Stage III is its ability to estimate precipitation when gauge densities are low or with convective precipitation for which spatial variations in precipitation are large and are not detected by the gauge network.

In contrast, P1 is designed to use radar information

as a means for spatially interpolating gauge estimates. The P1 does not take a weighted average of gauge and radar precipitation estimates but rather relies on gauge–radar precipitation ratios. The P1 will have problems when these ratios vary significantly in space. For example, P1 would have significant problems in an hour when a convective storm misses available gauges, because there would be no gauge–radar pairs from which to compute the precipitation ratios. The P1 works better in situations in which precipitation is fairly uniform (e.g., stratiform precipitation) and the gauge–radar ratios are consistent in space. Because P1 relies heavily on gauges for its quantitative estimates, the method is best suited for applications for which the gauge density is high, as in the eastern portion of the ABRFC region, and may perform poorly in gauge-sparse areas. The operational advantage of P1 is that it is less-computationally intensive and makes it easier for the HAS forecaster to adjust estimates, so that final products can be produced in a more timely manner.

e. Time line

Several operational realities make the evaluation of the NEXRAD multisensor precipitation products difficult. First, the algorithms used are not static. Updates and modifications have occurred during the study period. Next, because of the human interaction by the HAS forecaster during processing, both Stage III and P1 are not purely objective precipitation analyses. Another issue is that the gauge data and intermediary radar products used to produce the multisensor estimates are not archived during the analysis. This fact makes it impossible to reprocess the data for algorithm intercomparisons or to evaluate the effect of HAS-forecaster involvement on the final product. Last, indication of the methodology (i.e., Stage III or P1) used by the HAS forecaster to produce a final multisensor product is not archived. For periods during which two algorithms were used interchangeably at ABRFC, the methodology used might not be the same from one hour to the next.

Figure 3 shows a time line of major operational changes in NEXRAD precipitation processing at

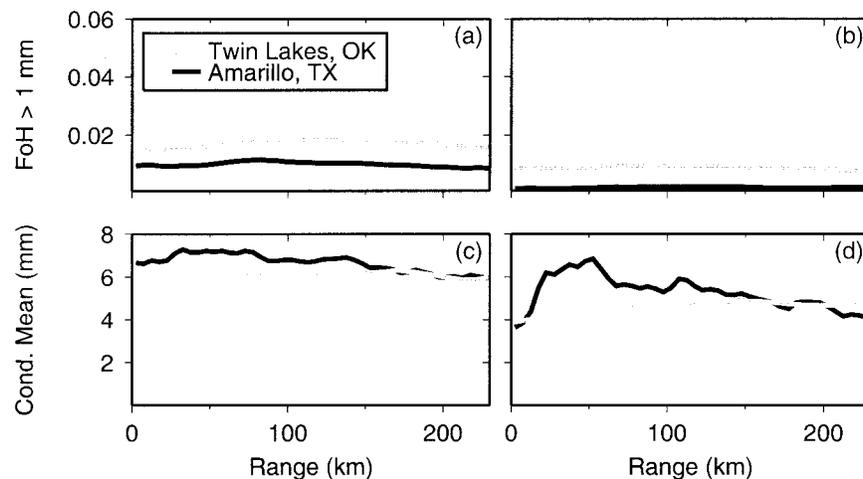


FIG. 4. Stage-III multisensor precipitation characteristics vs range (km) from the radar site for the TLX and AMA radars. Characteristics include the fraction of hours (FoH) recording greater than 1 mm of precipitation for (a) the warm season (Apr–Sep) and (b) the cold season (Oct–Mar), and the conditional mean hourly precipitation (mm) for (c) the warm season and (d) the cold season.

ABRFC. This time line is approximate and was constructed through conversations with NWS personnel at HRL and ABRFC and from information compiled at HRL (D. Johnson 1999, unpublished manuscript). Prior to the summer of 1996, the Stage-III methodology was used exclusively for multisensor precipitation estimation. During a transition period, both Stage III and P1 were used, to varying degrees. From 1997, P1 has been used almost exclusively. Stage III still can be run but is seldom used. Prior to the implementation of P1 algorithms, one major update occurred that affects the analysis of the Stage-III products. In February 1996, the bscan maximization option was turned off. This option used the maximum radar reflectivity from the first two radar tilts for certain ranges from the radar for precipitation estimation. The investigation by Smith et al. (1996), which recently was verified by Anagnostou and Krajewski (1998), showed that brightband effects produce anomalously high accumulations with this option, prompting the change in 1996. The next major update on the horizon is the implementation of a new RFC-wide algorithm for multisensor precipitation estimation that is being developed at HRL.

As a result of these changes, the multisensor products from ABRFC were divided into two major periods. The Stage-III period runs through January 1996, and the P1 period begins in January 1997. The period from February through December 1996 is not included in any of the analyses because of the major updates that occurred and the mixed use of Stage-III and P1 methodologies during this time.

4. Product intercomparisons

The initial evaluation compared NEXRAD multisensor estimates based on the Stage-III and P1 meth-

odologies. The effects of well-known biases in radar-only products were examined through analyses that look at variations in estimates with range from the radar, long-term precipitation accumulations over the entire ABRFC region, and gauge–multisensor comparisons using the mesonet and micronet data for Oklahoma.

a. Range-dependent variations

Previous evaluations of the radar-only HDP product have shown that there are significant range-dependent errors (Smith et al. 1996; Young et al. 1999). The study by Smith et al. (1996) examined HDP products in the southern plains region from September 1993 through December 1994. They found that characteristics such as precipitation detection and conditional mean hourly precipitation vary with range from the radar. One reason for range-dependent biases in HDP products is that the radar sampling volume and beam height increase with range from the radar. Another reason identified by Smith et al. (1996) was the bscan maximization algorithm, which enhances brightband echoes (e.g., anomalously high reflectivities from melting snow) and overestimates precipitation in the 50–75-km range.

Stage-III and P1 methodologies both use HDP products to produce multisensor precipitation estimates. Figures 4 and 5 show the range-dependent characteristics for Stage-III and P1 periods for the Twin Lakes (TLX) and Amarillo (AMA) radars. After Smith et al. (1996), results are shown for the warm (April–September) and cold (October–March) seasons. The locations of the two radars are highlighted in Fig. 1. The TLX radar is located in central Oklahoma in the more-humid, eastern half of the ABRFC region, in which the gauge density is very high. The AMA radar is located in the Texas

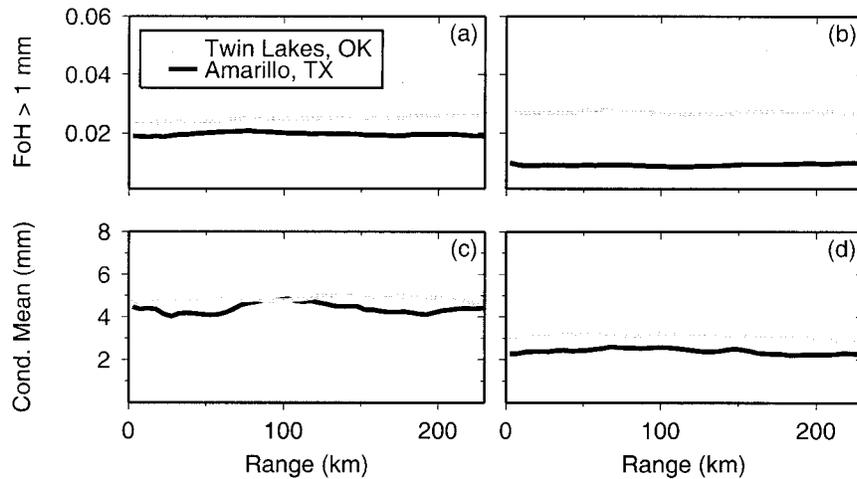


FIG. 5. Same as Fig. 4 but for P1.

Panhandle in the semiarid western half of the region, in which the gauge density is much lower.

For each methodology, the fraction of hours that record more than 1 mm of precipitation does not vary greatly with range from the radar in either season. As one would expect, the occurrence of precipitation greater than this threshold is higher around the TLX radar than around the AMA radar. The estimated conditional mean hourly precipitation (i.e., conditioned on occurrence) shows some slight variations with range for the Stage-III methodology. The variations are most pronounced around the AMA radar, especially in the cold season. The pattern of lower conditional averages very close to the radar, higher values in the 50–150-km range, and lower values at farther ranges also was observed for the HDP products (Smith et al. 1996). Still, such variations are not seen for the TLX radar. This result probably is due to the differences in gauge and radar densities for these two sites. With fewer gauges and radars in the western region, the multisensor product reduces but does not eliminate range-dependent biases in HDP products. For the P1 product, there are no important variations with range, even around the AMA radar. For this period, the biscan maximization algorithm was not in use, which reduces the range-dependent biases in the HDP products. More important, though, the P1 methodology uses the HDP information only to aid in interpolation of the gauge estimates, so biases in HDP can affect only the local interpolation between gauge observations.

Figures 4 and 5 also show that there are large differences in the precipitation estimates for the two methodologies, especially in the cold season. The occurrence of precipitation (>1 mm) is much higher for the P1 period. Conversely, the conditional mean is much higher for the Stage-III period. These differences tend to compensate for one another, resulting in similar estimates for the total accumulations (not shown). This result sug-

gests that the differences are related to the algorithm estimates for light precipitation. Differences in precipitation detection by radar and gauge, as well as the algorithm thresholds used in P1 and Stage-III processing to define (nonzero) precipitation, mostly likely create the observed differences. In addition, with the P1 methodology, the HAS forecaster has the capability to make snow or effectively to add areas of precipitation based on ancillary information of the storm situation. This capability also would contribute to the higher precipitation occurrence in the cold season.

b. Multisensor accumulations for the ABRFC region

Another means for comparing the two methodologies is to examine the total precipitation accumulation over the region. Figure 6 shows the accumulations for the Stage-III and P1 periods. For both periods, the strong west-to-east gradient in precipitation is observed. Still, some artifacts of the radar biases and data processing errors are present. For Stage III, there are pronounced circular features associated with the individual radars. Because the multisensor precipitation estimates are produced first for the individual radars and then are combined in Stage III, these features may be due to systematic biases between HDP products from neighboring radars (Smith et al. 1996). In the P1 processing, the HDP products first are combined by averaging the estimates from multiple radars and then are used to guide the gauge interpolation. The results therefore appear smoother and do not show the same circular features. Another artifact of the radar biases is the faint “spokes” seen emerging from the radars in the Stage-III period. Similar features are present in the P1 accumulations but are not as obvious without more detailed analysis. These spokes likely are due to systematic underestimation for specific beam radials resulting from partial blockage of

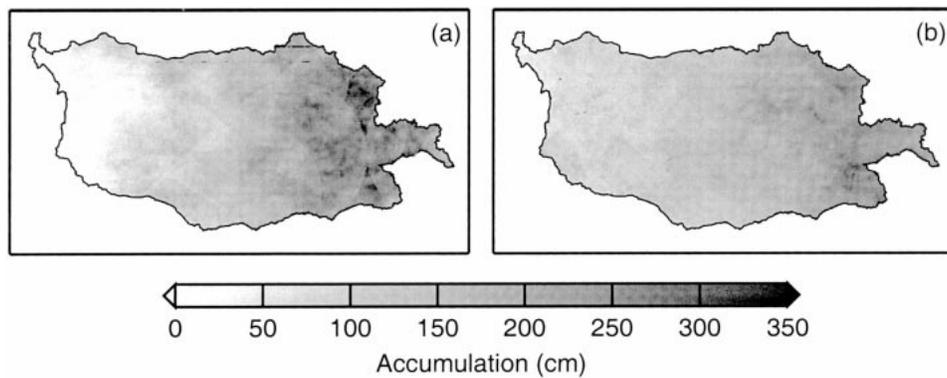


FIG. 6. NEXRAD multisensor precipitation accumulations (mm) for (a) Stage III, May 1993–Jan 1996 and (b) P1, Jan 1997–Sep 1998.

the radar beam (e.g., a tower or other obstacle in the beam path).

An artifact of the data processing is indicated by two horizontal black lines in the northern portion of the ABRFC region for the Stage-III products. These lines reflect abnormally high precipitation accumulations recorded for many hours in 1995. A second data-processing error that is not evident in the total accumulations but can be seen at shorter durations (e.g., 1 month), is the presence of isolated cells with zero accumulation. These cells are associated with gauge locations that failed to report precipitation but were assumed in the data processing to be a report of zero accumulation. Both of these data-processing errors are confined to the early stages of NEXRAD usage and have been largely corrected in more recent operational use (B. Lawrence 1999, personal communication).

c. Gauge–multisensor precipitation comparison

Any detailed comparison of gauge and NEXRAD products is complicated by the large difference in sampling areas for the two estimates. NEXRAD multisensor products are recorded on a 4 km × 4 km grid, and a typical rain gauge only covers an area of 0.3 m². Still, valuable information can be obtained from a simple

comparison. For example, Smith et al. (1996) found significant underestimation of precipitation by radar in the ABRFC region; gauge precipitation exceeded radar estimates by 14% to 100%, depending on season and range from the radar. Merging of radar and gauge data in the NEXRAD multisensor products should decrease these differences.

Figure 7 displays 3-hourly scatterplots of NEXRAD multisensor estimates for cells that contain gauges versus the corresponding gauge observations. Figure 7a compares hourly Stage-III estimates with observations from the 111 Oklahoma Mesonet gauges, and Figs. 7b and 7c plot P1 estimates versus gauge observations for the Oklahoma Mesonet and ARS Micronet, respectively. Table 1 presents the bias and root-mean-square difference (rmsd) for each of these graphs.

As was the case with radar precipitation estimates in this region (Smith et al. 1996), Stage-III estimates tend to underestimate gauge precipitation (−21.5% bias). In contrast, P1 estimates tend to overestimate gauge precipitation (5.2% bias). In both cases, this bias exists despite the fact that mesonet gauges are included in the multisensor processing. For the comparison between P1 and micronet gauges in Fig. 7c, the bias is higher (19.7%) than that recorded for the mesonet gauges. This difference could be due to seasonality or to the short

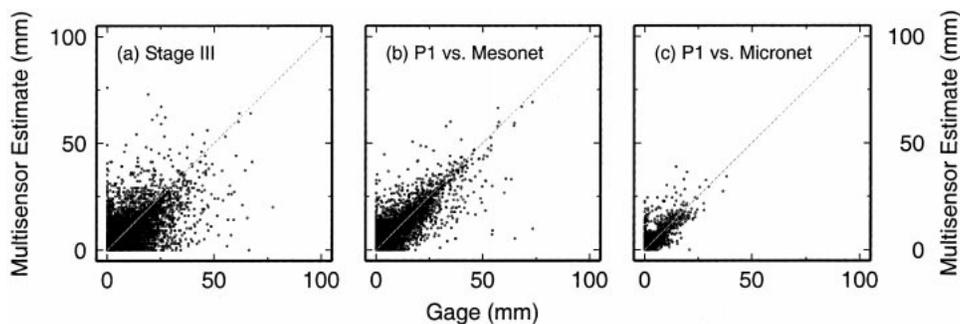


FIG. 7. Hourly NEXRAD multisensor precipitation accumulations (mm) plotted against gauge observations for (a) mesonet gauges (Jan 1994–Jan 1996), (b) mesonet gauges (Jan 1997–Jun 1998), and (c) micronet gauges (Jun 1997–Sep 1997).

TABLE 1. Multisensor bias and rmsd from gauge observations.

Gauge network	Algorithm	Bias (%)	Rmsd (mm)
Mesonet ^a	Stage III	-20.5	2.66
Mesonet ^b	P1	+5.17	1.56
Mesonet ^c	P1	+19.7	2.68

^a Jan 1994–Jan 1996.

^b Jan 1997–Jun 1998.

^c Jun–Sep 1997.

record length of the micronet data but also could reflect the fact that the micronet is an independent data source.

For graphs that compare multisensor products with mesonet gauges (Figs. 7a,b), the scatter is visibly lower for P1 estimates than for Stage-III estimates. Indeed, rmsd decreases from 2.66 to 1.56 mm from the Stage-III to the P1 period. This decrease is due to the heavy weighting that gauge estimates receive in P1 and is not necessarily an indication of superior performance. In fact, given that the P1 algorithm interpolates gauge–radar ratios, the large scatter in Fig. 7b is surprising. There are several possible explanations for this scatter. In many instances, gauge observations are not reported with sufficient lead time to be included in the P1 analysis. Human interaction in the estimation process also may lead to discrepancies. Unfortunately, information on which gauges are used in NEXRAD processing for each individual hour is unavailable, making it extremely difficult to sort out the causes for the scatter.

For both Stage III and P1, the scatterplots also reveal that the multisensor estimate occasionally is equal to zero despite positive gauge accumulation. Given the differences in sampling areas of the two estimates (i.e., gauge and radar), this occurrence should be rare. Table 2 presents the fraction of hours for which the multisensor estimate was greater than zero given that the corresponding gauge observation was greater than a given threshold. These conditional probabilities were estimated using the 111 mesonet gauges. The probabilities are very different for the two methodologies. The results for the Stage-III product show a poor correspondence at low gauge thresholds. The conditional probabilities are higher for the P1 period, especially for low gauge accumulations. These results are another indication of the differences in estimation for light precipitation by the two methods and suggest that the Stage-III method misses many occurrences of light precipitation. The differences likely lie in the relative weighting of HDP products and gauge information in the processing. Many known factors may prevent radar detection of precipitation at the ground surface, including beam overshoot of low precipitating clouds, horizontal advection of hydrometeors, false elimination of rain in anomalous propagation algorithms, and the use of thresholds in NEXRAD processing to distinguish precipitation. The heavier weighting of the HDP estimates by Stage III likely results in considerably fewer hours with light precipitation. Still, major and minor updates to the algorithms

TABLE 2. Fraction of multisensor hours recording precipitation, conditioned on mesonet gauge observations exceeding X .

X (mm)	Fraction of hours > 0	
	Stage III*	P1**
0	0.413	0.860
1	0.768	0.958
2	0.887	0.976
5	0.956	0.993
10	0.974	0.996
25	0.983	1.000

* Jan 1994–Jan 1996.

** Jan 1997–Jun 1998.

used to produce the HDP estimates also may contribute. Further research on multisensor estimation of light precipitation and the effects of changes in HDP algorithms would help to sort out these factors.

d. Summary

The comparisons of the Stage-III and P1 products presented in this section show that the range-dependent biases in HDP products are largely eliminated in NEXRAD multisensor products when the gauge density is high. For Stage-III products, however, HDP bias patterns still may be present, albeit strongly attenuated, with lower gauge densities. The two methodologies show considerable differences in precipitation occurrence and conditional means, due in large part to differences associated with light precipitation. The P1 product indicates precipitation more often than does Stage III, consistent with the gauge detection that drives the analysis. For Stage-III procedures, inconsistencies between the radar and the gauge estimates due to difficulties in observing light precipitation produce fewer indicated hours of precipitation. In addition, artifacts of radar biases and data-processing errors affect long-term accumulations for Stage III, producing visible radar circles and spokes and lines with erroneously high values. Although many of the data-processing errors have been largely eliminated in recent years, radar biases still produce faint radar spokes in long-term accumulations of P1 estimates.

Although this evaluation has identified some deficiencies in the NEXRAD multisensor products, it does not mean the products fail to produce quality precipitation estimates. For example, a gauge-only analysis for precipitation accumulations would not show any of the artifacts seen with the multisensor products but likely would have larger errors than do the multisensor estimates. By the same token, fewer artifacts in the P1 product does not imply higher accuracy in these estimates. Still, a motivation for implementing the P1 methodology at ABRFC was the obvious qualitative deficiencies in the Stage-III product. Another important consideration was operational time constraints. To establish an objective basis for assessing the two meth-

odologies, however, the uncertainties in precipitation estimates need to be evaluated. This issue is examined in the following section.

5. Uncertainty estimation

In section 4c, NEXRAD multisensor estimates were compared with gauge observations using scatterplots (see Fig. 7) and rmsd values (see Table 1). Because a gauge samples precipitation over a very small fraction of the 4 km × 4 km multisensor grid, much of the scatter (or rmsd) may be due to the natural spatial variability of precipitation. To use gauge precipitation data to evaluate the uncertainty of NEXRAD multisensor data, this spatial mismatch must be accounted for. Ciach and Krajewski (1999) describe a methodology, the error separation method (ESM), for separating the effects of spatial variability from the mean square difference between radar and gauge estimates. In the following section, an attempt to apply ESM to evaluate the error variance of NEXRAD multisensor estimates is discussed.

a. Error separation method

To apply ESM to NEXRAD multisensor and gauge observations, these two quantities first must be considered as estimates of areal average hourly precipitation over the multisensor grid. Then the error for each estimate is defined as its difference from the true average hourly rainfall. The objective of ESM is to derive the variance of the multisensor error from the variance (Var) of the multisensor and gauge difference. This objective is accomplished as follows:

$$\text{Var}(\hat{R}_m - \hat{R}_g) = \tag{1}$$

$$\text{Var}[(\hat{R}_m - R) - (\hat{R}_g - R)] = \tag{2}$$

$$\begin{aligned} &\text{Var}[(\hat{R}_m - R)] + \text{Var}[(\hat{R}_g - R)] \\ &- 2 \text{Cov}[(\hat{R}_m - R), (\hat{R}_g - R)], \tag{3} \end{aligned}$$

where \hat{R}_m and \hat{R}_g are the multisensor and gauge estimates of hourly precipitation, R is the true rainfall, and $\text{Cov}(\cdot, \cdot)$ is the covariance between multisensor and gauge errors. If this error covariance term can be neglected, based on independence of the multisensor and gauge-based errors, the variance of the hourly multisensor error can be calculated as

$$\text{Var}[(\hat{R}_m - R)] = \text{Var}(\hat{R}_m - \hat{R}_g) - \text{Var}[(\hat{R}_g - R)]. \tag{4}$$

In Eq. (4), the multisensor error variance is computed as the difference between the gauge–multisensor variance and the gauge error variance. The gauge–multisensor variance is estimated directly from the recorded data. The gauge error variance accounts for the gauge error in estimating the areal average precipitation over the 4 km × 4 km multisensor cell. After Ciach and Krajewski (1999) or Bras and Rodríguez-Iturbe (1976),

the variance of the difference between point and areal average precipitation is

$$\begin{aligned} \text{Var}(R_g - R) = \sigma_g^2 &\left[1 - \frac{2}{A} \int_A \rho(x_g, x) dx \right. \\ &\left. + \frac{1}{A^2} \int_A \int_A \rho(x_0, x_1) dx_0 dx_1 \right], \tag{5} \end{aligned}$$

where $\rho(\cdot, \cdot)$ denotes the spatial correlation function of the precipitation process; A is the averaging domain (i.e., a multisensor cell of about 4 km × 4 km); x_g is the location of the gauge within the radar pixel; x , x_0 , and x_1 are coordinate pairs within the domain A ; and σ_g^2 is the variance of the point rainfall and is assumed to be constant over A .

The applicability of Eq. (4) hinges on the assumption of independent errors. Ciach and Krajewski (1999) present a plausible justification for this assumption in the case of evaluating the uncertainty in radar-based precipitation estimates. The validation of multisensor data is more complex, however. Independence clearly will be violated if the gauges used to validate the multisensor product also are used in NEXRAD multisensor precipitation processing. Hence, a set of gauge precipitation estimates that is not used in NEXRAD processing must be used for validation. In the southern plains, one is limited to one such dataset: the ARS Micronet archive for the SGP97 campaign. As a result, a comprehensive comparison of the error characteristics for the Stage-III and P1 products cannot be performed. Instead, a sample evaluation using the available micronet data for June–September 1997 is presented here. With this example, a framework for determining the error characteristics of NEXRAD multisensor products is illustrated, and the data required for such an analysis are discussed. Yet, even with independent gauges for validation, the assumption of independent error still may be violated because neighboring gauges used in the multisensor product are correlated with the validation gauges. This result does not invalidate ESM but would mean that the covariance term is nonzero. Because it is not possible to evaluate the covariance of the gauge and multisensor errors with the available data, and because the effect of the covariance terms would be relatively small, we will proceed under the assumption that this covariance is negligible.

b. Spatial correlation of precipitation

The evaluation of Eq. (5) requires knowledge of the spatial correlation function of the precipitation process. Because this equation is applied only over a 4 km × 4 km grid cell for the hourly multisensor data, it is reasonable to assume that the intrinsic hypothesis, that is, that the variance is independent of location, and rainfall is a stationary process in time, is satisfied (Journel and Huijbregts 1978). The correlation function for this mod-

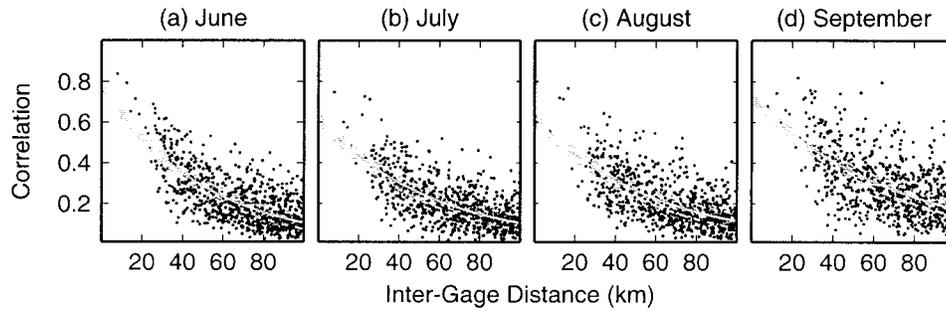


FIG. 8. Monthly correlograms of hourly precipitation for Jun through Sep for the Oklahoma Mesonet (based on data from Jan 1994 to Jun 1998). Each point represents the sample correlation for a gauge pair separated by an intergauge distance. The solid gray line displays a least squares fit of an exponential correlation model, and the dashed lines present the 95% confidence limits based on classic sampling theory (Stuart and Ord 1994).

el of the precipitation process will be estimated using the Oklahoma Mesonet and ARS Micronet gauge data.

Before this estimation could proceed, the existence of significant diurnal or seasonal variations in the precipitation correlation structure over the period of the available micronet data (June–September 1997) was investigated. Figure 8 presents monthly correlograms computed using the 4.5-yr record of the mesonet. In this figure, each point marks a sample correlation coefficient between time series of hourly precipitation observed by two gauges separated by a given distance d . The correlation plots are similar for each month, although September has slightly higher sample correlations than do the other months. Given the considerable scatter, however, the correlation function does not vary significantly over the four summer months. A similar analysis indicated that, although there are significant diurnal var-

iations in mean precipitation over the summer months, the correlation function does not exhibit significant diurnal fluctuations. Therefore, the correlation structure will be modeled using a single correlation function model for the entire 4-month period.

One obvious feature of the plots in Fig. 8 is the considerable scatter despite the relatively large sample size used in the calculations (around 3000 h of data for each point). Several possible sources for this scatter were investigated, including sampling variability and anisotropy. Classic sampling theory unfortunately is inappropriate for gauge correlations at the hourly timescale and cannot help to explain the large scatter observed in the sample correlation coefficients. This fact is illustrated by the dashed, gray lines in Fig. 8, which display the 95% confidence interval around the least squares fit of an exponential function to the data for a Gaussian process (Stuart and Ord 1994). The issue of anisotropy was investigated by computing directional correlograms for the mesonet gauge data. Anisotropy was evident in these graphs (not shown) but did not significantly reduce the scatter. Omnidirectional correlograms are used for the analysis in this section because the effects of anisotropy on ESM should be negligible over the scale of a single grid cell.

Another important observation can be made about the graphs in Fig. 8. The mesonet does not provide sufficient data to guide a model of the correlation function below the 20–30-km separation distance. To illustrate this fact, consider the mesonet and micronet data for June–September 1997. Figure 9 presents the gauge correlograms for these two networks. A least squares, two-parameter correlation model of the type

$$\rho(d) = r_0 \exp(-d/d_0) \tag{6}$$

was fit to the data from each network and from the two networks combined. Table 3 presents the correlation model parameters and the corresponding “fitting error,” the root-mean-square error (rmse) between the model and the data points. The model fit to the mesonet data

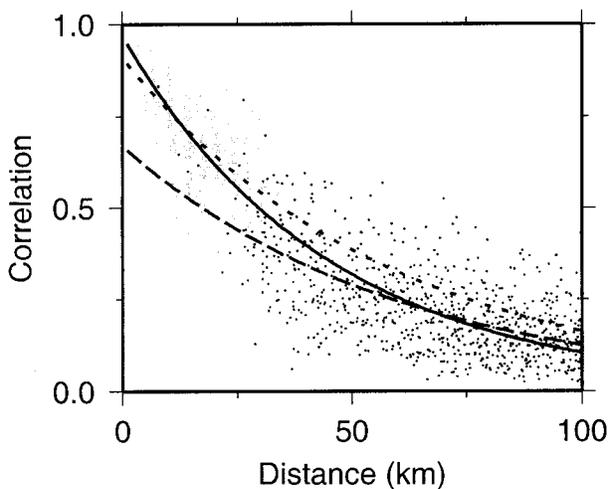


FIG. 9. Correlograms of hourly precipitation for the mesonet (black points) and micronet (gray points) gauges (Jun–Sep 1997). The three curves present fitted exponential correlation models to the mesonet data (long dashed), the micronet data (short dashed), and a combination of the two (solid).

TABLE 3. Summary of correlation model parameters and model rmse.

Gauge network	r_0	d_0 (km)	Rmse (mm)
Micronet	0.93	69.3	0.093
Mesonet	0.67	59.5	0.099
Both	0.99	46.9	0.107

results in an r_0 of 0.67, with an rmse of 0.10, while the micronet data result in an r_0 of 0.93. The results of the combined data are similar to those for the micronet alone, demonstrating that high-density gauge clusters are important for defining the correlation structure of rainfall at small scales.

c. Sensitivity analysis

To demonstrate how uncertainty in the small-scale correlation structure affects the error quantification, a sensitivity analysis of the ESM of Ciach and Krajewski (1999) was performed. First, the identifiability of the correlation parameter r_0 for the case of the mesonet data presented in Fig. 9 was investigated. The solid, black line in Fig. 10a shows the regression rmse for a least squares fit of Eq. (6) for multiple, fixed values of r_0 . It is apparent from the figure that the optimal location of r_0 is not well defined. Next, the effect of this poor parameter identifiability on the calculation of multisensor error variance was investigated. ESM was applied for each micronet gauge using the best-fit gauge correlation functions for multiple values of r_0 . The thin, gray lines in Fig. 10a show the results for each micronet gauge–cell pair. As r_0 decreases, $\text{Var}[(\hat{R}_g - R)]$ [see Eq. (5)] increases, thus reducing the estimated multisensor error [see Eq. (4)]. Over the wide range of r_0 highlighted in gray, the estimated multisensor error ranges from 0% to nearly 100% of the multisensor–gauge rmsd. Hence, it is not possible to quantify the multisensor error be-

cause of the uncertainty in small-scale correlation as estimated using only the mesonet gauges.

Figure 10b repeats this sensitivity analysis using both micronet and mesonet data to define the correlation function. Again, the solid, black line shows the identifiability of r_0 . Here, the range of r_0 is constrained through the addition of the micronet data, although the multisensor error still ranges from 50% to nearly 100% of the total rmsd. The bulk of the gauge–pixel pairs indicates that the multisensor error is 80% to 95% of the total rmsd. Hence, reducing the uncertainty in the small-scale correlation structure allows one to make approximate estimates of the multisensor error. Unfortunately, in this study, the micronet record available for estimating the multisensor error is only 4 months long. In practice, a longer independent gauge record would be needed to make meaningful quantitative estimates of the multisensor error magnitude.

From the analysis presented in this section, it becomes clear that, to estimate effectively the multisensor precipitation error, one should know the correlation function over the size of the radar pixel domain. This result implies that there is a need to establish small-scale clusters of precipitation gauges to provide the relevant information. This concept was put forward by Ciach and Krajewski (1999) and Steiner et al. (1999). Recently, we have deployed two such clusters, one in Iowa City, Iowa (Krajewski et al. 1998), and the other within the ARS Micronet in Oklahoma.

6. Discussion

This paper presents an attempt to evaluate NEXRAD multisensor precipitation products from ABRFC. Of course, the comparison of NEXRAD products with an independent source of precipitation data is needed for error estimation. Ironically, the ABRFC region has one of the highest gauge densities in the United States, but, because most available data are used in the multisensor

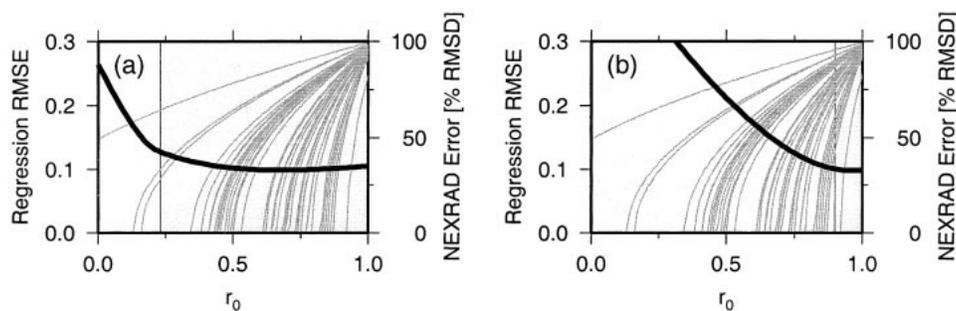


FIG. 10. Sensitivity analysis for ESM using (a) mesonet gauges and (b) mesonet and micronet gauges to define the correlation structure. In each plot, the solid black curve shows the sensitivity of the rmse of the regression model to the correlation parameter r_0 . The thin gray lines present (for each micronet gauge) the percentage of gauge–multisensor rmsd that is attributable to NEXRAD multisensor error. The rmse of the regression model is insensitive to the r_0 in the shaded areas. Estimates of the NEXRAD multisensor error are very sensitive to the estimate of r_0 over this range in both plots.

product, there are not many gauges available for validation.

Evaluating the uncertainties of forecasts and predictions is critical for the use of hydrometeorological products in operational and nonoperational applications. Within the meteorological forecast community, there is a long tradition of evaluating forecast uncertainty in terms of "skill scores," making it possible to quantify forecast improvements as technologies and methodologies have advanced. Even though the verification of hydrologic predictions usually is much more complex, similar efforts clearly are needed for NEXRAD precipitation products 1) to provide an objective basis for comparing the performance of competing methodologies, and 2) to quantify the uncertainty in precipitation estimates to be used as input in streamflow forecasting or other analyses. The following are some issues that need to be addressed to permit routine evaluation and verification of NEXRAD precipitation products.

Perhaps the most critical need is for development of archives of NEXRAD products and gauge data to facilitate routine evaluation, algorithm testing, and reanalysis. At a minimum, one would need an accessible archive of HDP products and gauge information. These archives could be used in cross-validation studies to estimate errors for different methodologies (e.g., Stage-III, P1, and gauge-only products). This archive also would permit the reanalysis of data to produce products based on a consistent methodology over a long period. This aspect is important for operational hydrologic forecasting because model parameters and forecast performance depend on the data sources for model calibration and prediction (Bradley 1997). Without a long, consistent record of precipitation products for recalibration of operational streamflow forecasting models, replacing current gauge products with improved precipitation products may not lead to improved forecasts (Bradley and Kruger 1998). The ability to carry out reanalyses also is important for hydrological and climatological studies to remove artifacts caused by changes in operational methodologies.

Still, there is a need to archive more fully the products from current operational analyses to aid in additional evaluations. This archive would include information on the gauge data actually used in processing (which can change from hour to hour with the availability of gauge reports), the resulting objective multisensor analyses (without human involvement), and the final products. For operational reasons, human involvement will continue to be necessary because it provides important quality control and assimilates other available information and knowledge that is not directly included in the processing algorithms. Comparisons with objective methods are needed to quantify improvements resulting from HAS-forecaster interaction, however.

Verification also will require accurate representation of the spatial correlation structure of precipitation within the RFC region. For the evaluation of hourly

precipitation estimates on the scale of an HRAP cell ($4 \text{ km} \times 4 \text{ km}$), the error estimates depend on the small-scale correlation structure. Although the Oklahoma Mesonet has a higher gauge density than most locations, the average gauge spacing ($\sim 35 \text{ km}$) cannot provide information on the correlation at small scales ($< 4 \text{ km}$). This study was aided by the limited data available from the ARS Micronet, for which the average gauge spacing is much smaller ($\sim 5 \text{ km}$). These data helped in modeling the small-scale spatial correlation structure and greatly reduced the uncertainty about the effect of spatial precipitation variability on the difference between gauge and multisensor estimates. There is still a need to install gauges on the scale of the micronet and smaller at locations across the country to develop information on correlation structure for verification. This information also would contribute to the design of algorithms for optimal merging of gauge and radar data (which depends on the uncertainty associated with the gauge estimate).

Further research also is needed to refine verification schemes for multisensor precipitation products. For example, in this application of the ESM approach, we assumed that the errors in the gauge and multisensor estimates of areal precipitation are statistically independent. Although Ciach and Krajewski (1999) argue that the independence assumption may be reasonable for uncertainty estimation for radar-only precipitation products, the assumption may be violated for ABRFC multisensor products because of the high density of gauges used to generate these estimates. The gauges not included in the multisensor analysis, which are used for verification, are correlated with nearby gauges used to produce the multisensor estimates. Further research is needed to evaluate the effect of this correlation on the error covariance term neglected in this application of the ESM.

Last, other approaches to uncertainty analysis could provide the best means for estimating prediction errors in space and time across a region. For example, with Stage III, one could begin by evaluating the HDP products and gauge-only estimates separately, then could propagate the uncertainties through the Stage-III processing to derive uncertainties for the multisensor fields. This approach has some theoretical appeal and could provide a tool for evaluating the error characteristics of various gauge and radar combinations for design of improved multisensor estimation methodologies.

7. Summary and conclusions

NEXRAD multisensor precipitation products from ABRFC for a 5.5-yr period were examined. During this period, two processing methods, Stage III and P1, were used to make multisensor precipitation estimates. Both methods used a radar-based precipitation product that has known systematic biases (Smith et al. 1996). An intercomparison of the products indicates that radar bi-

ases are reduced, although still present, in NEXRAD multisensor precipitation estimates. The Stage-III product, used predominately in the first half of the study period, shows more range-dependent biases than do the P1 estimates. The evaluation also revealed that additional processing artifacts are present when precipitation information from multiple radars is merged to produce the multisensor product for the ABRFC region. Despite these deficiencies and differences in quality, we cannot infer that one method is superior to the other. An evaluation of the error characteristics for each product is necessary to compare competing algorithms for multisensor precipitation estimation.

In section 5, the error separation method (Ciach and Krajewski 1999) was applied to try to evaluate the error variance of NEXRAD multisensor estimates. Several obstacles hampered this attempt. First, despite the high gauge density of the ABRFC region, data for gauges that are not included in the multisensor products are difficult to find for validation. Hence, the validation necessarily was limited to a short observation period for a small region within ABRFC and was based on gauge data from the ARS Micronet. Second, existing gauge networks do not provide sufficient data for characterizing small-scale (<4 km) spatial variability of precipitation. Detail at this scale is necessary to represent accurately the gauge correlation function. Third, the evolving nature of operational precipitation estimation results in a nonhomogeneous record. This fact makes evaluation of algorithms difficult and creates problems in the use of these data in applications.

To implement a validation framework for routine evaluation of NEXRAD multisensor products, several issues first must be addressed. An archive of gauge and radar data used in precipitation processing, in addition to the multisensor products, is needed. This archive would help in evaluating operational products and enable reanalysis of precipitation estimates using alternate algorithms. For operational multisensor precipitation estimates produced at RFCs, the archive should include a record of the processing methodology and human interactions that were used to produce the final product. Also, to gather the information on small-scale precipitation variability needed for validation, small-scale gauge networks must be established at several sites across the country. The development of such a validation framework is clearly vital for quantification of error in current estimation methodologies and for evaluation of proposed algorithms.

Acknowledgments. The work was carried out at the Iowa Institute of Hydraulic Research Computational Laboratory for Hydrometeorology and Water Resources. We are grateful to D.-J. Seo, Mike Smith, and Richard Fulton at the NWS Hydrologic Research Laboratory for their help on the operational use of NEXRAD in multisensor precipitation estimation and to Bill Lawrence and Suzanne Fortin at ABRFC for providing details on

NEXRAD precipitation processing and operational experiences at ABRFC. We also thank Grzegorz Ciach for his contributions to our efforts in evaluation of NEXRAD products and the three anonymous reviewers for their insightful comments and suggestions.

REFERENCES

- Anagnostou, E. N., and W. F. Krajewski, 1998: Calibration of the WSR-88D precipitation processing system. *Wea. Forecasting*, **13**, 396–406.
- Bradley, A. A., 1997: Use of weather radar with lumped parameter hydrologic models. Preprints, *13th Conf. on Hydrology*, Long Beach, CA, Amer. Meteor. Soc., J87–J88.
- , and A. Kruger, 1998: Recalibration of hydrologic models for use with WSR-88D precipitation estimates. Preprints, *Special Symp. on Hydrology*, Phoenix, AZ, Amer. Meteor. Soc., 302–305.
- Bras, R. L., and I. Rodríguez-Iturbe, 1976: Evaluation of mean square error involved in approximating the areal average of a rainfall event by a discrete summation. *Water Resour. Res.*, **12**, 181–184.
- Brock, F. V., K. C. Crawford, R. L. Elliott, G. W. Cuperus, S. J. Stadler, H. L. Johnson, and M. D. Eilts, 1995: The Oklahoma Mesonet—A technical overview. *J. Atmos. Oceanic Technol.*, **12**, 5–19.
- Ciach, J. G., and W. F. Krajewski, 1999: On the estimation of radar rainfall error variance. *Adv. Water Resour.*, **22**, 585–595.
- Crum, T. D., R. L. Alberty, and D. W. Burgess, 1993: Recording, archiving, and using WSR-88D data. *Bull. Amer. Meteor. Soc.*, **74**, 645–653.
- Fulton, R. A., J. P. Breidenbach, D.-J. Seo, and D. A. Miller, 1998: WSR-88D rainfall algorithm. *Wea. Forecasting*, **13**, 377–395.
- Habib, E., W. F. Krajewski, V. Nešpor, and A. Kruger, 1999: Numerical simulation studies of rain gauge data correction due to wind effect. *J. Geophys. Res.*, **104**, 19 723–19 734.
- , —, and A. Kruger, 2000: Sampling errors of fine resolution tipping-bucket rain gauge measurements. *J. Hydrol. Eng.*, in press.
- Hudlow, M. D., 1998: Technological development in real-time operational hydrologic forecasting in the United States. *J. Hydrol.*, **102**, 69–92.
- Johnson, D., M. Smith, V. Koren, and B. Finnerty, 1999: Comparing mean areal precipitation estimates from NEXRAD and rain gauge networks. *J. Hydrol. Eng.*, **4**, 117–124.
- Journal, A. G., and Ch. J. Huijbregts, 1978: *Mining Geostatistics*. Academic Press, 600 pp.
- 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.
- Krajewski, W. F., 1987: Cokriging of radar rainfall and rain-gauge data. *J. Geophys. Res.*, **92**, 9571–9580.
- , A. Kruger, and V. Nešpor, 1998: Experimental and numerical studies of small-scale rainfall measurements and variability. *Water Sci. Technol.*, **37**, 131–138.
- Kruger, A., and W. F. Krajewski, 1997: Efficient storage of weather radar data. *Software Prac. Exp.*, **27**, 623–635.
- NWS Office of Hydrology, 1997a: Stage II Precipitation Processing System. User's guide. Hydrologic Research Laboratory, 22 pp. [Available from Hydrologic Research Laboratory, National Weather Service, 1325 East–West Highway, Silver Spring, MD 20910; also available online at <http://hsp.nws.noaa.gov/oh/hrl/general/indexdoc.htm>.]
- , 1997b: Stage III Precipitation Processing System. System guide. Hydrologic Research Laboratory, 12 pp. [Available from Hydrologic Research Laboratory, National Weather Service, 1325 East–West Highway, Silver Spring, MD 20910; also available online at <http://hsp.nws.noaa.gov/oh/hrl/general/indexdoc.htm>.]

- Reed, S. M., and D. R. Maidment, 1999: Coordinate transformations for using NEXRAD data in GIS-based hydrologic modeling. *J. Hydrol. Eng.*, **4**, 174–183.
- Seo, D.-J., 1998: Real-time estimation of rainfall fields using radar rainfall and rain gage data. *J. Hydrol.*, **208**, 37–52.
- Smith, J. A., and W. F. Krajewski, 1991: Estimation of the mean field bias of radar rainfall estimates. *J. Appl. Meteor.*, **30**, 397–412.
- , D. J. Seo, M. L. Baeck, and M. D. Hudlow, 1996: An inter-comparison study of NEXRAD precipitation estimates. *Water Resour. Res.*, **32**, 2035–2045.
- Steiner, M., J. A. Smith, S. J. Burges, C. V. Alonso, and R. W. Darden, 1999: Effect of bias adjustment and rain gage data quality control on radar rainfall estimation. *Water Resour. Res.*, **35**, 2487–2503.
- Stuart, A., and J. K. Ord, 1994: *Kendall's Advanced Theory of Statistics*. Vol. 1, *Distribution Theory*, 6th ed., Edward Arnold, 676 pp.
- Young, C. B., B. R. Nelson, A. A. Bradley, J. A. Smith, C. D. Peters-Lidard, A. Kruger, and M. L. Baeck, 1999: An evaluation of NEXRAD precipitation estimates in complex terrain. *J. Geophys. Res.*, **104**, 19 691–19 703.