Evolving Multisensor Precipitation Estimation Methods: Their Impacts on Flow Prediction Using a Distributed Hydrologic Model

DAVID KITZMILLER,* SUZANNE VAN COOTEN,[+] FENG DING,* KENNETH HOWARD,[+]
CARRIE LANGSTON,[#] JIAN ZHANG,[+] HEATHER MOSER,[#] YU ZHANG,* JONATHAN J. GOURLEY,[+]
DONGSOO KIM,[@] AND DAVID RILEY*

*Office of Hydrologic Development, NOAA/National Weather Service, Silver Spring, Maryland
[+]NOAA/National Severe Storms Laboratory, Office of Oceanic and Atmospheric Research, Norman, Oklahoma
[#]Cooperative Institute for Mesoscale Meteorological Studies, Norman, Oklahoma
[@]National Climatic Data Center, NOAA/National Environmental Satellite, Data, and Information Service,
Camp Springs, Maryland

ABSTRACT

This study investigates evolving methodologies for radar and merged gauge–radar quantitative precipitation estimation (QPE) to determine their influence on the flow predictions of a distributed hydrologic model. These methods include the National Mosaic and QPE algorithm package (NMQ), under development at the National Severe Storms Laboratory (NSSL), and the Multisensor Precipitation Estimator (MPE) and High-Resolution Precipitation Estimator (HPE) suites currently operational at National Weather Service (NWS) field offices. The goal of the study is to determine which combination of algorithm features offers the greatest benefit toward operational hydrologic forecasting. These features include automated radar quality control, automated Z–R selection, brightband identification, bias correction, multiple radar data compositing, and gauge–radar merging, which all differ between NMQ and MPE–HPE. To examine the spatial and temporal characteristics of the precipitation fields produced by each of the QPE methodologies, high-resolution (4 km and hourly) gridded precipitation estimates were derived by each algorithm suite for three major precipitation events between 2003 and 2006 over subcatchments within the Tar–Pamlico River basin of North Carolina. The results indicate that the NMQ radar-only algorithm suite consistently yielded closer agreement with reference rain gauge reports than the corresponding HPE radar-only estimates did. Similarly, the NMQ radar-only QPE input generally yielded hydrologic simulations that were closer to observations at multiple stream gauging points. These findings indicate that the combination of Z–R selection and freezing-level identification algorithms within NMQ, but not incorporated within MPE and HPE, would have an appreciable positive impact on hydrologic simulations. There were relatively small differences between NMQ and HPE gauge–radar estimates in terms of accuracy and impacts on hydrologic simulations, most likely due to the large influence of the input rain gauge information.

1. Introduction

Improving the accuracy of both quantitative precipitation estimation (QPE) and high-resolution distributed hydrologic models are critical to the National Oceanic and Atmospheric Administration's (NOAA) mission. The experiments described herein provide a foundation for NOAA hydrometeorological service improvements for the Tar–Pamlico River basin of North Carolina and

eventually much of the United States. These service improvements will provide additional benefits to NOAA programs in the Carolinas focusing on ecosystem and water resource management, severe storm hazards, and estuary health.

This project is a joint scientific research effort conducted by the National Severe Storms Laboratory (NSSL); the National Weather Service (NWS) Office of Hydrologic Development (OHD); and the National Environmental Satellite, Data, and Information Service (NESDIS). These organizations are working to identify an optimum set of techniques and algorithms to serve as a state-of-the-science NOAA multisensor QPE. A key component of this collaborative research is the scientific validation

Corresponding author address: David Kitzmiller, Office of Hydrologic Development, NOAA/National Weather Service, 1325 East West Highway, Silver Spring, MD 20910.
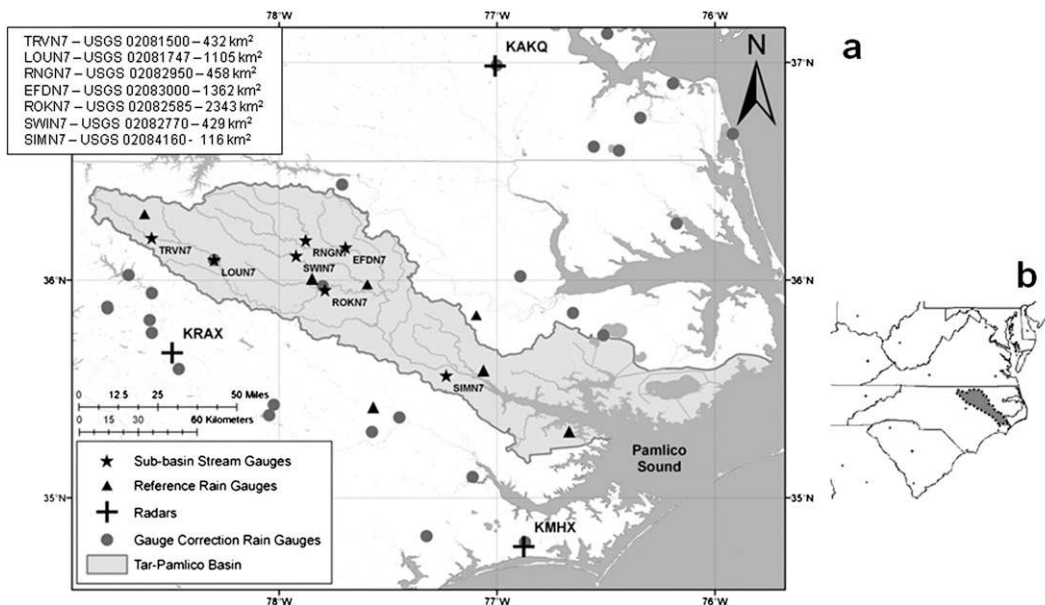E-mail: david.kitzmiller@noaa.gov

FIG. 1. (a) Rain gauge, stream gauge, and radar locations within and near the Tar–Pamlico basin, North Carolina, basin (shaded gray). Stream gauge sites used in this study are shown as labeled stars, input rain gauge locations are shown as circles, and reference rain gauges are shown as triangles. Input rain gauge locations are those employed in this study during the June 2006 event. WSR-88D sites are indicated by crosses. (b) The location of the basin on the United States east coast region, with WSR-88D radar sites indicated by dots.

of the techniques for use in NOAA forecast and warning operations.

The evaluation consists of three phases: first, evaluation of precipitation algorithms in post-case analysis in terms of accuracy relative to a set of reference rain gauges; second, identification of QPE algorithms and individual elements that provide substantial improvements in accuracy over current operational baseline QPE products; and third, evaluation of the impact of improved QPE on the quality of streamflow simulations produced by an advanced distributed hydrologic model.

The Tar–Pamlico River basin in North Carolina has been identified as a test bed region for several reasons. The basin and surrounding areas feature radar and rain gauge networks that are similar to those in many hydrologically sensitive areas of the United States (see Fig. 1). Furthermore, interdisciplinary, multiagency research efforts focused on improving coupled hydrologic, hydraulic, and water-quality models for both rivers and estuaries are ongoing in the basin and Pamlico Sound. These activities include the Coastal and Inland Flood Observation and Warning (CI-FLOW) project, which seek to leverage the outcomes of this multisensor QPE research effort to improve river and flash-flood forecasts for the Tar–Pamlico basin. CI-FLOW focuses on a number of problems related to precipitation–environment interactions including flooding, debris flow prediction, river–estuary interaction modeling, and water-quality prediction.

This article describes the evolution of this QPE research activity to date and presents initial results in terms of accuracy of radar-only precipitation estimation techniques relative to reference rain gauge reports and the sensitivity of streamflow predictions in headwater basins to these different precipitation inputs. This paper is structured so that sections 2–4 provide information on the experimental models, methods, and input. Section 2 describes the radar and gauge–radar estimation algorithm suites; section 3 describes the experimental methodology for rain gauge and hydrologic simulation evaluations, with a brief description of the hydrologic model framework; and section 4 describes precipitation, temperature, and stream discharge inputs.

Sections 5–9 present results of the research experiments with a summary of the research and conclusions presented in section 10. Sections 5 and 6 detail the radar and multisensor precipitation fields for three storm events and rain gauge verification results, respectively. Section 7 describes the hydrologic simulations and their overall accuracy relative to observations, and section 8 gives details for subbasins in the individual storm events. Further details of the impact of differences in QPE on the hydrologic simulations are provided in section 9.

2. Precipitation estimation algorithm packages

The subsections here present a brief description of existing NOAA QPE algorithm packages, their output,

and their capabilities. Each NOAA QPE system continues to evolve in response to user needs, which accounts for their different approaches and unique features as each system attempts to produce the most accurate QPE possible.

a. NMQ

This system (J. Zhang et al. 2011, 2006, 2004; Vasiloff et al. 2007) developed from a joint initiative between NSSL; the Salt River Project (SRP); the Federal Aviation Administration (FAA) Aviation Weather Research Program; and the NWS Office of Climate, Water, and Weather Services. The objective of the National Mosaic and QPE system (NMQ) research and development, which meets the objectives of NOAA's weather and water mission, was twofold. The first goal was to develop a seamless high-resolution national 3D grid of radar reflectivity for operational utilization in data assimilation, numerical weather prediction (NWP) model verification, and aviation product development. The second goal was to develop fully automated multisensor QPE techniques at high spatial and temporal resolutions and accuracy for use in operational flash-flood monitoring and prediction and water resource management.

The NMQ system is a collection of techniques and algorithms that facilitate the integration of multiple radar and gauge networks, including the Weather Surveillance Radar-1988 Doppler (WSR-88D), the Terminal Doppler Weather Radar, and Canadian radar networks, into a unified 3D grid. A suite of QPE products is produced at ~1-km resolution and is updated every 5 min. Using a combination of vertical reflectivity profiles and Rapid Update Cycle (RUC) model analyses, the NMQ system identifies whether the precipitation is convective, stratiform, or tropical for each grid cell and assigns appropriate Z–R relationships every 5 min to obtain a radar-based QPE product suite (Xu et al. 2008; Zhang et al. 2008). An alternative Z–R relationship is applied in conditions when the RUC analysis indicates snow at the surface (i.e., when the temperature is below $2°C$ and the wet bulb temperature is below $0°C$). Hail detection and appropriate adjustment of Z–R relationships are also incorporated.

NMQ produces a set of gauge–radar QPE products by applying a local rain gauge bias adjustment to the radar-based QPE, on an hourly basis. Real-time QPE products for the conterminous United States (CONUS) have been available to researchers since 2007. The system is scalable and could be configured for national implementation at an NWS national office such as the National Centers for Environmental Prediction (NCEP) as well as at regional or local offices.

NMQ is a complete end-to-end system and operates independently of current NWS field-office baseline hardware and software. A real-time prototype is currently functional, and, although it is not operational, it is being evaluated for transition to operations.

b. MPE and HPE

The High-Resolution Precipitation Estimator (HPE; Kitzmiller et al. 2008) functions operationally within the Advanced Weather Interactive Processing System (AWIPS). Multisensor Precipitation Estimator (MPE) and HPE use the AWIPS environment to integrate rain gauge, radar, and Geostationary Operational Environmental Satellite (GOES) precipitation estimates (Scofield and Kuligowski 2003) into fields covering the area of responsibility for individual Weather Forecast Offices (WFOs) and River Forecast Centers (RFCs). MPE, designed primarily for river prediction applications, includes a large suite of interactive tools for quality control (QC) of all inputs and resulting output products. All MPE rainfall estimates are spatially averaged to a 4-km grid and updated hourly.

HPE was designed primarily for flash-flood monitoring and creates 1-km grids of precipitation rate and accumulation on a subhourly update cycle. All radar rainfall accumulation and rain-rate products used by HPE are derived from the WSR-88D Precipitation Processing System (Fulton et al. 1998). For this research collaboration, HPE radar-only and gauge–radar estimates were evaluated. It should be noted that MPE and HPE presently ingest only precipitation estimates and do not have a capability for direct estimation of precipitation from raw remote sensor input (e.g., radar reflectivity, satellite radiance data) in terms of precipitation rate.

c. Operational gauge–radar precipitation analyses from the SERFC

These analyses are routinely produced, gridded gauge–radar QPE analyses using the Hydrologic Rainfall Analysis Project (HRAP) rectangular grid (Greene and Hudlow 1982; Reed and Maidment 1999). The HRAP grid is defined at the 4-km by 4-km resolution that corresponds directly to the NWS Next Generation Weather Radar (NEXRAD) precipitation products. HRAP gridded QPE data are produced operationally at the Southeast River Forecast Center (SERFC) and other NWS River Forecast Centers, through application of MPE or similar analysis packages. Analysts apply extensive quality control to gauge and radar input and often blend different input fields to produce a final estimated QPE grid. For this study, the research team compiled these analyses from internal OHD archives and from Stage IV mosaic composites created by the National Precipitation Verification Unit (Lin and Mitchell 2005). This QPE source is referred to as SERFC.

3. Project outline and methodology

The project steps can be summarized as follows:

1) identification of suitable historical heavy precipitation events in both cold and warm seasons;
2) creation of common radar, satellite, numerical weather prediction model, and rain gauge input datasets for all QPE algorithms and collection of stream gauge reports;
3) quality control of a common set of rain gauge reports, with some for input and some for validation;
4) execution of NMQ, and MPE–HPE algorithms to produce QPEs;
5) evaluation of precipitation algorithms relative to the reference rain gauge reports; and
6) evaluation of QPE in terms of impact on the quality of streamflow simulations from an advanced distributed hydrologic model.

a. Hydrometeorological events and input data

For this study, researchers selected three hydrometeorological events during the period from January 2003 to June 2006. The collaborators assembled the input and verification datasets required by all algorithms for each case. Each case features at least one major precipitation event over a period of at least 10 days. The precipitation events fell within the following periods:

- 1200 UTC 17 September–1200 UTC 20 September 2003 (Hurricane Isabel);
- 1200 UTC 9 December 2004–1200 UTC 17 January 2005 (three major precipitation events); and
- 0000 UTC 11 June–0000 UTC 16 June 2006 (major convective events, including Tropical Storm Alberto).

Event selection was contingent on availability of the radar and other data for the Tar–Pamlico basin and surrounding areas. Datasets included RUC model fields of surface temperature and melting level (1-h/20-km gridded fields), meteorological in situ data (precipitation and surface air temperature), and operational gridded MPE analyses from the SERFC. Level-2 data were collected for the WSR-88D sites of KRAX (NWS WFO Raleigh, North Carolina), KMHX (NWS WFO Morehead City, North Carolina), and KAKQ (NWS WFO Wakefield, Virginia), which are shown in Fig. 1a.

b. Hydrologic model

The Hydrology Laboratory-Research Distributed Hydrologic Model (HL-RDHM) served as the vehicle for testing the impact of the different QPEs on streamflow simulations. The HL-RDHM [formerly the Hydrology Laboratory Research Modeling System (HL-RMS);

Koren et al. 2004] consists of a framework integrating several components of streamflow modeling, including rainfall runoff [Sacramento Soil Moisture Accounting model (SAC-SMA); Burnash 1995; Koren et al. 2004], hill slope routing (Reed 2003), and snowmelt (a temperature index model commonly designated SNOW-17; Anderson 1976).

For these experiments, HL-RDHM was configured with the HRAP grid. A priori estimates of tunable parameters, such as those for the soil moisture and hill slope routing models, were used. These parameters were based on available soil type and land-use datasets (Y. Zhang et al. 2011). All simulations started with a zero base flow initial condition, though release of free water caused nonzero discharge prior to precipitation. Cell-to-cell connectivity for runoff water routing was based on evaluation of topography data from a 100-m digital elevation model. Evapotranspiration estimates were based on a climatic specification of the potential rate. Based on recent results from a survey administered by OHD, this methodology is used by most NWS RFCs in their forecasting operations.

It is important to note that, for this study, HL-RDHM, using a priori parameters only, was not calibrated to produce optimum discharge simulations for any one precipitation input source. However, we found that the three sets of model simulations based on combined gauge–radar input were all of comparable quality. Therefore, conclusions drawn from the series of assessments outlined in this paper are based on the impact of the different QPEs on the accuracy of the simulations in depicting hourly discharge time series and discharge and timing of flood peaks.

c. Scope of the study

Collecting and processing base radar data and then assessing the quality and accuracy of gauge datasets necessary for verification to determine a level of accuracy is both data and labor intensive. Thus, it was not practical for the scope of this study to produce long-term continuous time series of QPE grids from each of the algorithm suites. Rather, this study focused on creating analyses and verification datasets covering three active precipitation periods of ten to thirty days each and only for the hours when precipitation was observed over the basin. Because the initial assessment period was associated with Hurricane Isabel in September 2003, a basic HL-RDHM simulation was run from 1 January 2003 to 30 June 2006, using precipitation input from the SERFC operational 1-h datasets. The 8-month period beginning 1 January 2003 served as a "warm up" period for the hydrologic model. The SERFC datasets incorporate MPE gauge–radar analysis with forecasters' quality control

of input data and gridded output fields. For each of the evaluation–comparison events, the SERFC QPE grids were replaced with the experimental ones (i.e., NMQ or MPE–HPE). These simulations were then compared with hourly time series stream discharge observations in terms of linear correlation and Nash–Sutcliffe efficiency (Nash and Sutcliffe 1970). Separate evaluations of flood peak simulations were also made, in terms of the mean absolute error (MAE) and median absolute error of peak discharge and time of flood peaks. Finally, an assessment was made of the impact of QPE on the flood peak discharge per unit basin area, or specific discharge.

4. Precipitation and temperature input

The NMQ and HPE algorithm packages use a common set of radar and rain gauge inputs. In addition, the NMQ requires an externally supplied estimate of freezing level, used in Z–R corrections. The inputs and processing steps are described below.

a. Radar input and products

WSR-88D level-II reflectivity (1° azimuth × 1 km range resolution in local radar coordinates), for multiple elevation angles, from the NWS radar sites of KRAX, KMHX, and KAKQ served as the radar input. These data were used by the NMQ system to create 3D reflectivity grids at multiple levels, which were then used as input to additional QPE algorithms. The NMQ QPE grids are of approximately 1-km mesh spacing (0.01° latitude–longitude grid). Although NMQ produces a variety of multisensor products, this report includes only an analysis of the basic radar-only precipitation estimates. For MPE–HPE, the data were input to the Open Radar Product Generator (ORPG) version OB5.2, which generated digital storm-total precipitation (DSP) and digital precipitation array (DPA) products, which were input to an offline copy of MPE and HPE. Precipitation accumulations used for this study were based on time differencing of the DSP product.

For HPE, output was removed for the 2-h period ending at 2300 UTC 26 December, when a time accounting error caused a zero storm total precipitation amount to be indicated at the beginning of the hour. The correct storm total was still in place at the end of the hour; thus, the time difference scheme caused the erroneous value of storm total precipitation to be assigned to that hour. This resulted in some basin-average amounts of over 25 mm resulting in obviously unrealistic results. Because of the error not being physically reasonable, that hour's HPE grid was replaced with the MPE radar-only grid. The MPE product indicated no precipitation at that hour, as

did the other QPE sources. To date, this time accounting error has not been observed in operations.

b. Rain gauge reports for QPE input and reference evaluations

Hourly reports from rain gauges both inside and outside the basin boundaries were collected and input into the algorithm suites. Gauges well outside the boundaries of the basin have very little impact on the basin-average precipitation but must be included as the reports do influence bias corrections for the radar data.

Gauge sites in this study were primarily from four different networks: North Carolina Econet sites, which are maintained for environmental and other purposes; NWS Automated Surface Observing System (ASOS) sites; Cooperative Observer (COOP) sites whose reports were supplied by the National Climatic Data Center (NCDC); and real-time reporting sites operated by several federal and local authorities, commonly reporting through the GOES Data Communication System and collected and collated by the NWS Hydrologic Automated Data System (HADS). The distribution of rain gauges used in the analyses is shown in Fig. 1a.

Other rain gauge locations were used to provide validation reference observations for the three precipitation events (Fig. 1a). The observations from these gauges were not used in the gauge bias algorithms and were only used as validation points. For the 2003 event, three hourly rain gauge sites were selected to be reference gauges: Aurora (AURO; North Carolina Econet site), Oxford (OXFO; North Carolina Econet site), and the Lizzie (LZZN7) HADS site. For the 2006 event, reports from the rain gauge at the U.S. Geological Survey (USGS) Swift Creek at NC97 near Leggett, North Carolina (LEGN7), gauge; OXFO; and Williamston (WILL; North Carolina Econet site) served as references. For the 2004–05 event, the Tranters Creek (TRAN7) HADS site and KRWI (Rocky Mount-Wilson Airport ASOS site) were designated reference gauges. A set of 68 daily reporting sites, not collocated with hourly reporting ones, were available as a reference for 24-h precipitation amounts. These rain gauge data were collected from NCDC, North Carolina state government, and NWS sources.

c. Rain gauge report quality control

Although we attempted no modification or quality control of input radar data, the existence of certain systematic errors necessitated manual quality control of the input rain gauge data. The reports were compared manually with neighbors and with radar reflectivity and precipitation fields. All suspect reports were removed from the final analysis, consistent with current operational practice at RFCs.

Rain gauge input (for both multisensor analyses and reference validation sites) was quality controlled jointly by NSSL, OHD, and NCDC. A common set of input and validation reports was agreed upon and used in the verification statistics.

Of the available set of hourly gauges, only the ASOS units were equipped to report the water equivalent during frozen precipitation events. An examination of the hourly gauge time series indicated that some sites were affected by snow during the precipitation event of 14–16 December 2004, when several centimeters of snow accumulated over parts of the river basin. These gauges reported little or no precipitation during the latter part of the event. These reports were deleted from that portion of the overall record and were not used for input or validation. We found that the 14 daily reports also had questionable accuracy especially with zero values. The daily reports were carefully examined and suspect reports during two events were removed.

d. Temperature input

In addition to the freezing-level information required by NMQ, the HL-RDHM package requires gridded estimates of surface air temperature as input to its snowfall accumulation and melt model, SNOW-17. These were extracted from the long-term NCEP–National Center for Atmospheric Research (NCAR) reanalysis archive (Kalnay et al. 1996) and interpolated to the basin area. Although frozen precipitation and melting had little influence on the warm-season events, some snow was observed during the cool-season events, particularly around 27 December 2004.

e. Stream gauge reports

USGS hourly stream gauge reports for seven sites within the Tar–Pamlico basin were kindly provided by the North Carolina district office of the USGS. These sites are all forecast points within the NWS river forecast system. Time series from the EFDN7, LOUN7, RNGN7, TRVN7, SIMN7, SWIN7, and ROKN7 sites (Fig. 1) were used to evaluate HL-RDHM output. These basins range in size from 116 to 2343 km^2. All but one of the gauge sites reports discharge from unregulated headwaters. The ROKN7 site is immediately downstream of a reservoir and is subject to some regulation, which is not modeled by the HL-RDHM package. However, we have included the ROKN7 simulations in our results.

The hourly time series for all stream gauges included some missing data. For instances where the missing data sequences were less than 10 h and were in periods of slow recession, the discharges were estimated by interpolation. For LOUN7, the reported hydrograph appeared unrealistic during a period on 19–20 September 2003, following the passage of Hurricane Isabel. Alternative reports indicated the gauge was giving unreliable values (D. Kim 2010, personal communication), and these discharge data were excluded from our verification statistics.

5. Precipitation analyses and comparisons

The differing radar analyses yielded accumulation fields with both subtle and obvious differences over the course of all events. During all events, the spatial pattern of 24-h precipitation was very similar among the three estimation systems, as might be expected given the common sources of radar data. However, assumptions about Z–R relationships made by the logic within each of the algorithm suites caused large systematic differences in the magnitude of the estimated rainfall. Sample 24-h accumulations ending at 1200 UTC 19 September 2003 and 0000 UTC 27 December 2004 are shown in Figs. 2 and 3, respectively.

In these figures, the 4-km mesh HRAP projection is used, illustrating the granularity of the precipitation input to the distributed hydrologic model. The effective coverage of the various analyses (nongray areas) differs depending on the number of radar units used as input and on the radar QPE algorithms; the operational SERFC analyses (Figs. 2a and 3a) extend southward from central Virginia and northern West Virginia to Florida and the Gulf Coast, whereas the HPE analyses (Figs. 2c and 3c) used solely DPA data from the three radar units nearest the study basin. The NMQ analyses (Figs. 2b and 3b) used data from the same three radars, but they include radar estimates from beyond the 230-km range limit of the DPA product and thus encompass a slightly larger valid area than the HPE analyses.

An examination of all the days included in the study reveals differences among the various precipitation analyses were most pronounced during the September 2003 Hurricane Isabel event. As shown in Figs. 2a–c, the operational SERFC analysis contained significantly more rainfall than the NMQ or HPE radar-only analyses did. Over the Tar–Pamlico basin (white rectangle in Fig. 2a), these differences amounted to 25 mm over the 24-h period ending at 1200 UTC 19 September and 40 mm over the course of the entire storm event. Between the two radar-only analyses of NMQ and HPE (Figs. 2b,c), rainfall magnitude differences were most apparent over the extreme northern portion of the basin and south-central Virginia. The NMQ and HPE multisensor algorithms also yielded substantially different accumulations over the extreme northern portion of the basin and southern Virginia, with the HPE analysis (Fig. 2e) indicating a much larger area of rainfall exceeding 150 mm than the SERFC or NMQ analyses had. Underestimation by the radar algorithms was also evident during the June 2006
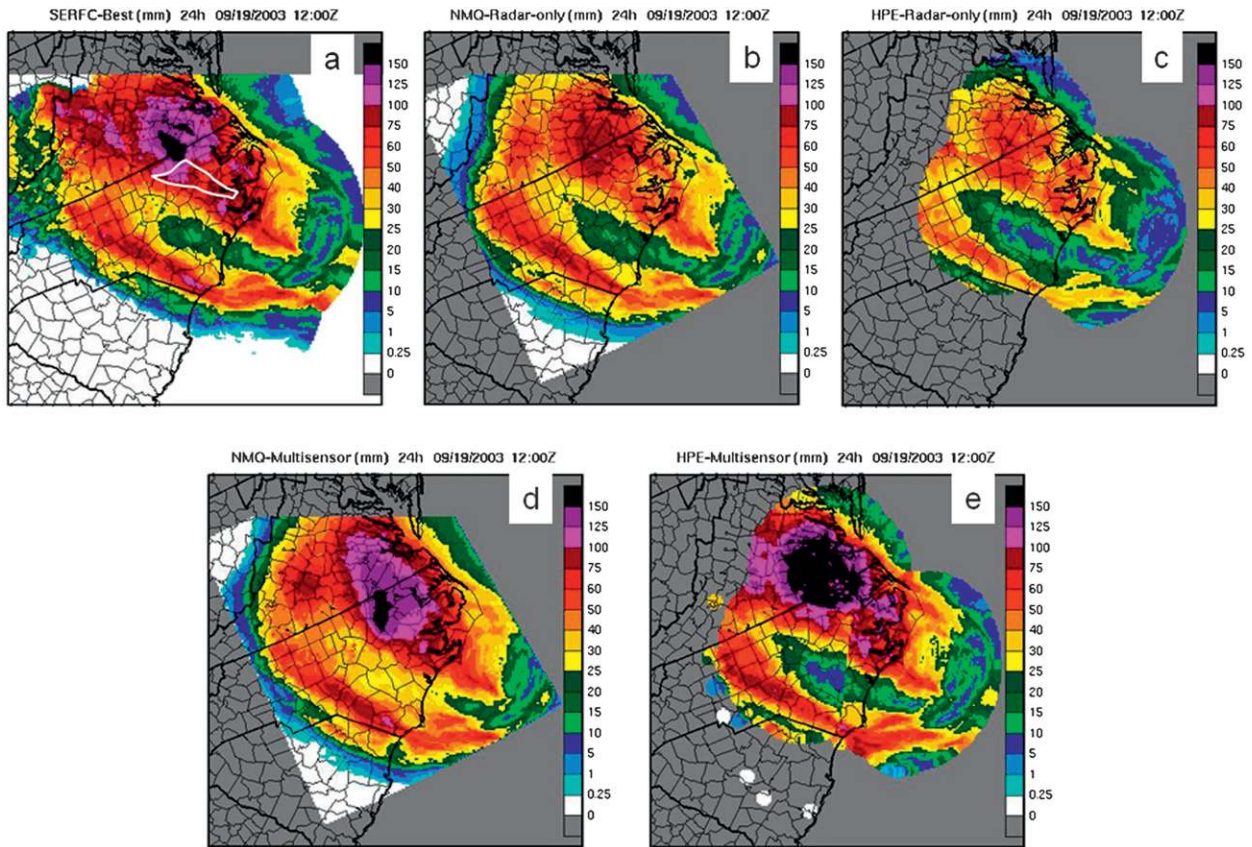
FIG. 2. Precipitation accumulations (mm) for the 24-h period ending 1200 UTC 19 Sep 2003, from (a) SERFC operational analysis, (b) NMQ radar only, (c) HPE radar only, (d) NMQ multisensor, and (e) HPE multisensor. White outline in (a) indicates location of the Tar basin.

event (not shown), where the HPE algorithm also yielded a greater level of underestimation.

The spatial distribution and amount of precipitation also differed appreciably among the SERFC and radar-only algorithms during the short event of 26 December 2004 (Fig. 3). The SERFC analysis contains some artifacts from radar coverage boundaries as illustrated by the northwest–southeast-oriented line in east-central North Carolina (Fig. 3a). This line does not appear in the NMQ or HPE analyses, which used no input from radars south of Morehead City, North Carolina. Some artifacts due to bright banding are plainly evident in the HPE radar only (Fig. 3c). Differing assumptions regarding Z–R relationships and multiple radar data merging are reflected in differences between Figs. 3b,c. As in the September 2003 example, differing treatment of gauge–radar merging produced multisensor fields with some obvious differences (Figs. 3d,e).

The introduction of gauge–radar bias correction greatly reduced the differences among the estimation packages. The NMQ and HPE gauge–radar multisensor algorithms differ in their treatment of gauge input. The NMQ radar–gauge analysis operates by calculating a multiplicative

gauge–radar bias factor separately at each gauge location with precipitation and then applying an objective analysis of the bias factor field to all points with nonzero radar precipitation (Ware 2005). The HPE gauge–radar analysis option applied in this experiment was the closest operationally available analog to that used in NMQ. In the HPE approach, a single mean-field bias correction factor was first estimated from the gauge–radar pairs in each radar umbrella and applied to the QPE data from that umbrella (Seo et al. 1999). The resulting radar QPE field was then merged directly with rain gauge observations, following the method described by Seo (1998). Both of these multisensor approaches adjusted the final precipitation estimates to values much closer to that of the SERFC analyses (Figs. 2d,e and 3d,e). In the case of HPE, some underestimation was still suggested for the September 2003 event (Fig. 2e).

6. Statistical analyses of radar-only and multisensor estimates relative to rain gauge reports

The research team compared grid values from NMQ and HPE to daily precipitation totals from cooperative
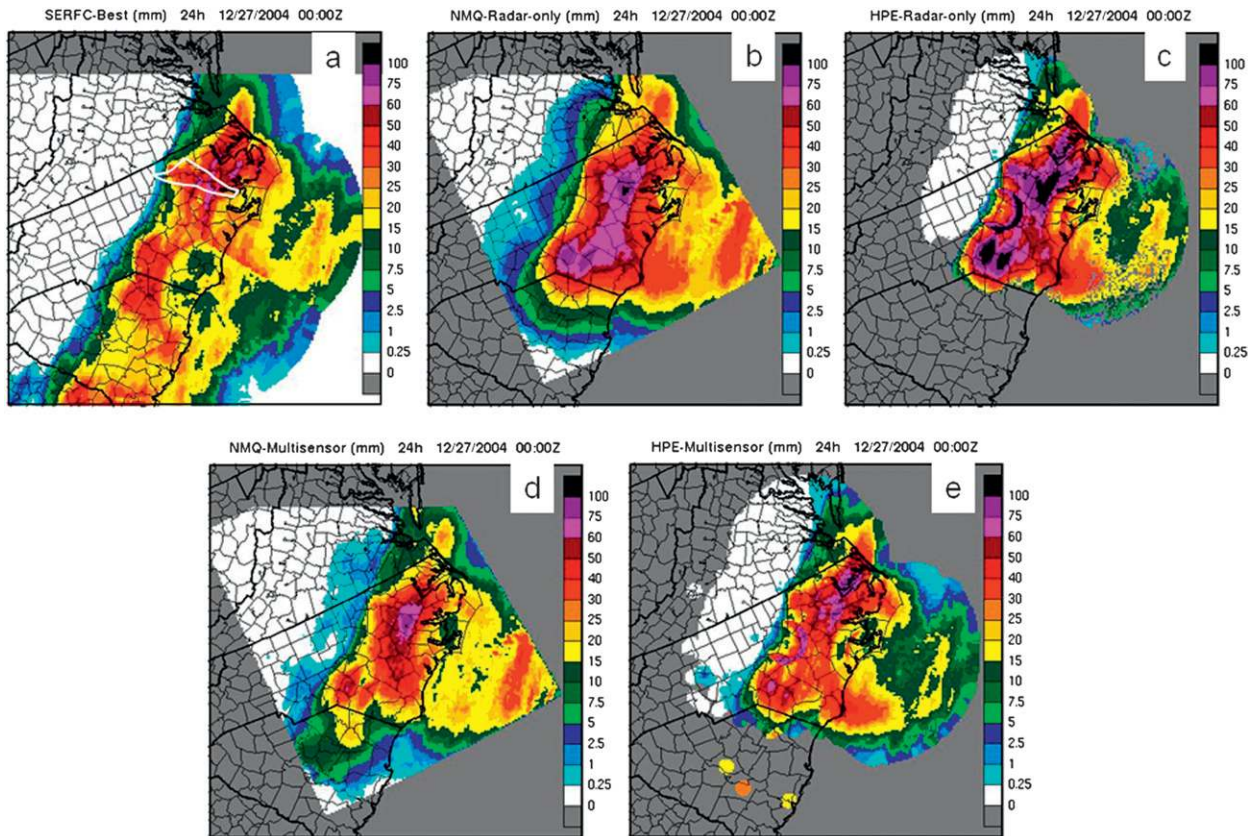
FIG. 3. As in Fig. 2, but for precipitation for the 24-h period ending at 0000 UTC 27 Dec 2004.

observer reports archived at NCDC to complete an objective comparison of the performance of features within each of the algorithm suites. Precipitation analysis and radar–gauge comparisons were carried out for 68 daily reporting sites collected from NCDC archives and from three hourly reporting sites collected from NCDC, USGS, and North Carolina State Department of the Environment archives. These hourly reference reports were withheld from the NMQ and HPE input for gauge-biasing algorithms. The statistics shown in Figs. 4 and 5 are based on precipitation pairs: that is, those pairs in which the reference gauge reported a nonzero value or at least one of the radar or multisensor estimates was nonzero. This criterion provided 425 daily pairs and 265 hourly pairs.

A comparison of 24-h values during each of the three storm periods (Fig. 4, black and gray squares for NMQ and HPE radar only, respectively) shows that radar-only NMQ and HPE both generally underestimated daily precipitation. The greatest percentage of underestimation occurred during the Isabel and Alberto (warm season) events (Figs. 4a,c), as inferred previously. For example, during the September 2003 Hurricane Isabel

event, the radar/gauge ratio was 0.75 for NMQ and 0.54 for HPE radar-only estimates, respectively. There was a smaller negative bias during the 2004–05 cool-season events (Fig. 4b). The NMQ radar only consistently gave biases closer to 1, suggesting that the adaptive Z–R relationship and freezing-level identification algorithms had a positive impact on the estimates.

During each storm period and over all the events combined, the correlation between gauge values and the radar-only estimates from NMQ and HPE was similar (0.82 and 0.80 over all the events for NMQ and HPE, respectively, as shown in Table 1). However, this difference of 0.02 in the correlation is still statistically significant at the 5% level, based on a t test that accounts for the two correlation coefficients being derived from a matched sample of rain gauge reports and for the high degree of correlation between the NMQ and HPE estimates (Carter-Clark 1997). The low bias in HPE radar-only precipitation had appreciable impact on hydrologic simulations, as will be shown.

The introduction of gauge bias correction appreciably corrected the biases, as indicated by the distribution of gauge–multisensor pairs closer to the zero-bias diagonal
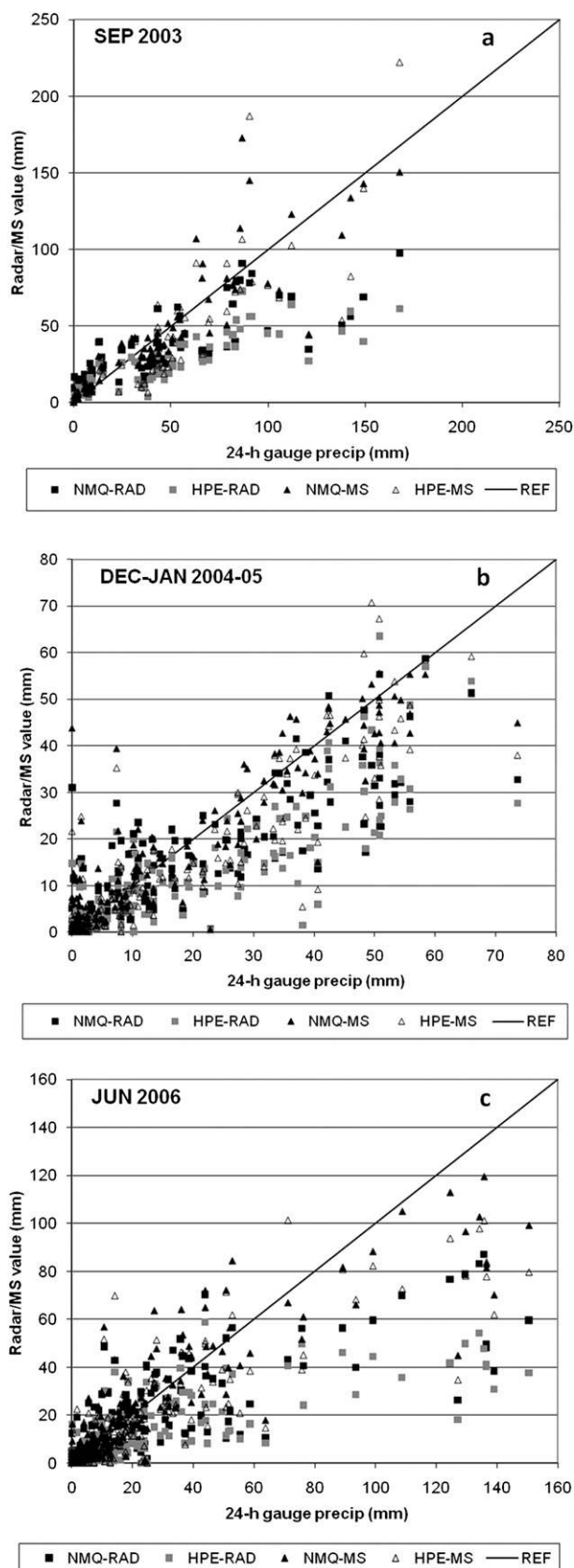
line (black and open triangles for NMQ and HPE in Fig. 4). The HPE multisensor estimates still had an appreciably low bias during all the events, with multiplicative bias factors between 0.8 and 0.85. The NMQ multisensor estimates had biases between 0.95 and 1. For each of the three events, the statistics presented in Table 1 show that the NMQ and HPE bias correction algorithms increased the accuracy of the estimates as the linear correlation of the estimates to reference gauge amounts increased to 0.89 and 0.85 for NMQ and HPE, respectively.

An analysis of multisensor estimates with hourly precipitation values yielded similar results, except that the apparent underestimation by the radar algorithms was smaller (Fig. 5 and Table 2). Within most of the individual events and overall, the HPE multisensor estimates had biases between 0.9 and 1. For both the radar and multisensor algorithms, the NMQ continued to yield biases closer to unity than HPE did.

These verification results, which incorporate data from multiple radars, clearly indicate a need to adjust radar algorithms to correct overall low bias, particularly during the warm season. The NMQ radar-only algorithm, which includes features for dynamic adjustment of Z–R relationships based on the detection of changes in hydrometeor types, performed better overall than HPE did. The NMQ gauge–radar multisensor algorithm also performed better than the HPE multisensor option applied in this phase of the experiment did.

7. Streamflow simulations based on differing precipitation inputs

The second phase of our study was carried out to document the impact of QPE accuracy on simulations of a state-of-the art distributed hydrologic model. Although our goal in improving QPE is improved hydrologic prediction, it is possible for improvements in input QPE to be masked by limitations to hydrologic models or for errors in the QPE to be magnified by the nonlinear nature of hydrologic response.

The impact of QPE on streamflow simulations could be most thoroughly documented by generating estimates through each of the four algorithm suites for extended periods of time, preferably 8 yr or longer, and then

←

FIG. 4. Comparison between NMQ and HPE gridded analyses to reference 24-h rain gauge reports. Statistics are for collocated gauge–radar pairs with at least one system reporting nonzero precipitation, for (a) September 2003, (b) December 2005–January 2005, and (c) June 2006.
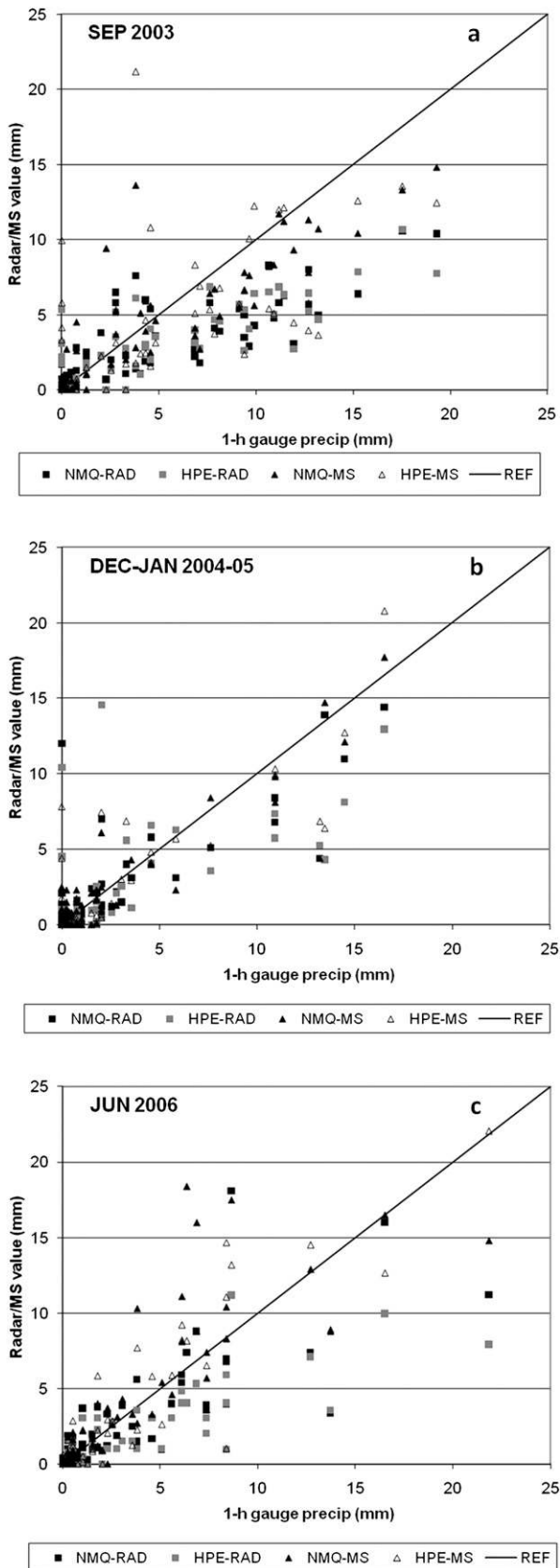
FIG. 5. As in Fig. 3, but for 1-h reference rain gauge reports.

calibrating a hydrologic model of the Tar–Pamlico River basin for subbasins well upstream of the tidal plain to optimize the accuracy of the simulations for each input QPE time series. We did not envision this level of effort and followed a simpler approach.

The Tar–Pamlico River system was modeled with a version of the SAC-SMA in which soil parameters were specified a priori through application of existing datasets for terrain, soil properties, and land use (Y. Zhang et al. 2011). As will be shown (in section 8), this approach gave reasonable results relative to discharge observations when driven by the operational SERFC QPE time series. The hydrologic model was then run with data from each of the remotely sensed QPE time series for the three storm events, by replacing the SERFC input for the relevant time periods. The hydrologic model output showed a realistic sensitivity to these differing QPE inputs. In general, the simulations based on SERFC input and NMQ and HPE multisensor input were fairly close to observed stream discharge values collected at USGS gauge locations. An exception was the smallest basin, gauged at SIMN7; possible reasons for this poor performance are explained below.

The hydrologic model was run for a continuous period from 1 January 2003 to 30 June 2006, with a 1-h time step and a computational mesh defined by the HRAP ($\sim$4 km) grid. Channel flow was assumed to be initially zero at all points. Except for the storm events, precipitation input was always taken from the SERFC operational hourly QPE grids.

Under the assumptions outlined above, HL-RDHM provided realistic simulations during the period September 2003–June 2006. The linear correlation between the simulated and observed flow over most basins (Fig. 6) was between 0.78 and 0.85 (explained variance between 0.60 and 0.72). Nash–Sutcliffe (Nash and Sutcliffe 1970) efficiency values, which are sensitive to both correlation and bias, ranged from 0.55 to 0.65. The quality of the simulations for the SIMN7 outlet was the poorest. The SIMN7 basin is small and in flat terrain. These factors make it possible that the gridcell connectivity was not optimally specified by the generating algorithm.

The lack of specific calibration and the assumption of no base flow had some effects on the absolute accuracy of the flow simulations. Overall, the observed total unit area discharge for all basins was about 9×10^{-3} m^3 s^{-1} km^{-1}, which the model underestimated by 10%–20% (Fig. 7). The simulations were also underdispersed, with flow standard deviations 20%–30% smaller than the observed in most of the basins. As will be shown in sections 8 and 9, simulations of magnitude of flood crests during our defined precipitation events were generally fairly close to the observed.

TABLE 1. Statistics for 425 nonzero 24-h gauge–radar precipitation estimate pairs.

	Gauge	NMQ radar only	HPE radar only	NMQ multisensor	HPE multisensor
Mean (mm)	25.9	19.4	14.4	25.6	21.2
Correlation with gauge values		0.82	0.80	0.89	0.85

8. Summary of subbasin simulation results based on NMQ and HPE QPEs

The impact of variations in QPE on hydrologic simulations in two of the Tar–Pamlico River subbasins, EFDN7 and TRVN7, are shown in Figs. 8 and 9. The simulations for these basins are characteristic of those for other basins and illustrate the effects of the major differences and similarities between the NMQ and HPE radar-only precipitation inputs. The secondary axis shows approximate stage values (m) corresponding to the discharges, based on recent USGS rating curves (U.S. Geological Survey 2010).

Both figures show simulations for 18–22 September 2003, 9 December 2004–18 January 2005, and 11–18 June 2006, with all simulations beginning and ending at 0000 UTC on the respective dates. These periods correspond to the start of the calendar day on which different precipitation sources were entered into the simulations, to approximately 48 h after the end of the precipitation event. Basin-average total precipitation estimates from all sources are shown for the storm events in September 2003, December 2004, January 2005, and June 2006 (Figs. 8a and 9a).

Differences among the observations and the simulations for these two particular subbasins are described in sections 8a and 8b. A narrative description of differences among the simulations for the other five subbasins is contained in section 8c.

a. Subbasin EFDN7

Simulations for the second-largest basin in our study, EFDN7, followed a pattern similar to that for most of the other subbasins. In general, the HPE radar-only QPE yielded lower simulated discharge than NMQ radar-only QPE did, particularly for the warm-season periods. The NMQ and HPE multisensor discharge hydrographs produced by the HL-RDHM simulations were usually close to each other and to the SERFC-forced simulation, in phase and magnitude. Because of the long time separation

between the experimental periods, the discharge for all four NMQ and HPE inputs were identical or nearly so at the start of these periods.

The radar-only QPEs were generally lower than the multisensor estimates, particularly for the September 2003 hurricane event (Fig. 8a). The NMQ radar-only QPE was closer to SERFC than the HPE radar only was, and that correlation was reflected in the streamflow simulations (Figs. 8b,c). Although the SERFC QPE produced a substantial overestimate of the first flood peak around hour 100, the radar-only QPEs produced a lower and later peak (Fig. 8b). Merging of gauge and radar data produced NMQ and HPE multisensor QPE data with a higher correlation to SERFC QPE values. This multisensor QPE data consequently produced hydrographs that were much closer to the SERFC simulations than the radar-only QPE did (Fig. 8c).

During the cool-season events, all simulations greatly overpredicted the magnitude of the observed flood peak around 27 December 2004 (hour 450, Figs. 8d,e). The discharge overestimates for the HPE radar-only precipitation input correspond to stage errors of ~3 m based on rating curves derived from current USGS stage–discharge charts. Of note, the application of rain gauge correction lowered the degree of overprediction. Otherwise, the discharge simulations forced by HPE radar-only QPE were generally lower than those forced by NMQ radar-only QPE, and the discharge simulations forced by NMQ and HPE multisensor QPE were close to the discharge values produced using SERFC QPE.

For the June 2006 event, the discharge simulation produced using NMQ radar-only QPE was consistently closer to that of the SERFC simulation and to USGS observations than the simulation using HPE radar-only QPE was, which generally under simulated discharge relative to the other time series. For the main storm peak around hour 120, the HPE radar-only QPE-driven discharge was about 50% of the observed, whereas NMQ radar-only QPE simulation was within 20% of the observed (Fig. 8f). Incorporation of gauge data greatly

TABLE 2. Statistics for 265 nonzero 1-h gauge–radar precipitation estimate pairs.

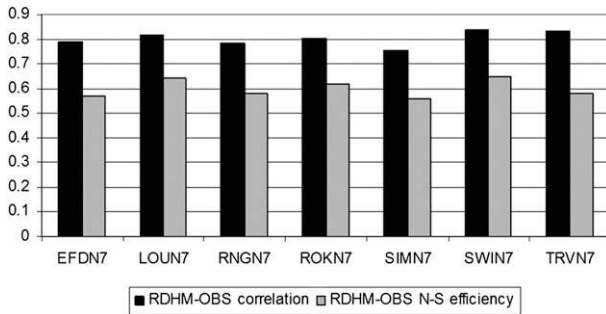|	Gauge	NMQ radar only	HPE radar only	NMQ multisensor	HPE multisensor
Mean (mm)	2.67	1.96	1.62	2.65	2.41
Correlation with gauge values		0.81	0.78	0.86	0.83

FIG. 6. Long-term linear correlation and Nash–Sutcliffe efficiency for RDHM simulations based on the SERFC operational gridded precipitation analysis, for each basin. Statistics are for the period September 2003–June 2006.
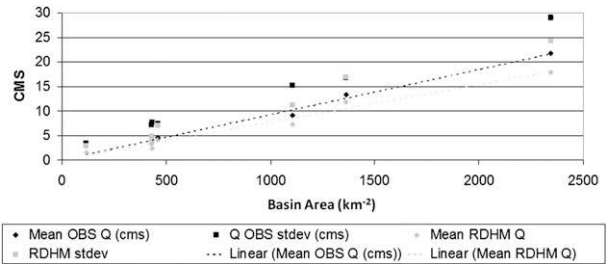

FIG. 7. Long-term mean discharge and standard deviation of discharge (in cubic meters per second, m³ s⁻¹, or cms), both observed and simulated, as functions of total area for the seven basins in the evaluation experiment. Statistics are for the period September 2003–June 2006. RDHM simulations tend to underpredict the mean and standard deviation of discharge.

increased the HPE precipitation and corresponding discharge, to values close to those simulations forced by SERFC precipitation estimates (Fig. 8g).

b. Subbasin TRVN7

During the September 2003 event, the NMQ and HPE radar-only QPE (Fig. 9b) produced nearly identical

streamflow discharge traces for the event period. Addition of rain gauge information (Fig. 9c) produced hydrographs very close to those of the SERFC simulation and the observed; however, all simulations underestimated the discharge peak.

For the cool season, all discharge simulations were very similar in the early December and January storm
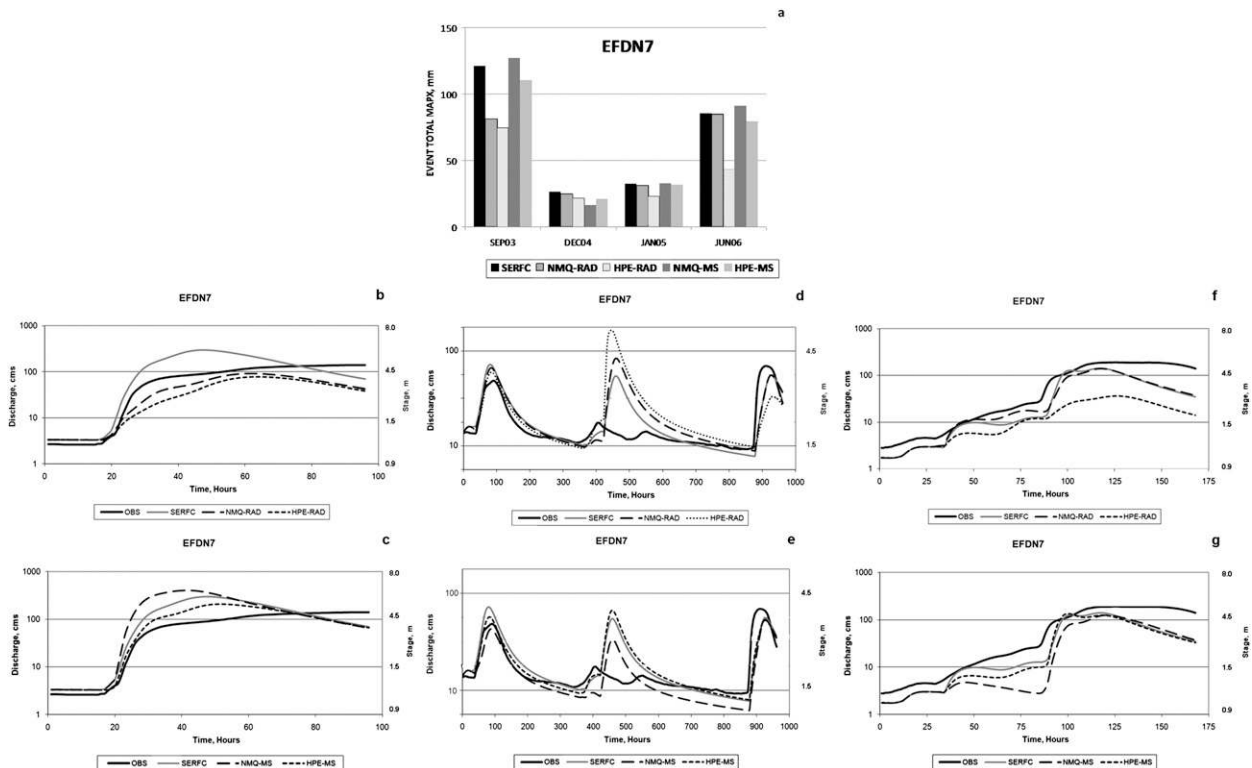

FIG. 8. For basin EFDN7, (a) mean areal precipitation from all algorithms for the four major precipitation events, and resulting streamflow simulations for (b),(c) September 2003, (d),(e) December 2004–January 2005, and (f),(g) June 2006. All hydrographs show observed discharge (black) and RDHM simulations based on SERFC (gray), NMQ (dashed), and HPE (dotted) traces. Featured are (b),(d),(f) radar-only QPE and (c),(e),(g) simulations from gauge–radar multisensor QPE. Approximate stage values (m) are shown on the right-hand scale.
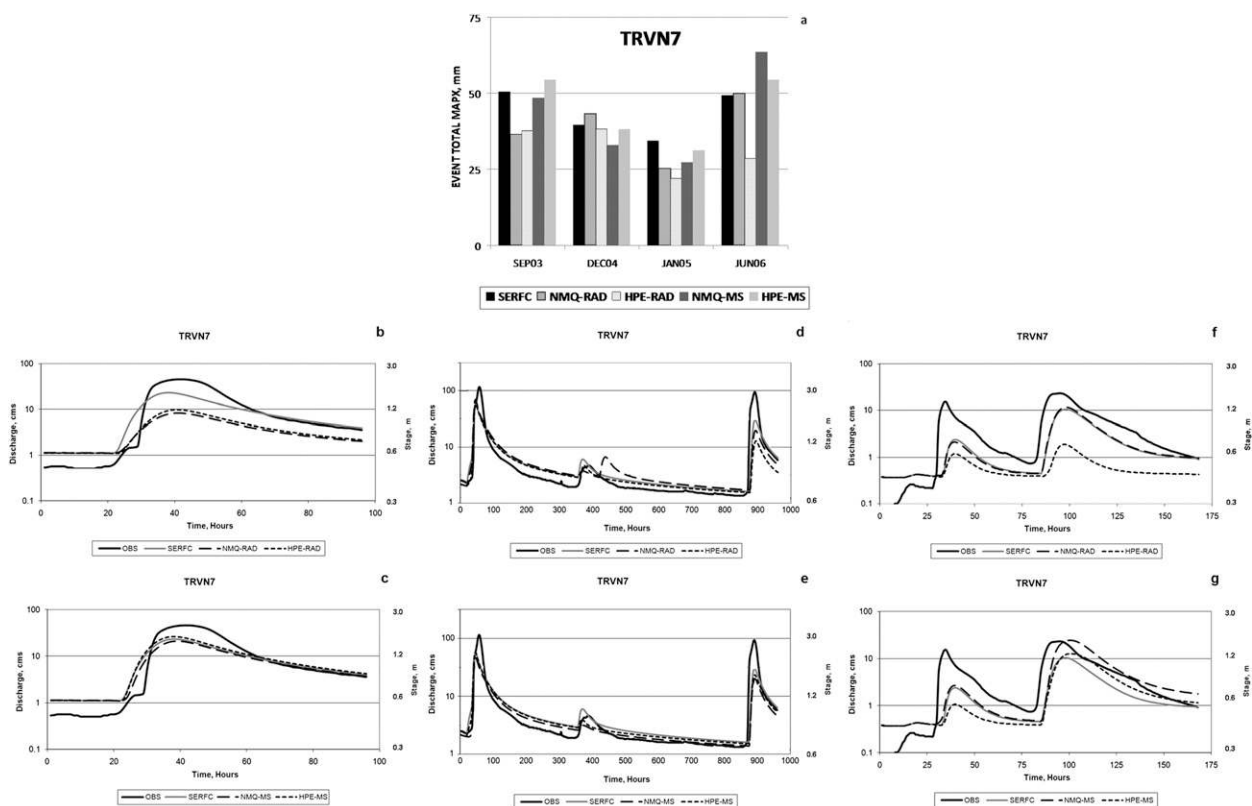
FIG. 9. As in Fig. 8, but for basin TRVN7.

events. The NMQ radar only was closer to the results of the SERFC simulation than the HPE radar only in January was (Figs. 9d,e). For the late December event, the NMQ radar-only precipitation produced an anomalous late peak in streamflow discharge, not reflected in the HPE radar-only or SERFC simulations, which captured the magnitude and timing of that minor precipitation event closely. This anomalous peak was associated with the overestimation of precipitation that affected some basins on 26 December (Fig. 3). The addition of rain gauge input substantially corrected the overestimate (Fig. 9e). The radar-only QPE simulations of the January peak were somewhat low and again the addition of rain gauge information to NMQ and HPE caused their respective HL-RDHM discharge hydrographs to more closely approximate that of the SERFC-QPE simulation (Fig. 9e).

Observed flow and all simulations started from a very low discharge state in the June 2006 event (Figs. 9f,g). The SERFC and radar-only QPE simulations underestimated peak discharge, which was observed at around 20 and 30 cms, following the two main precipitation events. Discharge simulations forced by the NMQ radar-only QPE were much closer to SERFC than those forced by the HPE radar-only QPE throughout the

event period. Incorporation of rain gauge data (Fig. 9g) had little impact on the NMQ and HPE simulations of the first peak but greatly increased the QPE and peak discharge during the second event (hour 120).

c. Distribution of flood peak simulation errors in terms of stage error

We can estimate the effects of the input QPE differences on river stage by applying current rating curves, supplied by the USGS, to the discharge time series. In Fig. 10, stage errors (in m) for the NMQ- and HPE-based simulations are shown as functions of the errors associated with the SERFC simulations. For both EFDN7 and TRVN7 (Figs. 10a,b, respectively), there is a strong correlation among the errors associated with different inputs. The largest errors for the HPE and NMQ simulations are larger than those associated with the SERFC simulations (2.8 m versus 1.5 m for EFDN7 and 1.75 m versus 1.25 m for TRVN7). In general, the errors for the multisensor QPE are smaller and closer to errors observed from the SERFC-based simulations than those for the radar-only simulations. Particularly for EFDN7, the HPE radar-only (denoted as HPE-RAD in Figs. 10a,b) simulations produced some very large errors in excess of 2 m, associated with the December 2004 events. The
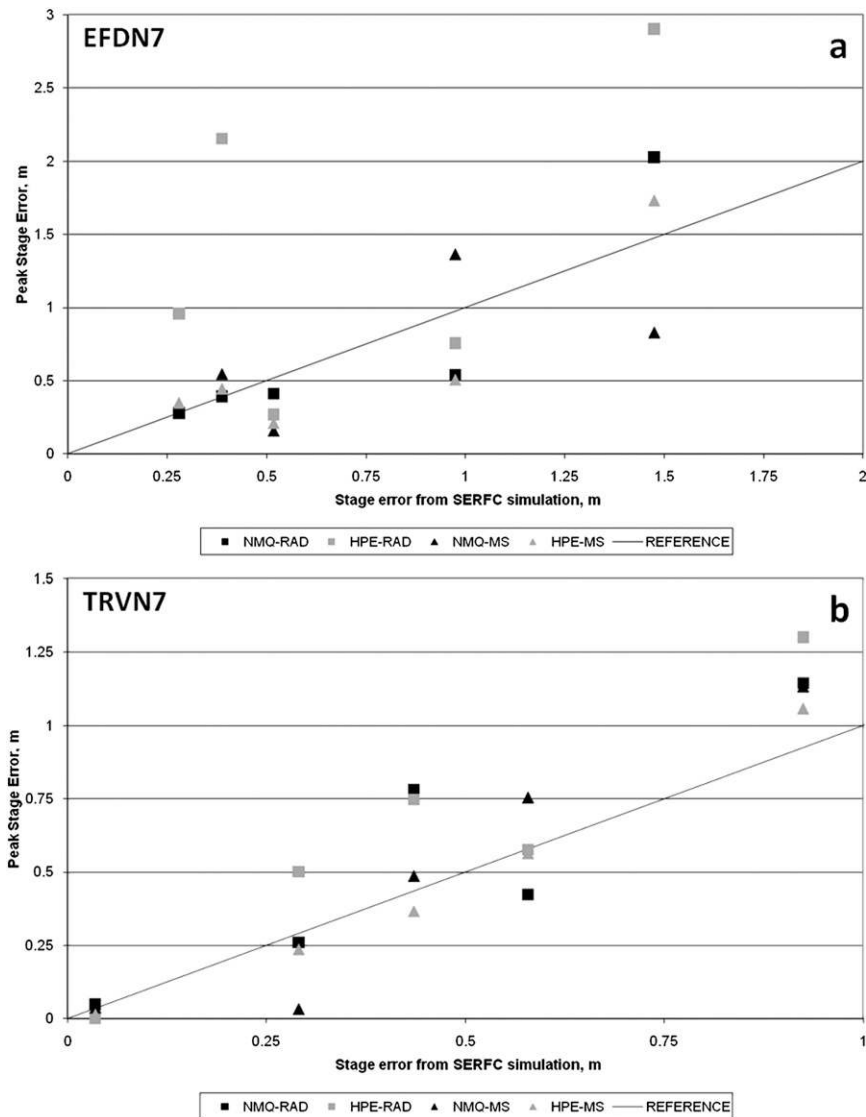
FIG. 10. Approximate peak stage errors for simulations driven with NMQ and HPE precipitation as functions of the stage error for the SERFC-driven simulation, at (a) gauge for EFDN7 and (b) gauge for TRVN7. Errors for radar-only precipitation are shown as black and gray squares, and errors for multisensor precipitation are shown as black and gray triangles. The diagonal line shows a no-bias reference.

various simulation errors for TRVN7 were more closely correlated among each other than those for EFDN7 were, with the largest and smallest stage errors generally falling within a range of 0.5 m.

d. Summary of results for other subbasins

The hydrologic simulations for these subbasins all followed a similar pattern. The simulations forced by NMQ radar-only QPE were generally closer to those based on SERFC QPE input and to observations than the simulations forced by HPE radar-only QPE were.

Incorporation of rain gauge data consistently adjusted both NMQ and HPE precipitation closer to SERFC QPE values, and the corresponding HL-RDHM simulations also more closely tracked those based on SERFC input. In several of the basins, the anomalously high precipitation indicated by the radar-only analyses on 26 December 2004 was substantially reduced by the introduction of rain gauge bias correction, with corresponding improvements in the hydrologic simulations.

The model simulations generally featured peaks that were 10–12 h too early. Differing precipitation input
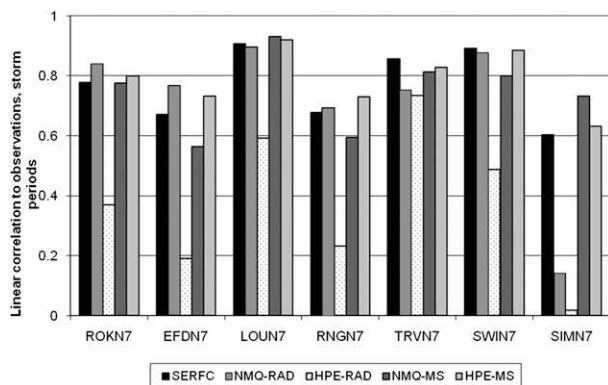
had little impact on this feature. As will be shown in section 9, the flood peaks in the simulations based on gauge–radar QPE were well correlated in magnitude with the corresponding observations, when all events were considered together.

The SIMN7 basin of only 116 km^2, near the southeastern boundary of the Tar–Pamlico basin, featured poor simulations during the storm periods, except for those forced with NMQ multisensor QPE. Given that the quality of hydrologic simulations driven by NMQ and HPE multisensor precipitation over the other basins was similar, this finding is possibly anomalous.

9. Quantitative assessment of the impact of QPE on streamflow simulations

As shown in the hydrograph traces in Figs. 8 and 9, the magnitude of the simulated streamflow and its correlation with observations is strongly dependent on the precipitation estimation source. The impact of the different QPEs on the streamflow simulations can be illustrated by considering the correlation between simulated and observed hourly streamflow over the three storm periods combined, a total of 1237 h. The correlation thus takes into account several precipitation events of markedly different character and runoff effects. Although the NMQ- and HPE-driven simulations were all nearly identical immediately prior to the events, because of the application of the same SERFC operational QPE time series during the months between events, the introduction of different QPE had significant impacts for several days thereafter.

In Fig. 11, the correlation between all simulations and observations, over the combined storm periods, is shown

separately for each basin. Except for SIMN7, there is a consistent pattern in that the highest correlations are from simulations forced by multisensor input: that is, SERFC QPE or NMQ or HPE multisensor QPE. The poorest correlations were for HPE radar-only QPE, with the NMQ radar-only input consistently yielding higher correlations. The differences in correlations between the NMQ and HPE radar-only simulations are statistically significant at the 5% confidence level. In five of the basins, the HPE multisensor QPE yielded slightly higher correlations than the NMQ multisensor did, in contrast to the results from the radar-to-gauge comparisons described in section 6. These differences in correlation coefficient are relatively small compared to those for the radar-only QPE algorithms and indicate the strong influence of rain gauge input on the multisensor products.

Very similar results (not shown) were obtained when the simulations were assessed in terms of MAE. Again, the lowest MAE was realized for the multisensor QPE-based simulations, and the NMQ radar-only QPE input consistently yielded lower MAE than the HPE radar only did. There were generally minor differences among the MAE values for the NMQ and HPE multisensor input.

These time series correlation statistics, which are based on hourly discharge values, are necessarily sensitive to timing errors introduced by precipitation errors and hydrologic model assumptions. In operational hydrologic prediction, it is also important to assess the reliability of simulations of the magnitude of floods regardless of the timing error (e.g., Reed et al. 2007). Results of this study demonstrate the positive impact of improved precipitation estimates on the simulation of the relative magnitude of discharge peaks during the storm periods.

The median error in simulated flood peaks for each basin, in terms of specific discharge and stage, is shown in Figs. 12a,b, respectively. These errors are for the five major peaks observed during the storm periods. As might be expected, the magnitude of the discharge errors was positively correlated with basin size. To simplify the comparison among basins, the errors were divided by the basin size to yield specific discharge errors. Because of the nonlinearity of the stage–discharge relationships, the smallest discharge errors did not necessarily produce the smallest stage errors. The HPE radar-only QPE generally produced larger discharge and stage errors than the NMQ radar only did, consistent with the time series correlations. Over some basins, the HPE radar-only simulations produced large errors in two or more events; hence, the median errors for the ROKN7 and RNGN7 basins were >1.5 m (Fig. 12b). By comparison, the differences between the HPE and NMQ multisensor simulations were relatively small and neither consistently
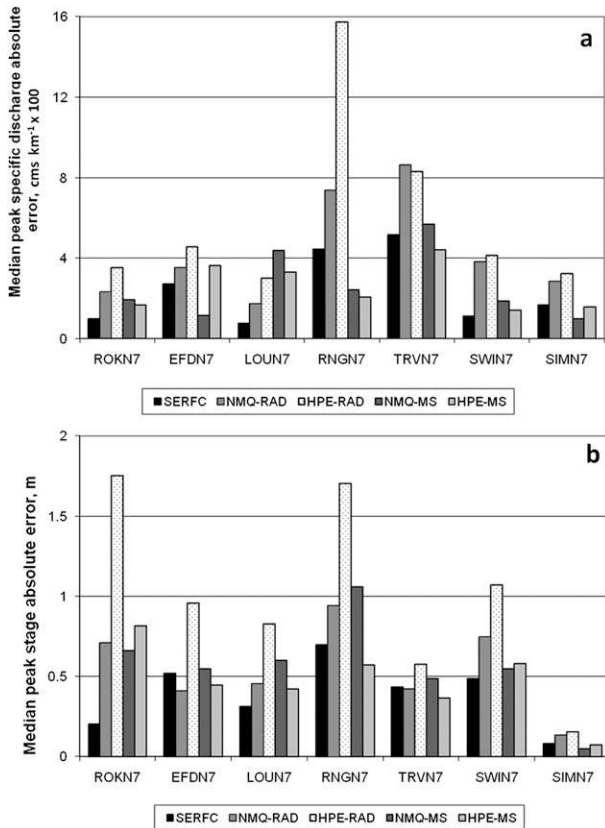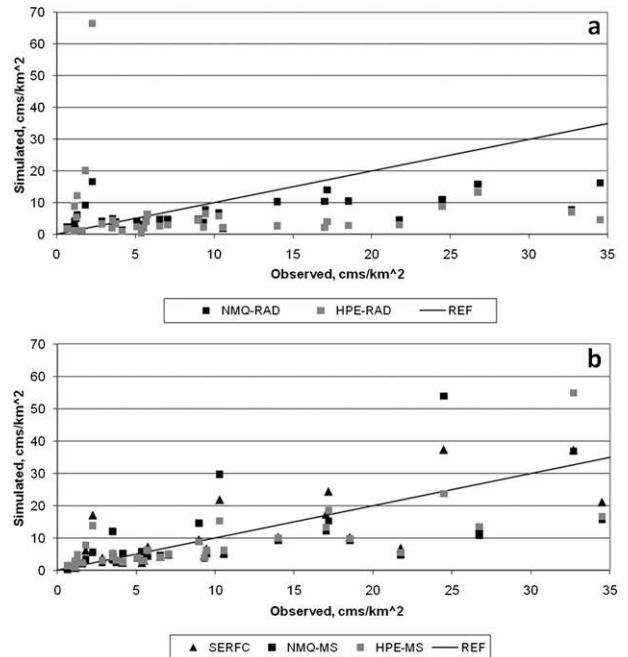
FIG. 13. Simulated storm peak discharge per unit area for all basins, five storm events, for (a) radar-only and (b) SERFC and multisensor precipitation. Values have all been multiplied by 100. The diagonal line shows a no-error reference.

produced smaller errors than the other (Figs. 12a,b). For some basins, either NMQ or HPE produced smaller median errors than the SERFC operational analyses did.

An analysis for all basins and storm events is shown in Fig. 13, where all simulated peak specific discharge values are displayed as a function of the observed. Note that all values in the figure were multiplied by 100. Consistent with the gauge verification and simulation studies, we note that the radar-only simulations (Fig. 13a) were consistently biased low and that the HPE radar only featured some larger errors in the poorly simulated minor events in the winter period, where only rather small peaks of 0.01–0.02 cms km^{-2} were observed (extreme left portion of Fig. 13a). The multisensor-based simulations (Fig. 13b) had less bias but also greater random errors for the larger events. We found, however, that both MAE and correlation relative to the observed specific discharge improved with the addition of rain gauge information, for both NMQ and HPE.

In general, the multisensor and NMQ radar-only simulations produced the smallest errors relative to observations and to the SERFC simulations. Although these

statistics are based on a very limited sample, they are consistent with the results for the rain gauge verification analysis and suggest that improved QPE has an appreciable impact on flood peak predictions.

Although there were pronounced differences between the simulations driven with the NMQ and HPE radar-only QPE, differences between the NMQ and HPE gauge–radar simulations were much smaller, and, over five of the seven basins, the HPE produced a closer approximation to observed flow in terms of overall correlation and peak error (Figs. 11 and 12). The rather dense gauge network applied in this experiment had a strong influence on the final multisensor QPE. As noted in section 5, the NMQ and HPE approaches to gauge–radar merging differ in some details. Although this result is in contrast to the rain gauge verification, where NMQ multisensor QPE had the smaller errors relative to rain gauge reports, it is possible that the limited number of geographic sampling points was not a completely accurate reflection of the two multisensor algorithms' potential for representing larger-scale precipitation patterns.

10. Summary and conclusions

This study compiled a set of gridded 1-h precipitation estimates for a study area centered on northeastern North Carolina using NSSL's National Mosaic and

Quantitative Precipitation Estimation algorithm suite (NMQ) and the NWS's operational High-Resolution Precipitation Estimator (HPE). For three event time series, the research team prepared gridded radar-only and gauge–radar precipitation estimates using the different QPE systems. The estimates covered storm events during September 2003, December 2004–January 2005, and June 2006. Although NMQ and HPE radar-only QPE algorithm suites both underestimated precipitation during Hurricane Isabel (September 2003), NMQ radar-only QPE showed less tendency, in general, to underestimate rainfall quantities for the other events and had a higher correlation to the reference rain gauge reports. The quality of both NMQ and HPE multisensor (gauge–radar) QPEs were considerably better, in terms of bias and correlation with reference gauge reports. The HPE multisensor QPE continued to underestimate the reference rain gauge reports.

An assessment of potential impacts of different QPE algorithms on hydrologic simulations for subbasins of the Tar–Pamlico, varying in size from 110 to 2300 km^2, was conducted. Four precipitation datasets, NMQ and HPE radar only and NMQ and HPE multisensor, were inserted for particular time periods within a 42-month time series compiled from SERFC QPE data obtained from the stage 4 archive. The QPE precipitation forcings were then input to a distributed hydrologic model (HL-RDHM). The resulting streamflow simulations were compared with a reference simulation based solely on the operational precipitation grids and discharge observations at seven stream gauge locations. A common set of a priori soil parameters, estimated from available soil and land-use datasets, were used for all five simulations. Results from this study indicate that the streamflow simulations were sensitive to differences in the QPE input and that the quality of the streamflow simulations was strongly correlated with the accuracy of the QPE.

The findings also indicate an operationally significant impact is provided by algorithm features unique to NMQ, particularly those that have the greatest effect in cool-season and tropical events. These features include dynamic Z–R adjustments and snow detection. The largest differences in data quality were evident between the NMQ and HPE radar-only QPE datasets, and such differences are likely to be even more significant in areas with fewer rain gauges than were available in this study area. The NMQ and HPE multisensor QPEs yielded simulation results generally close to those from the SERFC QPE, reflecting the major influence of rain gauge input on these estimates.

Further studies (Wu and Kitzmiller 2009) confirm the results of this study in terms of accuracy of the NMQ QPE relative to rain gauge reports. NMQ estimates are now available for use in real time at all RFCs in the conterminous United States and are being used operationally at some RFCs. Possibilities for operational implementation of NMQ by the NWS are being documented by a team of NWS headquarters, NCEP, and NSSL personnel.

REFERENCES

Anderson, E. A., 1976: A point energy and mass balance model of a snow cover. NOAA Tech. Rep. 19, 150 pp. [Available from Office of Hydrologic Development, W/OHD12, 1325 East West Highway, Silver Spring, MD 20910.]

Burnash, R., 1995: The NWS river forecast system—Catchment model. *Computer Models of Watershed Hydrology,* V. P. Singh, Ed., Water Resources Publications, 311–365.

Carter-Clark, D., 1997: *Doing Quantitative Psychological Research: From Design to Report.* Psychology Press, 665 pp.

Fulton, R. A., J. P. Breidenbach, D.-J. Seo, D. A. Miller, and T. O'Bannon, 1998: The WSR-88D rainfall algorithm. *Wea. Forecasting,* **13,** 377–395.

Greene, D., and M. Hudlow, 1982: Hydrometeorological grid mapping procedures. Preprints, *Int. Symp. on Hydrometeorology,* Denver, CO, American Water Resources Association, 20 pp.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.,* **77,** 437–471.

Kitzmiller, D., and Coauthors, 2008: A comparison of evolving multisensor precipitation estimation methods based on impacts on flow prediction using a distributed hydrologic model. Preprints, *22nd Conf. on Hydrology,* New Orleans, LA, Amer. Meteor. Soc., P3.4. [Available online at http://ams.confex.com/ams/pdfpapers/134451.pdf.]

Koren, V., S. Reed, M. Smith, Z. Zhang, and D.-J. Seo, 2004: Hydrology Laboratory Research Modeling System (HL-RMS) of the US National Weather Service. *J. Hydrol.,* **291,** 297–318.

Lin, Y., and K. E. Mitchell, 2005: The NCEP stage II/IV hourly precipitation analyses: Development and applications. Preprints, *19th Conf. on Hydrology,* San Diego, CA, Amer. Meteor. Soc., 1.2. [Available online at http://ams.confex.com/ams/pdfpapers/83847.pdf.]

Nash, J. E., and J. V. Sutcliffe, 1970: River flow forecasting through conceptual models. Part I: A discussion of principles. *J. Hydrol.,* **10,** 282–290.

Reed, S., 2003: Deriving flow directions for coarse-resolution (1–4 km) gridded hydrologic modeling. *Water Resour. Res.,* **39,** 1238, doi:10.1029/2003WR001989.

——, and D. R. Maidment, 1999: Coordinate transformations for using NEXRAD data in GIS-based hydrologic modeling. *J. Hydrol. Eng.,* **4,** 174–182.

——, J. Schaake, and Z. Zhang, 2007: A distributed hydrologic model and threshold frequency-based method for flash flood forecasting at ungauged locations. *J. Hydrol.,* **337,** 402–420.

Scofield, R. A., and R. J. Kuligowski, 2003: Status and outlook of operational satellite precipitation algorithms for extreme-precipitation events. *Wea. Forecasting,* **18,** 1035–1051.

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

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

U.S. Geological Survey, cited 2010: National Water Information System. [Available online at http://waterdata.usgs.gov/nwisweb/data/exsa_rat/02084160.rdb.]

Vasiloff, S. V., and Coauthors, 2007: Improving QPE and very short term QPF: An initiative for a community-wide integrated approach. *Bull. Amer. Meteor. Soc.,* **88,** 1899–1911.

Ware, E. C., 2005: Corrections to radar-estimated precipitation using observed rain gauge data. M.S. thesis, Cornell University, 87 pp. [Available from Cornell University Library, 201 Olin Library, Cornell University, Ithaca, NY 14853–5301.]

Wu, W., and D. H. Kitzmiller, 2009: Evaluation of radar precipitation estimates from NMQ and WSR-88D digital precipitation array products: Preliminary results. Preprints, *34th Conf. on Radar Meteorology,* Williamsburg, VA, Amer. Meteor. Soc., P14.4. [Available online at http://ams.confex.com/ams/pdfpapers/155669.pdf.]

Xu, X., K. Howard, and J. Zhang, 2008: An automated radar technique for the identification of tropical precipitation. *J. Hydrometeor.,* **9,** 885–902.

Zhang, J., K. Howard, W. Xia, C. Langston, S. Wang, and Y. Qin, 2004: Three-dimensional high-resolution national radar mosaic. Preprints, *11th Conf. on Aviation, Range, and Aerospace Meteorology,* Hyannis, MA, Amer. Meteor. Soc., 3.5. [Available online at http://ams.confex.com/ams/11aram22sls/techprogram/paper_81781.htm.]

——, ——, and S. Wang, 2006: Single radar Cartesian grid and adaptive radar mosaic system. Preprints, *12th Conf. on Aviation Range and Aerospace Meteorology,* Atlanta, GA, Amer. Meteor. Soc., P1.8. [Available online at http://ams.confex.com/ams/Annual2006/techprogram/paper_103897.htm.]

——, C. Langston, and K. Howard, 2008: Bright band identification based on vertical profiles of reflectivity from the WSR-88D. *J. Atmos. Oceanic Technol.,* **25,** 1859–1872.

——, and Coauthors, 2011: National Mosaic and Multisensor QPE (NMQ) System—Description, results and future plans. *Bull. Amer. Meteor. Soc.,* in press.

Zhang, Y., Z. Zhang, S. Reed, and V. Koren, 2011: An enhanced and automated approach for deriving a priori SAC-SMA parameters from the soil survey geographic database. *Comput. Geosci.,* **37,** 219–231.