A Real-Time Automated Quality Control of Hourly Rain Gauge Data Based on Multiple Sensors in MRMS System

YOUCUN QI AND STEVEN MARTINAITIS

Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, and NOAA/OAR/National Severe Storms Laboratory, Norman, Oklahoma

JIAN ZHANG

NOAA/OAR/National Severe Storms Laboratory, Norman, Oklahoma

STEPHEN COCKS

Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, and NOAA/OAR/National Severe Storms Laboratory, Norman, Oklahoma

(Manuscript received 1 October 2015, in final form 11 March 2016)

ABSTRACT

Automated rain gauge networks provide direct measurements of precipitation and have been used for numerous applications, such as generating regional and national precipitation maps, calibrating remote sensing quantitative precipitation estimation (QPE), and validating hydrological and meteorological model predictions. However, automated gauge observations are prone to be affected by a variety of error sources and require a careful quality-control (QC) procedure. Many previous gauge QC techniques were based on spatiotemporal checks within the gauge network itself, and their effectiveness can be dependent on gauge densities and precipitation regimes. The current study takes advantage of the multisensor data sources in the Multi-Radar Multi-Sensor (MRMS) system and develops an automated and computationally efficient gauge QC scheme based on the consistency of hourly gauge and radar QPE observations. Radar and gauge error characteristics related to radar sampling geometry, precipitation regimes, and freezing-level height is utilized within this scheme. This QC scheme is evaluated by testing its capability to identify suspect gauges and comparing the ability to quality-controlled gauges through statistical and spatial comparisons of gauge-influenced gridded QPE products. Spatial analysis of the gridded QPE products in MRMS resulted in a more physical spatial QPE distribution using quality-controlled gauges versus the same product created with non-quality-controlled gauge data.

1. Introduction

Accurate and timely precipitation estimates are critical for flood and flash flood forecasting and operational warning decision-making. They are also crucial input for hydrological and meteorological numerical models and for the verification of resulting predictions. Direct precipitation measurements provided by automated gauge networks have been used for these applications, as well as generating regional and national precipitation maps and calibrating quantitative precipitation estimation (QPE) products from remote sensing systems; however, these applications are entirely dependent upon the accuracy of the gauge measurements.

Numerous factors can prevent a gauge from providing accurate precipitation measurements. Strong winds and turbulence can result in gauge undercatch (e.g., Larson and Peck 1974; Wilson and Brandes 1979; Sevruk 1989; Essery and Wilcock 1991; Yang et al. 1998; Habib et al. 1999; Neŝpor and Sevruk 1999). Significant wind events, such as tropical cyclones, could result in missing gauge data because of structural damage to the gauge or loss of automated transmission (Martinaitis 2008). Partial or full blockage of the gauge orifice can greatly underestimate the measured precipitation (Sevruk 2005; Sieck et al. 2007). Instrumentation malfunctions, transmission

Corresponding author address: Youcun Qi, National Severe Storms Laboratory, 120 David L. Boren Blvd., Norman, OK 73072. E-mail: youcun.qi@noaa.gov

DOI: 10.1175/JHM-D-15-0188.1

and coding/decoding errors, telemetry issues, poor calibration, and improper timing of reported gauge values due to datalogger failure can also yield significant impacts on gauge measurements (Groisman and Legates 1994; Steiner et al. 1999; Marzen and Fuelberg 2005; Kondragunta and Shrestha 2006; Sieck et al. 2007). Poor siting or off-level placement of the gauge can yield systematic biases in the gauge dataset (Sieck et al. 2007).

Additional challenges are presented when rain gauges are tasked to measure solid winter precipitation. While some gauge configurations are capable of handling winter precipitation, the instrumentation can be subject to blockage of the gauge orifice and/or accumulation on the side of the orifice walls (Goodison et al. 1998). Unmeasured accumulations collected on the orifice of weighing gauges are not recorded until falling into the weighing bucket, usually after an increase of ambient air temperature (Goodison et al. 1998). Some tipping-bucket gauge networks utilize electrical heaters to melt solid precipitation for real-time measurement; however, increased evaporative loss of melted solid precipitation and enhanced sublimation of newly fallen snow have been observed (Metcalfe and Goodison 1992). Wetting losses on the internal walls of the gauge can also result in the undercatch of snow (Groisman and Legates 1994; Goodison et al. 1998).

Sevruk et al. (2009) provided a detailed summary of the magnitude of the aforementioned overall limitations on gauge observations through intercomparison gauge studies by the World Meteorological Organization (WMO). Wind-induced error was one of the predominant factors in observation error. Gauge undercatch in rainfall was typically around 3%, with some events approaching 20%. These wind-induced errors were magnified with snowfall. Shielded gauges had up to 40% undercatch in winds of $5\,m\,s^{-1}$ with an ambient temperature greater than $-8°C$; moreover, the undercatch of snow doubled with unshielded gauges (Goodison et al. 1998). Similar results of up to 15% undercatch for rain and 80% undercatch for snow in windy conditions were found by Sevruk (1996). Wetting losses and evaporation in rain were attributed an additional 1%–3% increase in undercatch. The use of proper correction software was also shown to improve the accuracy of gauge measurements (Lanza et al. 2005).

Steiner et al. (1999) demonstrated how non-quality-controlled gauge data could substantially affect the statistical validation of radar-based rainfall estimations; thus, a comprehensive gauge quality-control (QC) procedure is necessary. Some previous studies addressed gauge errors through intercomparisons to derive correction factors. Sevruk (1989, 1996) developed a graphical relationship and subsequent correction model for adjusting daily precipitation accumulations for tipping-bucket gauges as a function of wind speed. Upton and Rahimi (2003) proposed a set of programmable diagnostic checks to identify bad gauge observations using either data from a single gauge or by using additional information from neighboring gauges. This QC algorithm used the bucket tip frequency as input with a required temporal resolution on the order of seconds; however, this would be a huge challenge to QC gauge accumulations on large time scales.

Other studies used spatial and/or temporal analysis of gauge observations to evaluate the accuracy of gauge measurement. Kondragunta and Shrestha (2006) developed a four-tier model to QC rain gauge data that involved checks for instrument malfunction or transmission errors, comparisons against local climatological extremes, spatial and temporal consistency, and expert judgment on possible outlier values. Tollerud et al. (2005) developed a method to assess 24-h and 1-h gauge accumulation from the Hydrometeorological Automated Data System (HADS). This work first assessed multiple hours to determine if there are nonphysical repeating patterns and extreme values, followed by a neighbor check with the eight closest gauges depending on distance, and a 30-day performance check. Kim et al. (2009) also developed a methodology for enhancing and improving the quality of end-of-day HADS hourly precipitation data through temporal checks, similar to that of Tollerud et al. (2005). Decoded cumulative precipitation data were checked for temporal inconsistencies to recover missing values along with the correction of spikes, and noise in the hourly data significantly reduced the number of missing values found in monthly HADS data.

The QC of gauges can also be conducted through comparisons against other QPE sources, most notably radar-derived QPEs. One such example by Marzen and Fuelberg (2005) developed an objective rain gauge QC scheme for reanalysis using gauges from the Florida water management districts and corresponding radar-based QPE from a modified Multisensor Precipitation Estimation (MPE) algorithm. Checks within the QC algorithm included logic to remove false observations and nonzero gauge measurements with large disparities compared to the radar-based QPE through a simple classification of four radar versus gauge comparisons, followed by a nine-point check (i.e., the gauge is compared to the collocated MPE grid cell and its immediate neighboring grid cells). Variations between gauge and radar QPE values were based on strict thresholds or bias ratios. This QC scheme generally retained more than 95% of all hourly gauge observations per month, but the

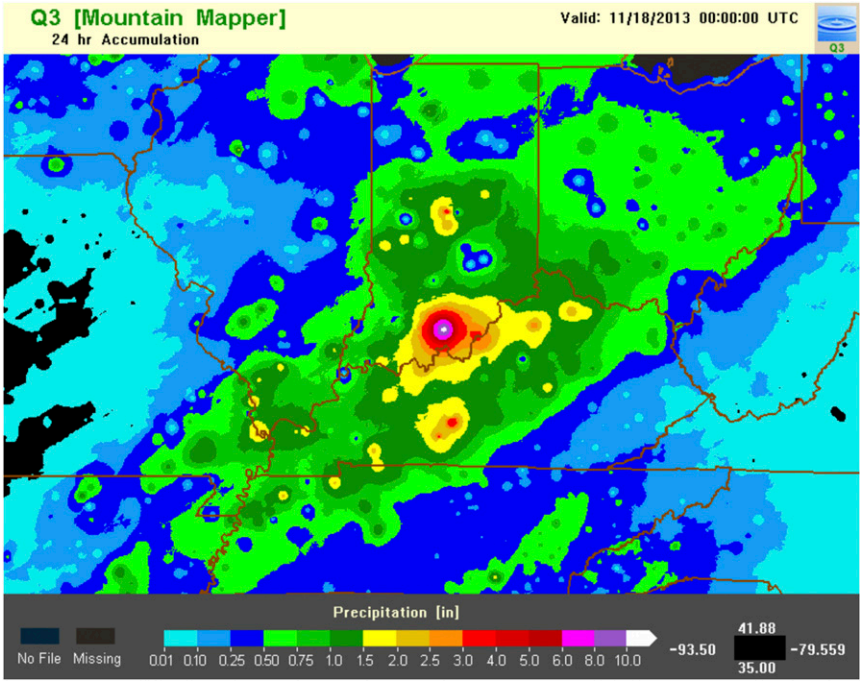

FIG. 1. The 24-h QPE accumulations of the MRMS Q3MM (see section 3 for a description) ending at 0000 UTC 18 Nov 2013.

removal of erroneous gauge observations improved the quality of a locally gauge-corrected MPE QPE.

Spatial or temporal consistency checks generally require additional computer resources, especially in instances where multiple computational iterations are needed to check each gauge with its neighbors in space or its own values in time. The computational resources and time required would be impractical for any real-time system that processes thousands of gauge observations per hour and creates products on an hourly time scale. Additionally, spatial consistency checks are limited within convective regimes as large gradients in precipitation rates and total accumulation can exist in and around convective cells (Kursinski and Mullen 2008). Temporal consistency checks are limited during some fast-moving precipitation events as valid gauge observations during these events can vary substantially with time.

The Multi-Radar Multi-Sensor (MRMS; http://mrms.ou.edu) system is an automated, real-time operational system that integrates radar, numerical prediction, and hourly gauge observations from a variety of sensor networks over the United States and southern Canada to generate seamless, three-dimensional radar and QPE mosaics at high temporal and spatial resolution (Zhang et al. 2011, 2016). The gauge sites ingested by the MRMS system utilize a variety of instrumentation types, such as tipping-bucket, weighing, and catchment gauges, as well as different instrumentation configurations (i.e., shielded vs nonshielded, heating elements, etc.). The MRMS system has been running in real time at the National Severe Storms Laboratory (NSSL) since June 2006 and operationally since September 2014.

Within an automated QPE system like MRMS, it is imperative to have quality gauge observations for use in product generation. Figure 1 depicts an example 24-h QPE accumulation for a gauge-derived product with no QC scheme applied to the gauges. This resulted in a spatially nonrealistic precipitation distribution across the region. Gauge observations that were either erroneously high or reporting no accumulations during the precipitation event generated spikes in the QPE or nonphysical holes across the area. Given the potential volume of gauge sites that can be ingested by MRMS, an automated and computationally efficient gauge QC algorithm is required to identify and remove erroneous gauge reports. This study describes and examines the use of a multisensor gauge QC scheme for real-time use in MRMS. The MRMS gauge QC scheme is designed to be straightforward and computationally efficient, since it only traverses each gauge once and does not require any spatial or temporal associations. An evaluation of the MRMS gauge QC algorithm will establish the efficiency of the gauge QC algorithm, the effectiveness of

TABLE 1. Description of the gauge instrumentation types and configurations from each network ingested into the MRMS gauge QC algorithm. The ASOS network is separated between heated tipping-bucket (HTB) gauges and All Weather Precipitation Accumulation Gauge (AWPAG) weighing gauges. DFIR stands for double fence intercomparison reference. The HADS and MADIS networks are operated by various agencies with a variety of gauge configurations. Adapted from Martinaitis et al. (2015).

Gauge source	Gauge type	Shielding	Heating element	Measurement sensitivity
HADS	Multiple gauge instrumentation configurations			0.254 mm
ASOS HTB	30-cm tipping bucket	Vinyl alter style	Yes	0.254 mm
ASOS AWPAG	Weighing	Tretyakov or double-alter style	No (antifreeze)	0.254 mm
OCS	30-cm tipping bucket	Alter style	No	0.254 mm
MADIS	Multiple gauge instrumentation configurations			
USCRN	Tipping bucket or Geonor weighing bucket	Alter style and/or DFIR windscreen	Unknown	0.1 mm (with 0.2 mm min)

capturing erroneous gauge observations, and the impacts of a QC scheme on the generation of gauge-derived QPE products.

2. Gauge data and gauge QC

a. Gauge data and other associated datasets for QC

The gauge data used in this study are a series of automated hourly gauge networks obtained from the Oklahoma Climatological Survey Mesoscale network (OCS Mesonet; McPherson et al. 2007), the National Centers for Environmental Prediction (NCEP), and the National Operational Hydrologic Remote Sensing Center (NOHRSC). Gauge networks provided by NCEP and NOHRSC include the HADS network (Kim et al. 2009), which is a real-time data acquisition, processing, and distribution system operated by the Office of Hydrologic Development (OHD) of the National Oceanic and Atmospheric Administration (NOAA)/National Weather Service (NWS), the U.S. Climate Reference Network (USCRN; Diamond et al. 2013), the Automated Surface Observing System (ASOS), and dozens of other local and regional automated networks via the Meteorological Assimilation Data Ingest System (MADIS; Helms et al. 2009). The aforementioned gauge networks utilize a variety of gauge instrumentation types and configurations that are not well documented and not distinguishable within the MRMS system (Table 1). Figure 2 shows the distribution of all automated hourly gauge stations from the aforementioned networks, which add up to 17 015 gauge sites.

The MRMS gauge QC scheme is dependent upon the use of multiple sources and products. The hourly gauge observations are compared against the MRMS gridded

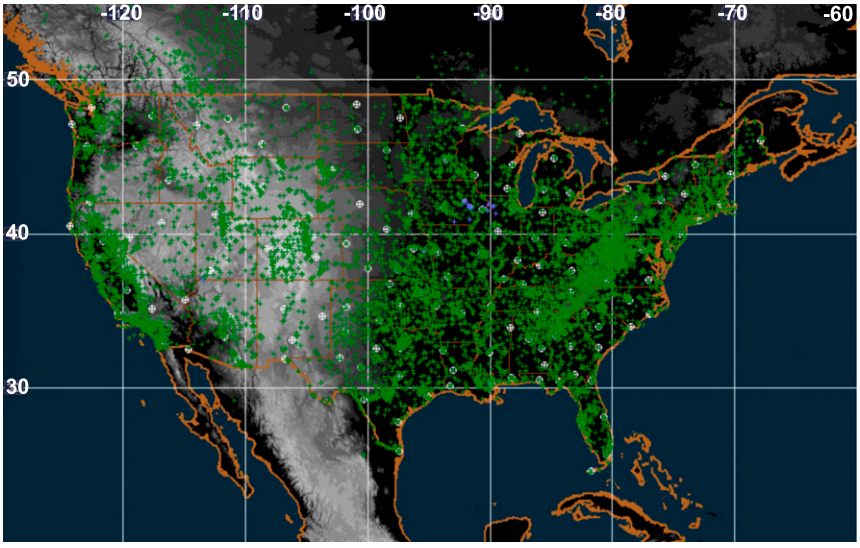

FIG. 2. Location of all hourly gauge sites (green) ingested by the MRMS system. The white circles denote the location of the WSR-88D radars in the CONUS.

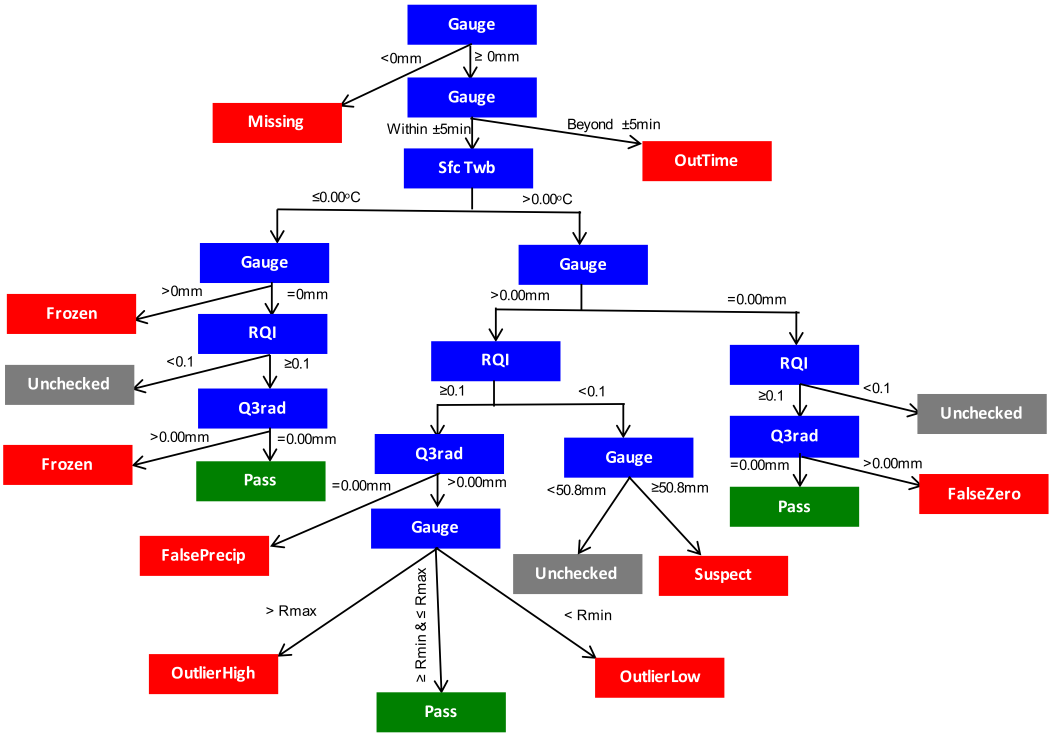

FIG. 3. Rain gauge QC decision tree used to identify erroneous or suspect gauge observations in the MRMS system.

radar-based hourly QPE (Q3rad; Zhang et al. 2011, 2016). The Q3rad product is generated every 2 min at a spatial resolution of approximately 1 km × 1 km. Q3rad is generated through a mosaicked seamless hybrid scan reflectivity and an automated precipitation classification scheme where each grid cell can be classified as one of seven precipitation types: warm and cool stratiform rain, convective rain, tropical convective rainfall, tropical stratiform rainfall, hail, and snow. Different reflectivity–rain rate (Z–R) relationships are applied to each grid cell to obtain an instantaneous precipitation rate field, and the rate fields are aggregated into hourly accumulations.

Other products utilized by the MRMS gauge QC scheme are the Radar Quality Index (RQI; Zhang et al. 2012) and the Rapid Refresh (RAP) model surface wet-bulb temperature T_{wb}. RQI was developed to indicate the potential uncertainty of radar-based QPE related to beam obstruction because of blockage along with the beamwidth/height and position with respect to the melting layer. RQI values range from zero (highest uncertainty) to one (lowest uncertainty). Values closer to zero would be due to either significant blockage, the radar beam positioned well above the melting layer, or a combination of blockage and melting layer impacts. Values closer to one would mean minimally obstructed or nonobstructed beam sampling below the melting layer. Coverage of high RQI values vary seasonally because of the height of the melting layer above the radar level. An RQI threshold of 0.1 is used in the gauge QC logic. While this would signify that the radar data are either subject to blockage and/or placed above the melting layer, the authors believe that there are enough radar data there to provide basic radar comparisons against the gauge observations and provide a greater areal coverage of gauge QC.

The RAP surface T_{wb} is utilized to delineate between environments that can produce rainfall versus environments conducive for solid, winter precipitation (Martinaitis et al. 2015). This allows for the assessment of winter weather impacts on gauges in areas of poor radar coverage based on RQI and to account for winter weather impacts on gauges when the surface ambient temperature is above freezing. The threshold value to delineate between the two environments was set at 0.00°C per Martinaitis et al. (2015).

b. *MRMS gauge QC scheme*

The logic tree for the MRMS gauge QC scheme is depicted in Fig. 3. The QC of the decoded gauge observations in MRMS begins with removing all missing observations. From there, each gauge measurement and the time of the measurement are evaluated based on the multiple sensors and products. The comparison between the hourly gauge observation and Q3rad is based on the best match between the gauge observation and a Q3rad

TABLE 2. Designation of QC flags for hourly gauge observations in the MRMS system.

QC flag	QC flag designation	Retained for MRMS use
−2	Out time window	No
−1	Unchecked	Yes
0	Pass	Yes
1	False precip	No
2	False zero	No
3	Outlier high	No
4	Outlier low	No
5	Frozen	No
6	Suspect	No

grid cell within a 5 km × 5 km area around the grid containing the gauge location. Approximately 90% of the comparisons between the gauge observation and Q3rad used the Q3rad grid cell containing the gauge, while almost all of the remaining gauge–Q3rad pairings were with an immediate neighboring grid cell. The use of a single comparison of the gauge observation to a 5 km × 5 km area of Q3rad would help mitigate any issue of gauge location sensitivity or horizontal displacement of rainfall in windy conditions.

The following subsections describe the different branches of the logic tree and how the hourly gauge measurements are flagged in the MRMS system (Table 2). Hereinafter, the hourly gauge observation will be denoted as G, and the Q3rad value will be denoted as R in the QC scheme description.

1) OUTSIDE OF TIME WINDOW

Hourly gauge data are assessed to the Q3rad product at the top of each hour. To ensure that the gauge observations temporally match Q3rad, the gauge observations must have been taken within ±5 min of the top of the hour. If the gauge observations fall outside of this window, then the gauge is flagged as "out time window" and is not used in the MRMS system. This ensures that a potential mismatch between the time of the Q3rad and gauge measurements would introduce an incorrect bias adjustment of the radar-based QPE, especially if there is an ongoing precipitation event.

2) WINTER ENVIRONMENTS AND $G = 0.00$ mm

The next check within the gauge QC scheme is if the gauge is located in an environment conducive for winter precipitation, as previously defined by the RAP model surface $T_{wb} \leq 0.00$°C. When $G = 0.00$ mm in an environment of RAP model surface $T_{wb} \leq 0.00$°C, the gauge is then compared against R and RQI. For $G = 0.00$ mm observations in areas where RQI < 0.1, the radar data from Q3rad are considered inadequate for comparison against the gauge; thus, the gauge is then flagged as "unchecked." Gauges that are marked as unchecked are still used in the MRMS system to provide observations in areas lacking radar coverage. When RQI ≥ 0.1, it is assumed that there are enough adequate radar data and coverage to QC the gauge. For observations when $G = 0.00$ mm and $R = 0.00$ mm, the gauge is flagged as "pass" and is used in MRMS. When $G = 0.00$ mm and $R > 0.00$ mm, the gauge is flagged as "frozen" and is not used in the MRMS system. The assumption being made is that most gauges have a blocked gauge orifice or are unable to register a liquid accumulation because of winter precipitation (Martinaitis et al. 2015).

3) WINTER ENVIRONMENTS AND $G > 0.00$ mm

For $G > 0.00$ mm in an environment of RAP model surface $T_{wb} \leq 0.00$°C, the gauge observation is flagged as frozen regardless of RQI and the value of Q3rad. These observations are not utilized in the MRMS system. Martinaitis et al. (2015) demonstrated that operational gauges could not provide reliable water equivalent measurements during winter precipitation events (i.e., when $R > 0.00$ mm), even if the gauges are reporting a nonzero observation. While an unknown percentage of these $G > 0.00$ mm measurements might provide an accurate value during a winter precipitation, it is difficult to distinguish these observations in a fully automated real-time system given the high uncertainty in both the gauge and radar-based observations. Martinaitis et al. (2015) demonstrated that in the current state, the retaining of gauges when $G > 0.00$ mm and $R > 0.00$ mm reduced the quality of QPE products generated in MRMS. Other observations that could fall into this category are when $G > 0.00$ mm and $R = 0.00$ mm. These observations could be a result of a mechanical or reporting malfunction or a result of thawing of solid winter precipitation.

4) NONWINTER ENVIRONMENTS AND $G = 0.00$ mm

The remainder of the gauge QC logic described here assesses gauge observations in an environment where the RAP model surface $T_{wb} \leq 0.00$°C. The assessment of when $G = 0.00$ mm is dependent upon the RQI value. For a $G = 0.00$ mm observation in areas where RQI < 0.1, the gauge is then flagged as unchecked and is retained in the MRMS system. When the RQI ≥ 0.1, the $G = 0.00$ mm is flagged as pass when $R = 0.00$ mm. If the $R > 0.00$ mm, then the gauge is flagged as "false zero" and would not be used in the MRMS system.

5) NONWINTER ENVIRONMENTS AND $G > 0.00$ mm

The assessment of $G > 0.00$ mm observations is also dependent upon the RQI value. For instances when

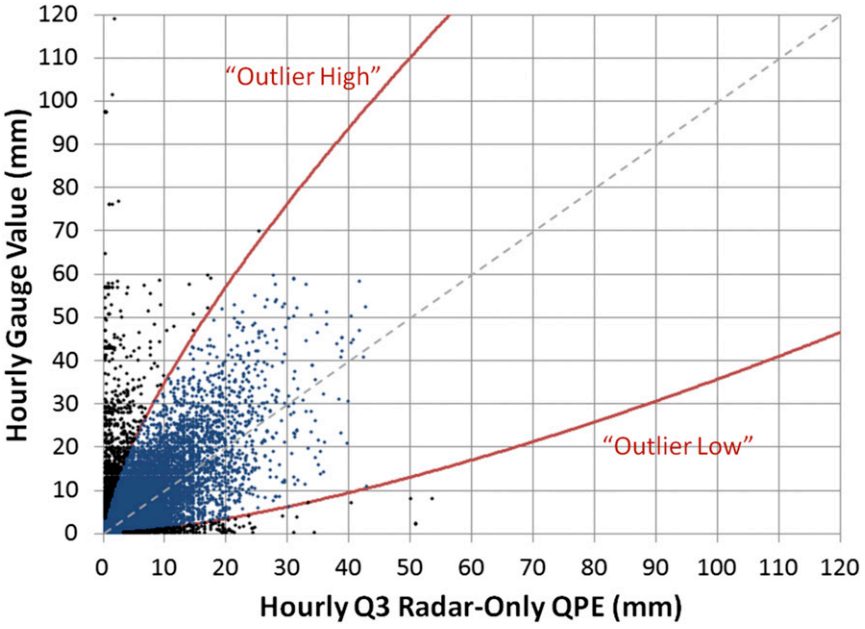

FIG. 4. Scatterplot of approximately 95 000 Q3rad and gauge pairs with the outlier high and outlier low curves shown in red.

RQI < 0.1, there is inadequate radar information to assess the nonzero gauge measurement. To provide a way to remove the more egregious nonzero gauge values, an internal study was conducted by the authors to define a threshold value to remove erroneous gauge values. An assessment of hourly gauge observations in areas of RQI < 0.1 found that approximately 97% of observations that were at least 50.8 mm (2.00 in.) were likely erroneous; thus, in the gauge QC algorithm, if $G \geq 50.8$ mm in areas where RQI < 0.1, then the observation is flagged as "suspect" and would not be used in the MRMS system. The use of a lower threshold could possibly remove a greater percentage of potential quality gauge observations. For $G < 50.8$ mm in areas where RQI < 0.1, the gauge observation will be flagged as unchecked and be retained for use in MRMS.

The evaluation of nonzero gauge observations becomes more complex when RQI ≥ 0.1. When $G > 0.00$ mm and $R = 0.00$ mm, the gauge observation is flagged as "false precip" and is not used by the MRMS system. When $G > 0.00$ mm and $R > 0.00$ mm, the gauge observation is compared against a potential maximum and minimum value based on Q3rad. The authors developed two power curves that define maximum threshold $R_{\max}$ and minimum threshold $R_{\min}$ values:

$$R_{\max} = 6.4 \times \text{Q3rad}^{0.725} + 1.0 \quad \text{and} \tag{1}$$

$$R_{\min} = 0.045 \times \text{Q3rad}^{1.45}. \tag{2}$$

Equations (1) and (2) were subjectively derived from a control dataset of approximately 95 000 nonzero G, R pairs located in all ranges of radar quality and fitted to acceptable error ranges of baseline values (Fig. 4). Approximately 2.3% of the nonzero G, R pairs from the control dataset were classified as outlier values. These equations were independently evaluated against six rainfall events (Table 3) where a time series of 130 Q3rad versus gauge pairs totaling 1337 evaluated hours was assessed. If $G \geq R_{\min}$ and $G \leq R_{\max}$, then the gauge is flagged as pass and used in the MRMS system. If $G < R_{\min}$ ($G > R_{\max}$), then the gauge observation is flagged as "outlier low" ("outlier high") and is not used by MRMS.

3. Methodology to evaluate QC gauges

Of the 17 015 gauge sites identified through the various gauge networks used in this study, the MRMS system would typically get approximately 65%, or about 11 000 observations in real time each hour (Fig. 5a). Some of the hourly gauges may not get to the MRMS system at any given hour because of the instrument malfunction, transmission issue, incorrect formatting, or some other technological issue. A drop-off of approximately 3000 gauge observations from 1600 UTC 11 June to 2300 UTC 21 June 2014 was noted in the MRMS system due to a file format change with one of the gauge data suppliers (Fig. 5a). Based on an analysis from 0000 UTC 1 January to 2300 UTC 31 December 2014, approximately 28% (4765) of hourly gauge sites reported

TABLE 3. Summary of events used for evaluation of the power curves used to define R_{max} and R_{min} to flag erroneous gauge observations as outlier high and outlier low.

Event No.	Date	Meteorological summary of event
1	5–6 Dec 2013	Winter weather/moderate-to-heavy rain
2	21–22 Dec 2013	Squall line with severe wind reports
3	22–23 Dec 2013	Heavy rainfall in southeastern United States
4	9–10 Jan 2014	Heavy rainfall from intense MCS in Florida
5	4–5 Feb 2014	Winter weather/moderate-to-heavy rainfall
6	2–3 Apr 2014	Severe storms/tornadoes, MCS event over Great Plains/Midwest
7	5–6 Sep 2014	Isolated convective rainfall in Florida

less than 9% (788 h) of the time, and approximately 24% (4083) of sites reported 100% of the time (Fig. 5b). The QC scheme was first applied retroactively to the hourly gauges ingested by MRMS in January and August 2014, and then applied to the gauge dataset for the 2014 calendar year. The gauge QC algorithm ran on a Linux workstation with sixteen 2.27-GHz processors and 12 GB random-access memory (RAM). The average time for the central processing unit (CPU) to process the gauges per hour was approximately 4.78 s, and the average RAM usage was 100 MB.

Hourly gauges observations that were flagged as pass or uncheck in the gauge QC algorithm are used to generate a number of QPE products within the MRMS system. Two such products are the locally gauge-corrected radar QPE (Q3rad_gc) and the Mountain Mapper (Q3MM) QPE (Zhang et al. 2011, 2016). Q3rad_gc is generated based on the method that is developed by Ware (2005), where a bias is generated between the gauge and collocated radar grid cell and the bias is then interpolated locally over the radar field using an inverse distance weighting (IDW) scheme. Q3MM interpolates hourly rain gauge observations using monthly precipitation climatologies from the Parameter-Elevation Regressions on Independent Slopes Model (PRISM; Daly et al. 1994, 2008) as a background field. The monthly climatologies are reduced down to an hourly scale, and a ratio between the gauge observation and the hourly PRISM value is generated and interpolated onto the background climatologies via a similar IDW scheme by Ware (2005).

To evaluate the performance of the gauge-derived QPE products using the pre-QC and post-QC gauge datasets, the Q3MM QPE outputs were evaluated against 24-h gauge observations from the Community Collaborative Rain, Hail and Snow Network (CoCoRaHS; Cifelli et al. 2005). CoCoRaHS is a unique, nonprofit, community-based network of volunteers that measure and map precipitation. The 24-h CoCoRaHS observations are generally taken at 0700 local time, but this can vary from user to user and sometimes from day to day. CoCoRaHS is utilized as an independent validation data. No QC is performed on the CoCoRaHS observations in the MRMS system. Both Q3MM products were compared against the independent CoCoRaHS gauge observations. The performance of the gauge QC scheme was quantified using a series of statistical measures. The mean bias ratio is defined as the following:

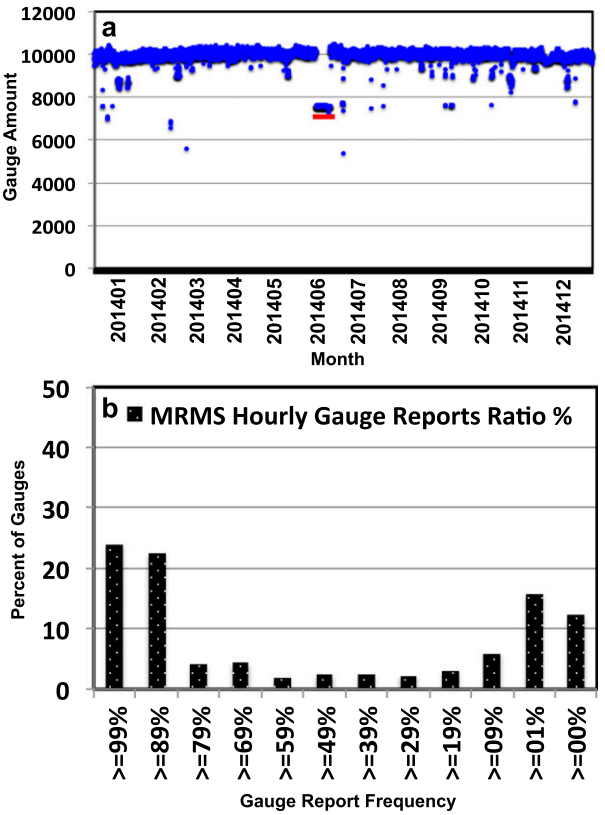

FIG. 5. (a) The number of the hourly gauge observations ingested by the MRMS system prior to gauge QC and (b) a histogram of gauge report frequency for the gauge sites processed by the MRMS gauge QC scheme. The red line in (a) denotes a change in the number of gauges ingested by the MRMS system due to a file format change with one of the gauge data suppliers.

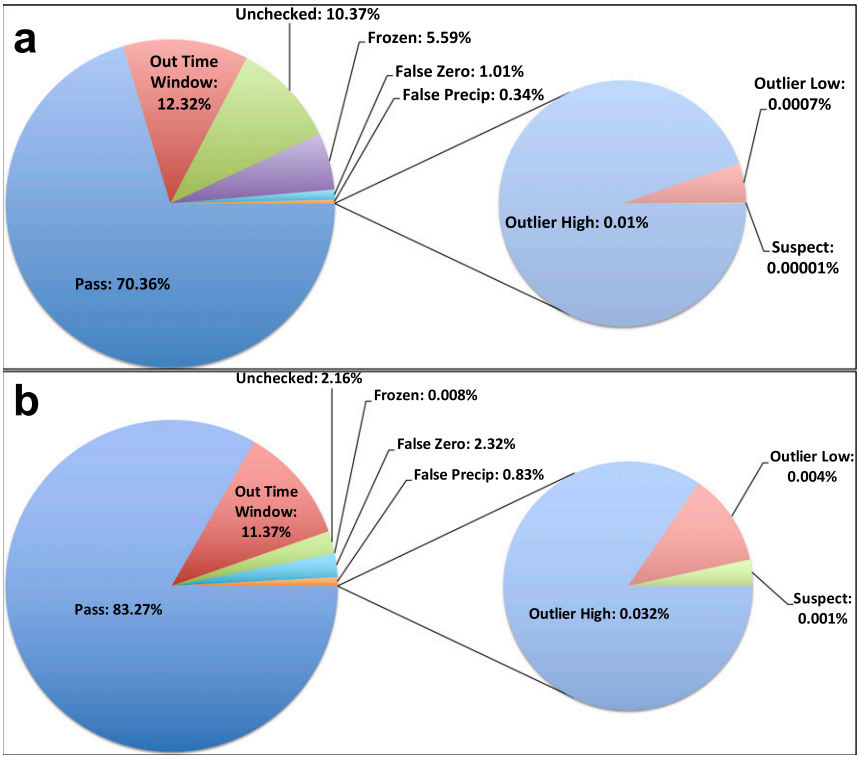

FIG. 6. Percent distribution of all gauge QC flags from the MRMS gauge QC scheme for (a) January 2014 and (b) August 2014.

$$\text{bias} = \overline{R}/\overline{G}, \tag{3}$$

$$\overline{R} = \frac{1}{N}\sum_{k=1,N} r_k, \quad \text{and} \tag{4}$$

$$\overline{G} = \frac{1}{N}\sum_{k=1,N} g_k. \tag{5}$$

Here, $\overline{R}$ and $\overline{G}$ are the averaged 24-h radar and gauge precipitation, respectively; r_k and g_k represent a matching pair of the radar-derived and gauge-observed rainfall; and N represents the total number of matching gauge and radar pixel pairs in the entire domain. A matching radar–gauge pair is found if the gauge location is within the boundary of a 1 km × 1 km radar pixel, and both the radar estimate (i.e., r_k) and the gauge observation (i.e., g_k) are greater than zero. A mean bias ratio greater (less) than 1.0 indicates that the radar has overestimated (underestimated) the rainfall, assuming the CoCoRaHS gauge reports were accurate. The root-mean-square error (RMSE) and correlation coefficient (CC) are defined as the following:

$$\text{RMSE} = \left[\frac{1}{N}\sum_{k=1,N}(r_k - g_k)^2\right]^{1/2} \quad \text{and} \tag{6}$$

$$\text{CC} = \frac{\left[\frac{1}{N}\sum_{k=1,N}(r_k g_k)\right] - \left(\frac{1}{N}\sum_{k=1,N} r_k\right)\left(\frac{1}{N}\sum_{k=1,N} g_k\right)}{\left\{\left[\frac{1}{N}\sum_{k=1,N} r_k^2 - \left(\frac{1}{N}\sum_{k=1,N} r_k\right)^2\right]\left[\frac{1}{N}\sum_{k=1,N} g_k^2 - \left(\frac{1}{N}\sum_{k=1,N} g_k\right)^2\right]\right\}^{1/2}}. \tag{7}$$

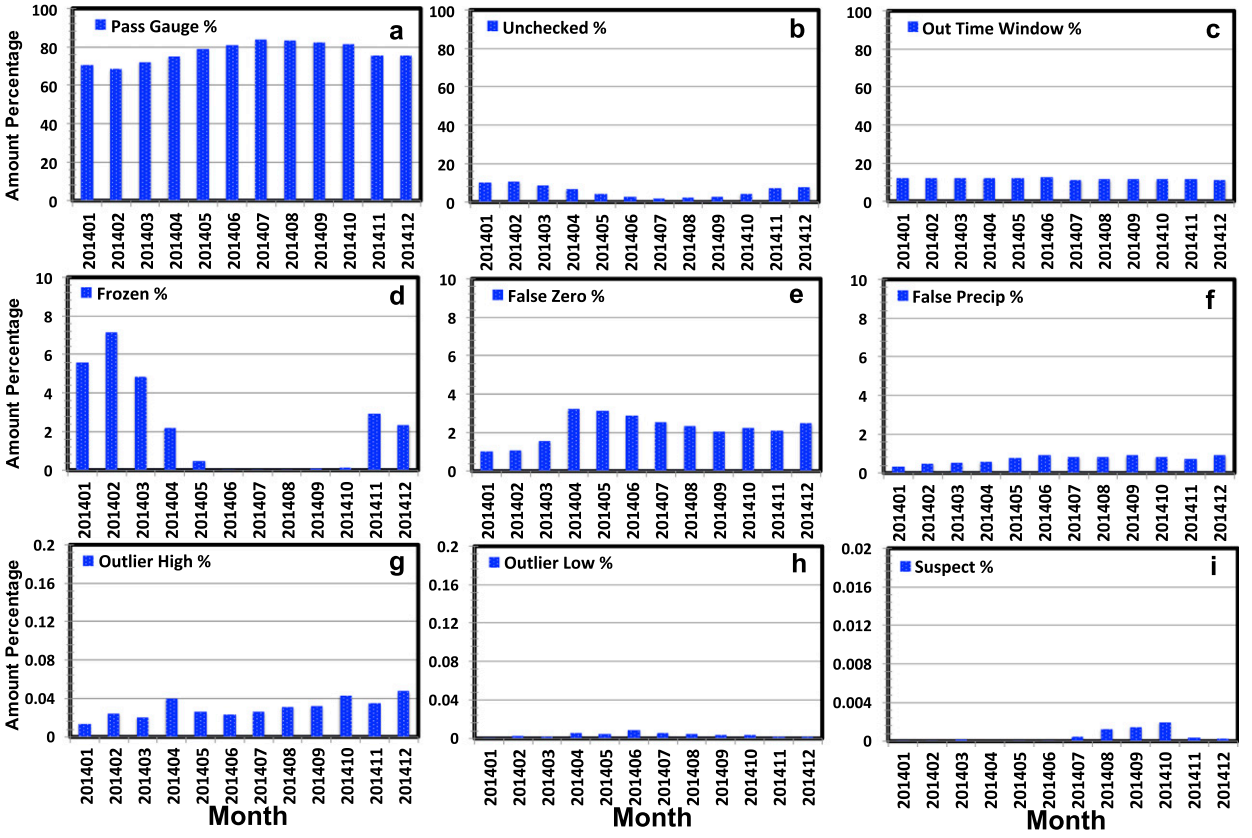

FIG. 7. Monthly percentage variation of hourly gauge observations for the following gauge classification from the MRMS gauge QC scheme: (a) pass, (b) unchecked, (c) out time window, (d) frozen, (e) false zero, (f) false precip, (g) outlier high, (h) outlier low, and (i) suspect. The year and month are represented in the YYYYMM format below each chart. Note the varying percentage scales on each graph.

4. Results

a. Gauge QC procedure

The distributions of the percent of hourly gauge observations that fell into each gauge QC criteria based on Table 2 for the months of January and August 2014 are shown in Fig. 6. There were a few gauge classifications that varied significant between January and August. The percent of gauge observations flagged as pass was 70.36% in January but was a much greater 83.37% in August. In contrast, the percent of gauge observations flagged as unchecked was 10.37% in January and 2.16%

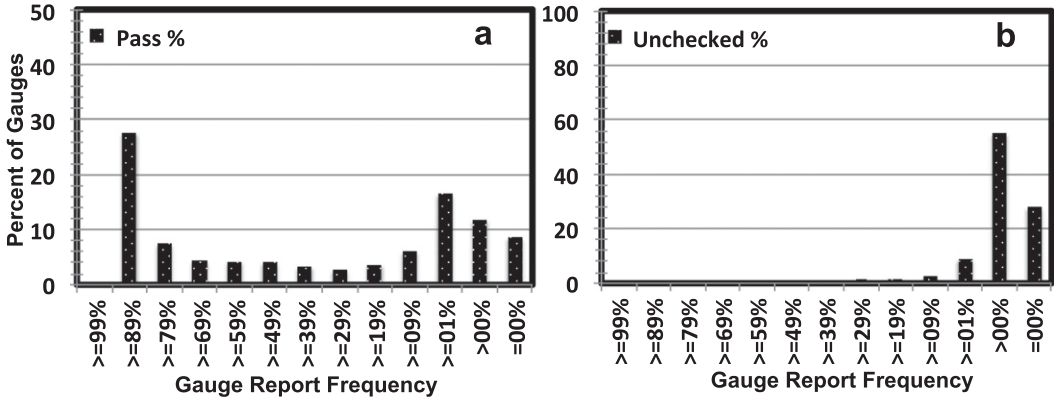

FIG. 8. Histogram of gauge report frequency for the QC flags (a) pass and (b) unchecked.

in August, while the percent of gauge observations flagged as frozen was 5.59% in January and 0.008% in August. There are two primary reasons as to why the distribution of these gauge QC flags varied. First, the performance of the radar detection via RQI is greater in warmer months (e.g., August) than in cooler months (e.g., January) because of the greater height of the freezing level. Second, the employment of the RAP model surface T_{wb} threshold of 0.00°C to flag gauges impacted by winter precipitation covered a greater areal extent across the contiguous United States (CONUS) in cooler months. All other gauge QC flags had negligible differences between January and August 2014.

A number of trends were shown when evaluating the gauge QC algorithm over the entire 2014 calendar year (Fig. 7). The greatest variability was seen in the number of hourly gauge observations that were classified as pass (Fig. 7a), unchecked (Fig. 7b), and frozen (Fig. 7d). The trends of these three gauge QC flags follow that of the aforementioned seasonal variability of the RQI coverage and the areal extent of where the RAP model surface $T_{wb} \leq 0.00$°C. Figure 8 shows the report frequency of pass and unchecked for all gauge sites. There were approximately 28% of gauge sites that had 89%–99% of their hourly observations flagged as pass, yet, there were approximately 10% of gauge sites that had nearly 0% of their observations flagged as pass (Fig. 8a). The majority of these sites that had almost no gauge observations flagged as pass were due to being outside of the ±5-min window. Approximately 58% of gauge sites had gauge observations that were flagged as unchecked less than 1% of the time over the 2014 calendar year, and 29% of gauge sites did not have a single observation that was flagged as unchecked (Fig. 8b).

The other classification had negligible variation during the 2014 calendar year. The most prominent result was the out time window classification (Fig. 7c), which accounted for 12% of the hourly gauge observations per month. Most gauges report their observations within the ±5-min window for at least 99% of their hourly observations; however, approximately 7.0% of gauge sites (1191) reported outside of the ±5-min window for at least 89% of their hourly observations. There was also some seasonal signal in the false zero classification (Fig. 7e). A local maximum can be seen during the months of April and May, which coincides with the seasonal increase in temperatures and the thawing of snow in higher altitudes and higher latitudinal regions. The following classifications also had a consistent distribution with little variation throughout the 2014 calendar year: false precip, outlier high, outlier low, and suspect. Hourly gauge observations flagged as false precip accounted for 0.4%–1.0% of all observations per month (Fig. 7f). Even though the outlier high (Fig. 7g), outlier low (Fig. 7h), and suspect (Fig. 7i) classifications attributed to a very low percentage of the total number of hourly gauge observations, they did affect the accuracy of the QPE products in MRMS that rely on the use of gauge observations.

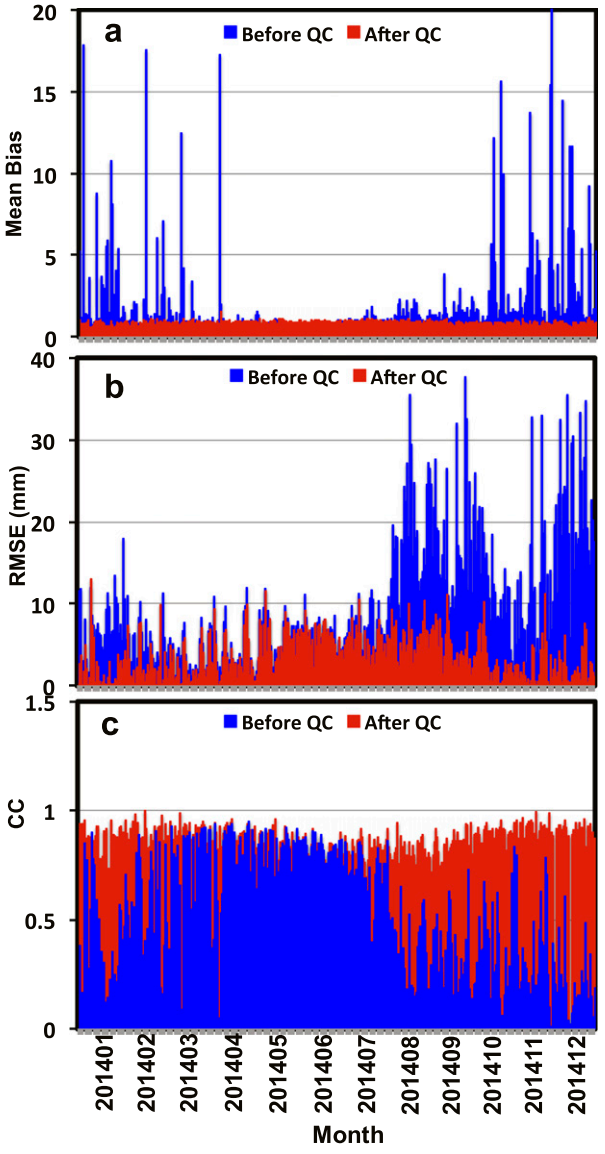

FIG. 9. The daily (a) mean bias ratio, (b) RMSE, and (c) CC values of Q3MM QPE across the CONUS before gauge QC (blue) and after gauge QC (red) has been applied.

b. *Gauge QC impact on QPEs*

Figure 9 displays the three statistical measures of the Q3MM QPE product accumulated over a 24-h period with respect to the CoCoRaHS observations for the 2014 calendar year before and after gauge QC. The

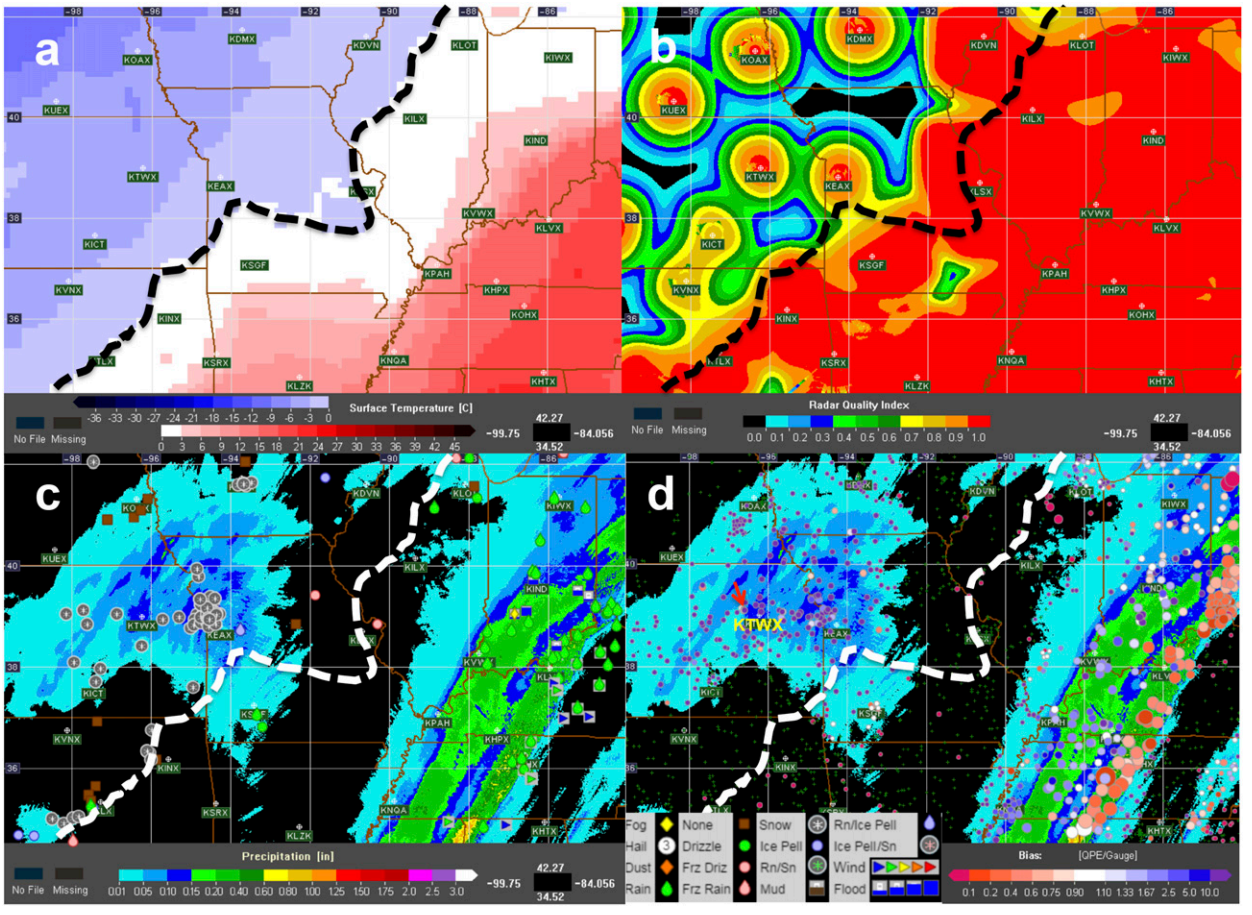

FIG. 10. (a) RAP model surface T_{wb}, (b) RQI, (c) 1-h Q3rad rainfall accumulation with mPING observations, and (d) 1-h Q3rad rainfall accumulation with gauge vs Q3rad bias ratio bubble plots ending at 0400 UTC 22 Dec 2013. The blue (red) shaded bias ratio bubble plots between the hourly radar and gauge observations indicated that the radar value at the gauge site overestimated (underestimated) when compared to the gauge observation. The dashed line in (a)–(d) denotes where RAP model surface $T_{wb} = 0.00°C$.

gauge QC algorithm consistently reduced the Q3MM QPE errors for all days in 2014 after the gauge QC scheme was applied. The mean bias ratio and CC values were closer to 1.0, and the RMSE of Q3MM was significantly reduced from an average of 20 to 5 mm h^{-1}. Improvements of the statistical measures in the cold season were greater than in the warm season. This can be attributed to the influence of winter precipitation on the gauge observations from partially or completely blocked gauge orifices, wind turbulence, or other impacts.

An example of gauge bias bubble plots of hourly gauge observations prior to gauge QC for a precipitation event at 0400 UTC 22 December 2013 along with the RAP model surface T_{wb}, RQI, and public reports of precipitation type from the Meteorological Phenomena Identification Near the Ground project (mPING; Elmore et al. 2014) is shown in Fig. 10. Hourly gauge observations in the area where snow is occurring have an extremely high bias, which could be a result of the gauge becoming stuck or blocked. The introduction of these gauges into a QPE without QC would skew the results in the cool season, as depicted in Fig. 9. Postevent thaw of accumulated winter precipitation is also problematic in QPE generation. Figure 11 shows an event over the southern Great Plains that received mixed winter precipitation types that resulted in numerous gauges reporting $G = 0.00$ mm during the event at 2100 UTC 21 December 2013. The same area 24 h later showed gauges reporting nonzero values while there was no precipitation in the Q3rad product. Without QC, these gauges would generate false QPE in the Q3MM product during these postevent thaw periods. A full study on the impacts of winter precipitation both during and after an event within the MRMS system can be found in Martinaitis et al. (2015).

Other features of the gauge QC algorithm in MRMS improved the QPE products. Significant hot spots

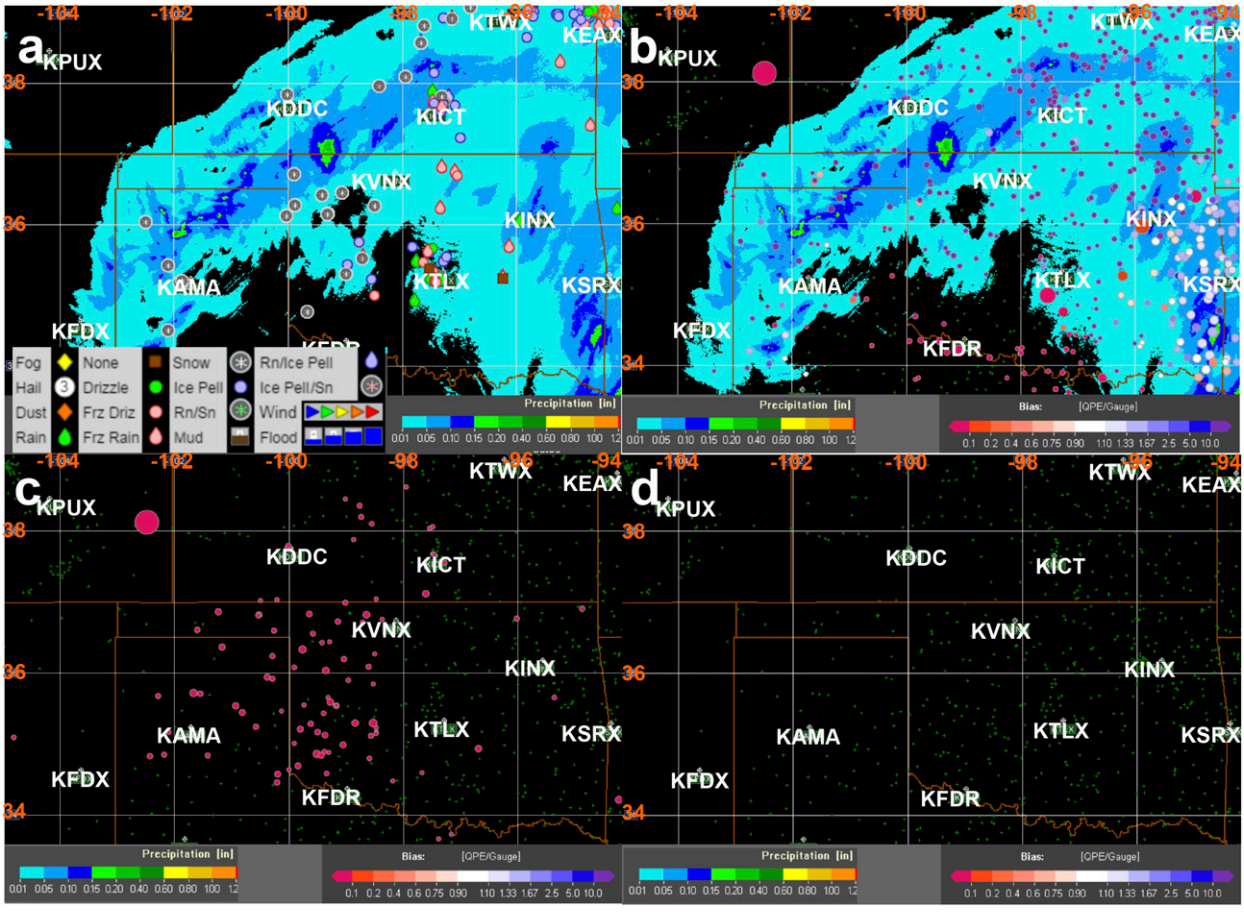

FIG. 11. The 1-h Q3rad rainfall accumulations ending at 2100 UTC 21 Dec 2013 overlaid with (a) mPING observations and (b) gauge ratio bubble plots, and 1-h Q3rad with gauge ratio bubble plots ending at 2100 UTC 22 Dec 2013 (c) before gauge QC and (d) after gauge QC. The blue (red) shaded colors of the bubbles indicated that the Q3rad overestimated (underestimated) when compared to the gauges. Radar locations are highlighted in white text.

caused by high outlier gauges and false precipitation gauges in the Q3MM product (Fig. 12a) were removed after the gauge QC is applied (Fig. 12b). The inclusion and subsequent removal of erroneous hourly gauge observations made a large contribution to the evaluation statistics and greatly affect the evaluation scores of mean bias ratio, RMSE, and CC (Figs. 12c,d). Another example from a precipitation event ending at 1200 UTC 14 October 2014 shows the influence of both erroneously high and low gauge observations in the generation of Q3MM over a widespread QPE event over the eastern United States (Fig. 13). Local maximum and minimum QPE values can be seen throughout the QPE field, especially those that were influenced by gauge observations that were erroneously high or a false zero value. An examination of a manually quality-controlled QPE product (not shown) generated by NWS River Forecast Centers confirmed that those outlier high gauges were incorrect. The removal of these gauges by the gauge QC algorithm slightly increased the mean bias but reduced the RMSE by 8.4 mm and increased the CC to 0.92 when compared to CoCoRaHS observations.

5. Summary and future work

The operational objective of MRMS is to create accurate, high-spatiotemporal-resolution QPE products. The MRMS system currently utilizes multiple sensors, including a network of overlapping radars and hourly rain gauge observations, to generate a suite of QPE products in real time. While hourly rain gauge observations are considered as ''ground truth'' and are used in the generation and validation of QPE products, there are a number of limitations and challenges that could result in erroneous observations. The gauge QC scheme developed for MRMS is a simple technique designed to ensure high-quality gauge observations are ingested into MRMS product development. This gauge QC scheme

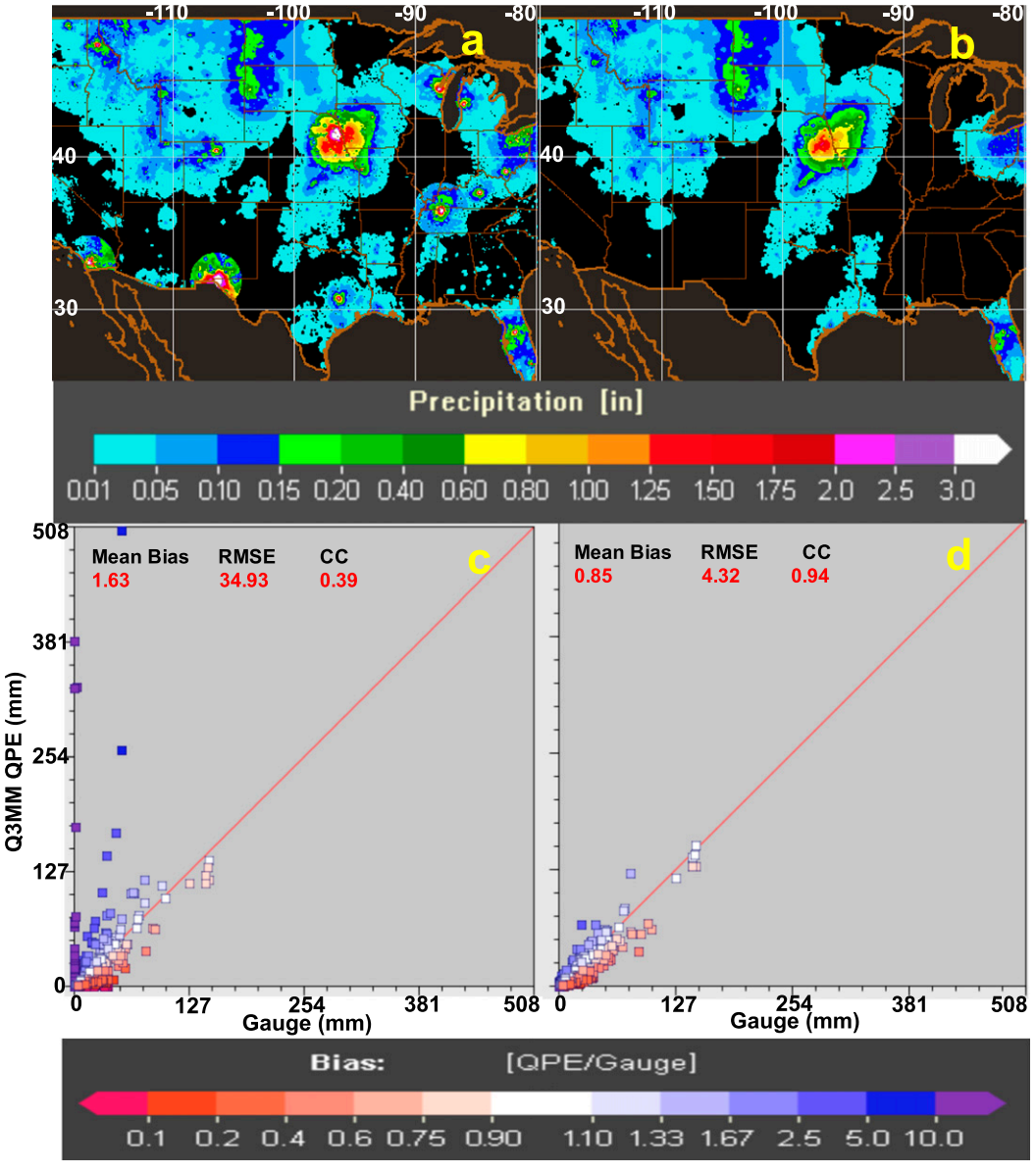

FIG. 12. Q3MM 24-h accumulated QPE ending at 1200 UTC 1 Oct 2014 (a) before gauge QC and (b) after gauge QC has been applied with the scatterplots of the 24-h Q3MM precipitation estimates vs CoCoRaHS gauge values (c) before gauge QC and (d) after gauge QC has been applied. The red line in the scatterplots is the line of perfect correlation. The mean bias ratio, RMSE, and CC values are also show above each scatterplot.

performs an analysis for each gauge observation using radar-based QPE, RAP model surface T_{wb}, and the RQI field to identify erroneous or suspect gauge values in a computationally efficient manner.

Approximately 80%–85% of the hourly gauge observations over the CONUS per month were identified for use in the MRMS system (i.e., those that were flagged as pass or unchecked). This varied seasonally because of the coverage of RQI changing based on the height of the freezing level and the areal extent of RAP model surface $T_{wb} \leq 0.00°C$, which is used to evaluate winter impacts on gauges. The remaining observations were flagged as erroneous by the MRMS system and were not used in QPE product generation. Approximately 12% of the gauge observation per month was flagged in the gauge QC scheme as being outside of the ±5-min window; thus, it was the greatest contributor to gauges that did not meet the criteria for use in MRMS. The next greatest contributors to gauges being removed from the system was those flagged as frozen, which varied greatly by season, and

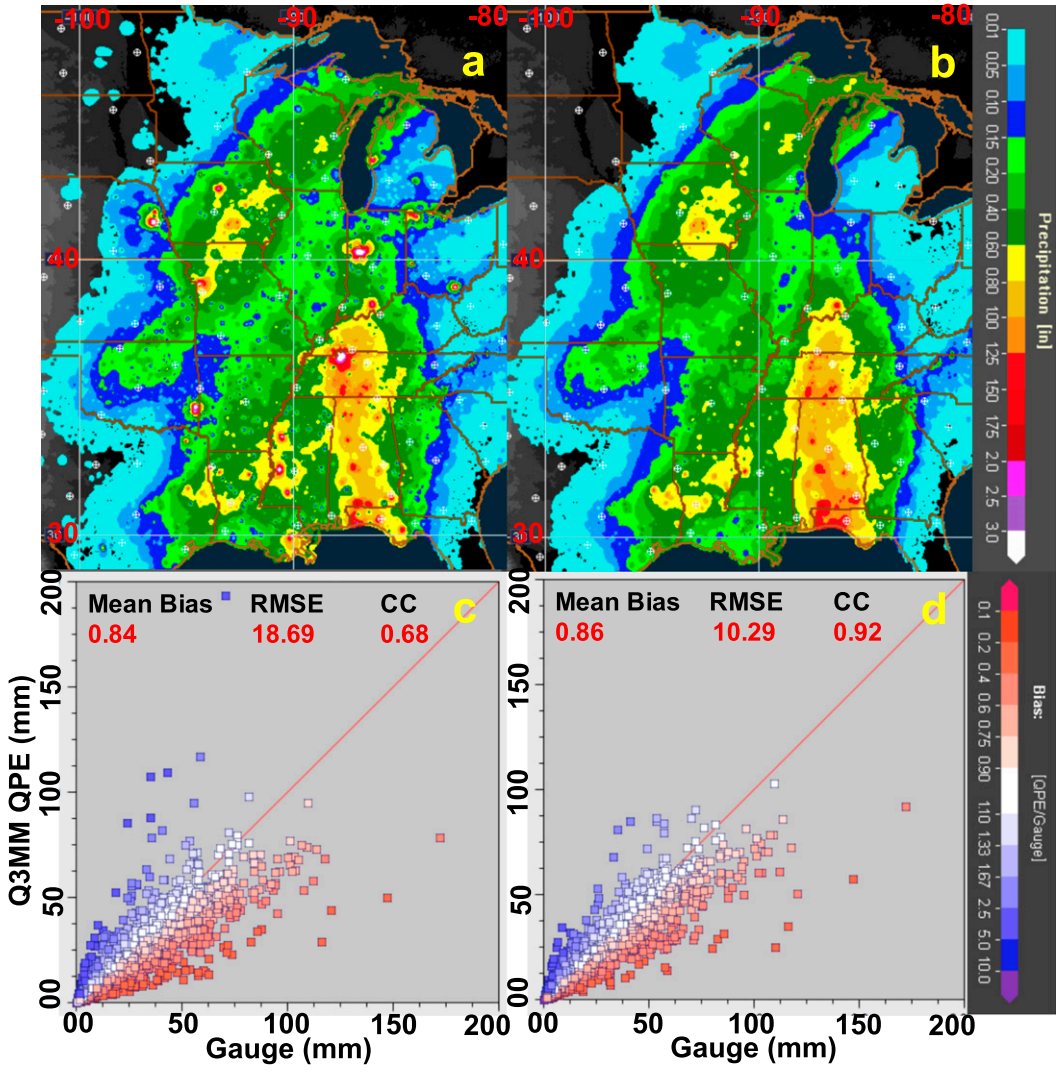

FIG. 13. As in Fig. 12, but for the 24-h period ending at 1200 UTC 14 Oct 2014.

those flagged as false zero, which had a local maximum during the spring months as a result of thawing snow. All other gauge QC flags accounted for approximately ≤1% of the gauge observations per hour.

The effectiveness of the gauge QC algorithm was evaluated by comparing Q3MM products generated using gauge datasets prior to and after being processed by the gauge QC algorithm. These Q3MM products were independent daily CoCoRaHS gauge observations. The Q3MM product using the gauges that passed the gauge QC scheme yielded more physically realistic precipitation distributions. By removing gauges that consistently reported zero precipitation observations, the number of nonphysical holes embedded in precipitation areas was minimized. The identification and removal of gauges reporting abnormally high amounts prevented unrealistic local maxima in the Q3MM product. Three statistical scores showed considerable improvements in the mean bias ratio, RMSE, and CC of the Q3MM product using only gauges that passed the QC logic versus a gauge dataset that did not undergo QC.

Further refinements to the MRMS gauge QC scheme are being investigated. The logic for removing gauges impacted by winter precipitation is broad and removes all observations except when G, R = 0.00 mm in environments conducive for winter precipitation. All nonzero gauge observations in these regions were considered as stuck or potentially impacted by winter precipitation because of the high uncertainty in the gauge observation and the radar-based QPE it would be compared to. Future work will look at identifying specific gauge instrumentation and gauge configurations to study the uncertainties of their observations through investigations with polarimetric radar observations,

wind field, and disdrometer observations. The future algorithm will assess three-dimensional reflectivity profiles and Q3rad values with respect to environmental conditions to determine if any precipitation accumulations from these specific situations are possible. The application of QC logic for hail impacts and the potential inclusion of a wind correction factor are also being investigated. Future studies will also assess the use of satellite-based QPE to evaluate gauges in areas of poor or no radar coverage.

Acknowledgments. Funding was provided by NOAA/Office of Oceanic and Atmospheric Research under NOAA–University of Oklahoma Cooperative Agreement NA11OAR4320072, U.S. Department of Commerce.

REFERENCES

Cifelli, R., N. Doesken, P. Kennedy, L. D. Carey, S. A. Rutledge, C. Gimmestad, and T. Depue, 2005: The Community Collaborative Rain, Hail, and Snow Network: Informal education for scientists and citizens. *Bull. Amer. Meteor. Soc.*, **86**, 1069–1077, doi:10.1175/BAMS-86-8-1069.

Daly, C., R. P. Neilson, and D. L. Phillips, 1994: A statistical–topographic model for mapping climatological precipitation over mountainous terrain. *J. Appl. Meteor.*, **33**, 140–158, doi:10.1175/1520-0450(1994)033<0140:ASTMFM>2.0.CO;2.

——, M. Halbleib, J. Smith, W. Gibson, M. Doggett, G. Taylor, J. Curtis, and P. Pasteris, 2008: Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *Int. J. Climatol.*, **28**, 2031–2064, doi:10.1002/joc.1688.

Diamond, H. J., and Coauthors, 2013: U.S. Climate Reference Network after one decade of operations: Status and assessment. *Bull. Amer. Meteor. Soc.*, **94**, 485–498, doi:10.1175/BAMS-D-12-00170.1.

Elmore, K. L., Z. L. Flamig, V. Lakshmanan, B. T. Kaney, V. Farmer, H. D. Reeves, and L. P. Rothfusz, 2014: MPING: Crowd-sourcing weather reports for research. *Bull. Amer. Meteor. Soc.*, **95**, 1335–1342, doi:10.1175/BAMS-D-13-00014.1.

Essery, C. I., and D. N. Wilcock, 1991: Variation in rainfall catch from standard U.K. meteorological office raingauges. *Hydrol. Sci. J.*, **36**, 23–24, doi:10.1080/02626669109492482.

Goodison, B. E., P. Y. T. Louie, and D. Yang, 1998: WMO solid precipitation intercomparison. Instruments and Observing Methods Rep. 67, WMO/TD 872, 212 pp. [Available online at https://www.wmo.int/pages/prog/www/IMOP/publications/IOM-67-solid-precip/WMOtd872.pdf.]

Groisman, P. Ya., and D. R. Legates, 1994: The accuracy of United States precipitation data. *Bull. Amer. Meteor. Soc.*, **75**, 215–227, doi:10.1175/1520-0477(1994)075<0215:TAOUSP>2.0.CO;2.

Habib, E., W. F. Krajewski, V. Nespor, and A. Kruger, 1999: Numerical simulation studies of rain gauge data correction due to wind effect. *J. Geophys. Res.*, **104**, 19 723–19 733, doi:10.1029/1999JD900228.

Helms, D., P. Miller, M. Barth, D. Starosta, B. Gordon, S. Schofield, F. Kelly, and S. Koch, 2009: Status update of the transition from research to operations of the Meteorological Assimilation Data Ingest System. *25th Conf. on Int. Interactive Information and Processing Systems*, Phoenix, AZ, Amer. Meteor. Soc., 5A.3. [Available online at https://ams.confex.com/ams/89annual/techprogram/paper_149883.htm.]

Kim, D., B. Nelson, and D. J. Seo, 2009: Characteristics of reprocessed Hydrometeorological Automated Data System (HADS) hourly precipitation data. *Wea. Forecasting*, **24**, 1287–1296, doi:10.1175/2009WAF2222227.1.

Kondragunta, C. R., and K. Shrestha, 2006: Automated real-time operational rain gauge quality-control tools in NWS hydrologic operations. *20th Conf. on Hydrology*, Boston, MA, Amer. Meteor. Soc., P2.4. [Available online at https://ams.confex.com/ams/Annual2006/techprogram/paper_102834.htm.]

Kursinski, A. L., and S. L. Mullen, 2008: Spatiotemporal variability of hourly precipitation over the eastern contiguous United States from stage IV multisensor analyses. *J. Hydrometeor.*, **9**, 3–21, doi:10.1175/2007JHM856.1.

Lanza, L., M. Leroy, J. van der Meuken, and M. Ondras, 2005: The WMO laboratory intercomparison of rainfall intensity gauges. Instruments and observing methods, Rep. 82, WMO/TD 1265, 8 pp. [Available online at https://www.wmo.int/pages/prog/www/IMOP/publications/IOM-82-TECO_2005/Papers/3(09)_Italy_Lanza.pdf.]

Larson, L. W., and E. L. Peck, 1974: Accuracy of precipitation measurements for hydrologic modeling. *Water Resour. Res.*, **10**, 857–863, doi:10.1029/WR010i004p00857.

Martinaitis, S. M., 2008: Effects of multi-sensor radar and rain gauge data on hydrologic modeling in relatively flat terrain. M. S. thesis, Florida State University, 99 pp. [Available online at http://diginole.lib.fsu.edu/islandora/object/fsu%3A180955.]

——, S. B. Cocks, Y. Qi, B. Kaney, J. Zhang, and K. Howard, 2015: Understanding winter precipitation impacts on rain gauge observations within an automated real-time system. *J. Hydrometeor.*, **16**, 2345–2363, doi:10.1175/JHM-D-15-0020.1.

Marzen, J., and H. E. Fuelberg, 2005: Developing a high resolution precipitation dataset for Florida hydrologic studies. *19th Conf. on Hydrology*, New Orleans, LA, Amer. Meteor. Soc., J9.2. [Available online at https://ams.confex.com/ams/Annual2005/techprogram/paper_83718.htm.]

McPherson, R. A., and Coauthors, 2007: Statewide monitoring of the mesoscale environment: A technical update on the Oklahoma Mesonet. *J. Atmos. Oceanic Technol.*, **24**, 301–321, doi:10.1175/JTECH1976.1.

Metcalfe, J. R., and B. E. Goodison, 1992: Automation of winter precipitation estimates: The Canadian experience. Instruments and observing methods, Rep. 49, WMO/TD 462, 81–85.

Nešpor, V., and B. Sevruk, 1999: Estimation of wind-induced error of rainfall gauge measurements using a numerical simulation. *J. Atmos. Oceanic Technol.*, **16**, 450–464, doi:10.1175/1520-0426(1999)016<0450:EOWIEO>2.0.CO;2.

Sevruk, B., 1989: Wind-induced measurement error for high intensity rains. *Proc. Int. Workshop on Precipitation Measurements*, St. Moritz, Switzerland, Swiss Federal Institute of Technology, ETH Zurich, 199–204.

——, 1996: Adjustment of tipping-bucket precipitation gauge measurements. *Atmos. Res.*, **42**, 237–246, doi:10.1016/0169-8095(95)00066-6.

——, 2005: Rainfall measurement: Gauges. *Encyclopedia of Hydrological Sciences*, M. G. Anderson, Ed., Wiley, 529–536.

——, M. Ondras, and B. Chvila, 2009: The WMO precipitation measurement intercomparisons. *Atmos. Res.*, **92**, 376–380, doi:10.1016/j.atmosres.2009.01.016.

Sieck, L. C., S. J. Burges, and M. Steiner, 2007: Challenges in obtaining reliable measurements of point rainfall. *Water Resour. Res.*, **43**, W01420, doi:10.1029/2005WR004519.

Steiner, M., J. A. Smith, S. J. Burges, C. V. Alonso, and R. W. Darden, 1999: Effect of bias adjustment and rain gauge data quality control on radar rainfall estimation. *Water Resour. Res.*, **35**, 2487–2503, doi:10.1029/1999WR900142.

Tollerud, E., R. Collander, Y. Lin, and A. Loughe, 2005: On the performance, impact, and liabilities of automated precipitation gage screening algorithms. *21st Conf. on Weather Analysis and Forecasting*, Washington, DC, Amer. Meteor. Soc., P1.42. [Available online at http://ams.confex.com/ams/pdfpapers/95173.pdf.]

Upton, G. J. G., and A. Rahimi, 2003: On-line detection of errors in tipping-bucket rain gauges. *J. Hydrol.*, **278**, 197–212, doi:10.1016/S0022-1694(03)00142-2.

Ware, E. C., 2005: Corrections to radar-estimated precipitation using observed rain gauge data. M.S. thesis, Dept. of Earth and Atmospheric Sciences, Cornell University, 87 pp. [Available online at https://ecommons.cornell.edu/handle/1813/2115.]

Wilson, J. W., and E. A. Brandes, 1979: Radar measurement of rainfall: A summary. *Bull. Amer. Meteor. Soc.*, **60**, 1048–1058, doi:10.1175/1520-0477(1979)060<1048:RMORS>2.0.CO;2.

Yang, D., B. E. Goodison, J. R. Metcalfe, V. S. Golubev, R. Bataes, T. Pangburn, and C. L. Hanson, 1998: Accuracy of NWS 8″ standard nonrecording precipitation gauge: Results and application of WMO intercomparison. *J. Atmos. Oceanic Technol.*, **15**, 54–68, doi:10.1175/1520-0426(1998)015<0054:AONSNP>2.0.CO;2.

Zhang, J., and Coauthors, 2011: National Mosaic and Multi-Sensor QPE (NMQ) system: Description, results, and future plans. *Bull. Amer. Meteor. Soc.*, **92**, 1321–1338, doi:10.1175/2011BAMS-D-11-00047.1.

——, Y. Qi, K. Howard, C. Langston, and B. Kaney, 2012: Radar Quality Index (RQI)—A combined measure of beam blockage and VPR effects in a national network. *Weather Radar and Hydrology*, R. J. Moore, S. J. Cole, and A. J. Illingworth, Eds., IAHS Publ. 351, 388–393.

——, and Coauthors, 2016: Multi-Radar Multi-Sensor (MRMS) quantitative precipitation estimation: Initial operating capabilities. *Bull. Amer. Meteor. Soc.*, **97**, 621–638, doi:10.1175/BAMS-D-14-00174.1.