# Advancements and Characteristics of Gauge Ingest and Quality Control within the Multi-Radar Multi-Sensor System

STEVEN M. MARTINAITIS,[a,b] STEPHEN B. COCKS,[a,b] MICHEAL J. SIMPSON,[a,b] ANDREW P. OSBORNE,[a,b]
SEBASTIAN S. HARKEMA,[c,d] HEATHER M. GRAMS,[b] JIAN ZHANG,[b] AND KENNETH W. HOWARD[b]

[a] Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, Norman, Oklahoma
[b] NOAA/OAR/National Severe Storms Laboratory, Norman, Oklahoma
[c] National Weather Center Research Experience for Undergraduates, Norman, Oklahoma
[d] Central Michigan University, Mt. Pleasant, Michigan

ABSTRACT: This study describes recent advancements in the Multi-Radar Multi-Sensor (MRMS) automated gauge ingest and quality control (QC) processes. A data latency analysis for the combined multiple gauge collection platforms provided guidance for a multiple-pass generation and delivery of gauge-based precipitation products. Various advancements to the gauge QC logic were evaluated over a 21-month period, resulting in an average of 86% of hourly gauge observations per hour being classified as useful. The fully automated QC logic was compared to manual human QC for a limited domain, which showed a >95% agreement in their QC reasoning categories. This study also includes an extensive evaluation of various characteristics related to the gauge observations ingested into the MRMS system. Duplicate observations between gauge collection platforms highlighted differences in site coordinates; moreover, errors in Automated Surface Observing System (ASOS) station site coordinates resulted in >79% of sites being located in a different MRMS 1-km grid cell. The ASOS coordinate analysis combined with examinations of other limitations regarding gauge observations highlight the need for robust and accurate metadata to further enhance the quality control of gauge data.

SIGNIFICANCE STATEMENT: This study examines an advanced quality control technique for the MRMS system and how it performs against manual quality control by forecasters, which showed >95% match for the reasoning of flagging a gauge; moreover, this study examines other characteristics pertaining to the ingest and quality control of automated gauge observations, including duplicate observations, errors in location, and the need for more robust metadata to improve hydrometeorological product verification and corrections.

KEYWORDS: Data quality control; Surface observations; Gauges

## 1. Introduction

Rain gauge observations provide critical information for various hydrometeorological needs, ranging from flood operations to water resource management. Precipitation accumulations recorded from rain gauges are considered as "ground truth" and are utilized for validation and correction of remotely sensed quantitative precipitation estimations (QPEs). Previous studies highlighted the positive impacts of combining rain gauge observations with radar-derived QPE to improve precipitation estimation accuracy (e.g., Seo and Breidenbach 2002; Smith and Krajewski 1991). The Multi-Radar Multi-Sensor (MRMS) system has relied upon hourly gauge observations to generate high spatiotemporal resolution gridded locally gauge-corrected radar QPE that has demonstrated improved statistical results over its radar-only counterpart (Zhang et al. 2016).

The utilization of gauge observations for QPE generation or validation requires the use of quality control (QC) algorithms to prevent erroneous observations from negatively influencing the validation or correction of gridded QPEs (e.g., Steiner et al.

1999). Previous research analyzed the QC of gauges through spatiotemporal checks (e.g., Tollerud et al. 2005; Kondragunta and Shrestha 2006; Kim et al. 2009). Other studies have outlined methodologies for identifying and removing erroneous gauge observations using gridded QPEs (e.g., Marzen and Fuelberg 2005; Chen and Xie 2008; Lewis et al. 2018). The MRMS system has employed its own gauge QC algorithm to identify potentially erroneous observations (Qi et al. 2016). Factors such as radar-derived QPE values, radar data adequacy, and surface wet-bulb temperatures to delineate between environments conducive of rainfall versus solid winter precipitation were utilized to help determine if hourly gauges observations were outlier or false values.

Gauge observations that fail to pass QC algorithms can be a result of physically based limitations inhibiting the retrieval of an accurate observation. Strong surface winds can result in precipitation undercatch due to turbulence (e.g., Wilson and Brandes 1979; Sevruk 1989; Sevruk et al. 1991; Yang et al. 1998; Habib et al. 1999). Previous research has noted challenges with the tipping process, notably in significant rain rates, that resulted in loss of liquid or double tipping (e.g., Parsons 1941). Blockage of the gauge orifice can result in the underestimation of or inability to record precipitation accumulations (e.g., Sevruk 2005; Sieck et al. 2007). Poor siting, malfunctions,

calibration errors, and other mechanical problems can also result in erroneous observations (e.g., Groisman and Legates 1994; Steiner et al. 1999; Kondragunta and Shrestha 2006; Sieck et al. 2007). Additional challenges have been documented when rain gauges were tasked to measure the liquid-equivalent of frozen precipitation. Previous studies by Goodison et al. (1998), Rasmussen et al. (2012), and Martinaitis et al. (2015) presented detailed discussions and examples of winter precipitation impacts on rain gauge observations.

Considerations beyond physical limitations must also be taken into account when utilizing gauge observations for QPE verification and correction, especially within a real-time system. The timing and data latency of gauge observations should be investigated to maximize the real-time ingest of data for product generation and availability. Qi et al. (2016) utilized a time window parameter to prevent the temporal mismatching of gauge observations to top-of-the-hour gridded QPEs for bias corrections during ongoing precipitation. This time window check accounted for the majority of the gauge observations that failed to pass the MRMS gauge QC scheme. The acquisition of multiple gauge networks necessitates the identification of duplicate observations; moreover, the accuracy of gauge metadata should also be verified prior to use.

The version 12.0 (v12.0) build of the MRMS system contained the most significant update to the system since the initial operating capability (IOC) in September 2014. This study details the gauge ingest and QC performance characteristics of the v12.0 build of the MRMS system over the CONUS domain, which includes southern Canada and northern Mexico. Included are the addition of new gauge data ingests and an expansion of the original gauge QC algorithm adapted from Qi et al. (2016) with proposed new QC flag designations. The results from the expanded fully automated gauge QC algorithm are presented in comparison to manual QC by National Weather Service (NWS) forecasters. The study also describes various characteristics of the gauges ingested into the MRMS system, including data latency, observation times, duplications of observations, impacts of gauge location errors, and gauge network metadata.

## 2. MRMS gauge ingest

The MRMS system IOC ingested gauges from the Hydrometeorological Automated Data System (HADS; Kim et al. 2009) network, primarily due to its availability at the National Centers for Environmental Prediction (NCEP) Central Operations (NCO). The HADS network ingest provided approximately 7000 hourly gauge observations per hour pre-QC at a 70-min latency for the CONUS, which allowed gauge-derived and gauge-corrected products to be available with a latency of 85 min. One significant change with the MRMS gauge ingest post-IOC was the inclusion of two additional gauge networks made available in the operational MRMS real-time feed as of May 2018. Rain gauge observations from the Meteorological Assimilation Data Ingest System (MADIS; Helms et al. 2019) network collection were downloaded from five different MADIS data
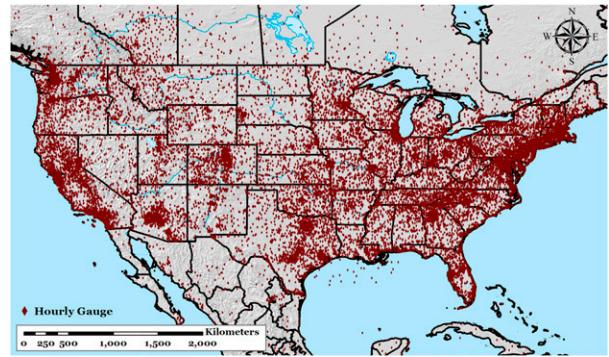


FIG. 1. Locations of available hourly gauge observations within the CONUS domain of the MRMS system for v12.0.

directories. This included a directory containing the Automated Surface Observing Systems (ASOS; https://www.weather.gov/asos/) stations. Gauge observations were reported at various accumulation periods and were decoded to hourly observations. The other gauge network was the Flood Control District of Maricopa County (FCDMC; e.g., Mascaro 2017). The FCDMC network consisted of 355 Automated Local Evaluation in Real Time (ALERT) rain gauges over Maricopa County and within neighboring counties across central Arizona. The addition of the MADIS and FCDMC gauge data feeds into the MRMS system increased the number of available unique hourly gauge observations for the CONUS domain to approximately 18 000–20 000 per hour pre-QC (Fig. 1).

The timing of MRMS gauge-based and gauge-corrected products is based upon optimal gauge data availability and product development latency. The availability of gauge observations within the HADS and MADIS networks increased at a steady rate for the first 65 min past valid time (Fig. 2). The stair-stepping increase of the data was primarily based on the update cycle of the MADIS files every 5 min. All gauges from the FCDMC became available within five minutes past the top of the hour. The prev12.0 capability of MRMS ingested gauges at 70 min past valid time, which contained approximately 91% of all observations. The v12.0 build of the MRMS system contained the capability of creating two passes of gauge-derived and
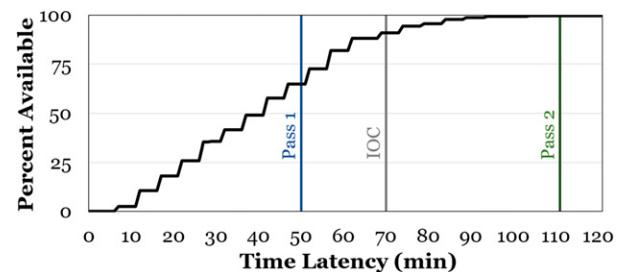


FIG. 2. The percent of available gauge observations from the gauge networks ingested by the MRMS system for the 0–120-min period following the top-of-the-hour observation time. The ingest times for the IOC version of MRMS along with the two-pass paradigm of MRMS v12.0 are denoted in the graph.

TABLE 1. List of proposed gauge QC flags for the MRMS system. Included are the description of each flag, if the gauge is retained for use in the MRMS system based on the flag designation, and the original MRMS gauge QC flags that were depicted in Qi et al. (2016).

| Proposed QC flag | Gauge QC description | Retain in MRMS | MRMS IOC QC flag |
|---|---|---|---|
| −20 | Outside time window of $>\pm 7$ min | No | −2 |
| −10 | Conditional pass—Zero observation (w/sources present) | Yes | −1 |
| −11 | Conditional pass—Zero observation (w/sources missing) | Yes | −1 |
| −12 | Conditional pass—Nonzero observation (w/sources present) | Yes | −1 |
| −13 | Conditional pass—Nonzero observation (w/sources missing) | Yes | −1 |
| 0 | Pass—Zero value | Yes | 0 |
| 1 | Pass—Nonzero value | Yes | 0 |
| 10 | False zero observation (rain only) | No | 1 |
| 20 | False nonzero observation (rain only) | No | 2 |
| 30 | Outlier observation—High value (rain only) | No | 3 |
| 31 | Outlier observation—High value (snow only) | No | 3 |
| 40 | Outlier observation—Low value (rain only) | No | 4 |
| 50 | Winter impacts—Zero observation | No | 5 |
| 51 | Winter impacts—Nonzero observation | No | 5 |
| 60 | Suspect observation—Nonzero observation (w/sources present) | No | 6 |
| 61 | Suspect observation—Nonzero observation (w/sources missing) | No | 6 |

gauge-corrected products (hereinafter denoted as Pass 1 and Pass 2). Pass 1 allowed for gauge-based products to be available earlier than in previous builds. The Pass 1 gauge ingest occurred at 50 min, which captured approximately 65% of gauge observations and allowed products to be ready prior to 60 min past valid time. The Pass 2 ingest occurred at 111 min to incorporate approximately 99.8% of gauge observations available to the MRMS system. Pass 2 products were available at a latency of 120 min.

## 3. Gauge quality control

### a. Logic advancements

The MRMS system employed a fully automated gauge QC algorithm to remove evidently erroneous gauge observations

from the generation or correction of various precipitation products. Qi et al. (2016) describes the automated gauge QC process as part of the IOC implemented in September 2014. The decision tree within the IOC-based MRMS gauge QC algorithm contained a total of 28 nodes that allowed gauge observation to be classified by one of nine available QC flags (Qi et al. 2016).

New advancements to the MRMS gauge QC algorithm were applied throughout the entire automated process. Two significant changes involved the precipitation sources utilized in the QC algorithm. The radar QPE used in this study was the MRMS dual-polarization synthetic QPE scheme (Cocks et al. 2019; Wang et al. 2019; Zhang et al. 2020) with an evaporation correction algorithm (Martinaitis et al. 2018), hereinafter denoted as $Q_{DP}$. This replaced the reflectivity-based
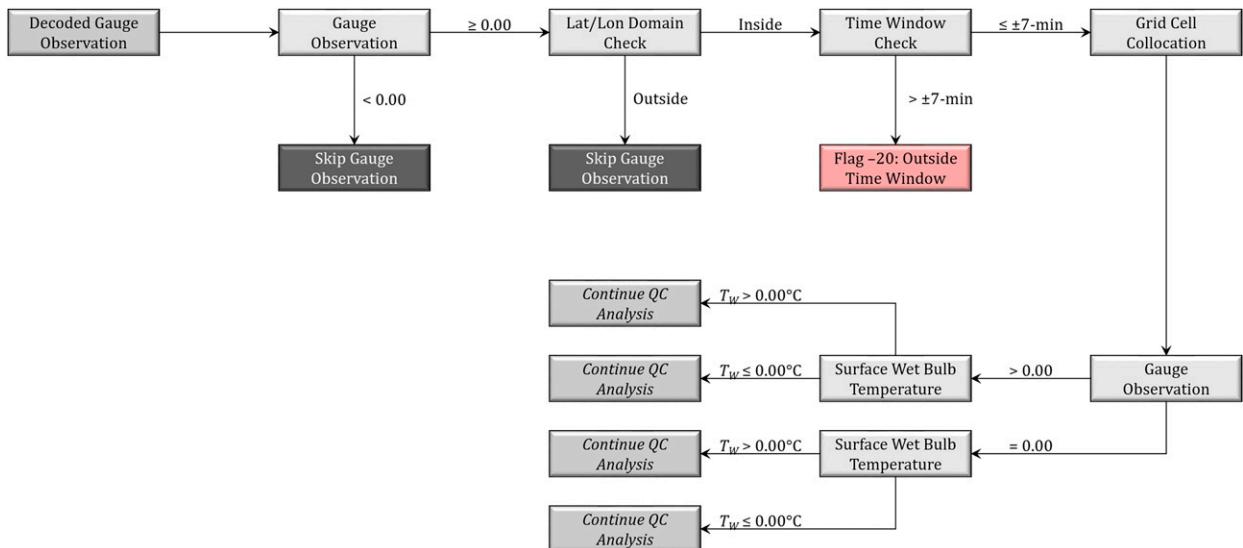


FIG. 3. Initial section of the MRMS gauge QC decision tree.
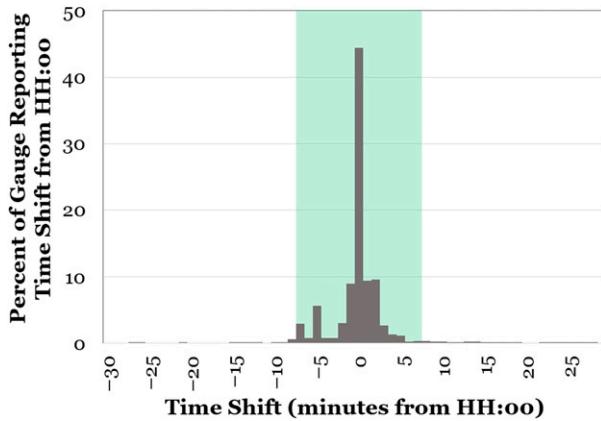
FIG. 4. Distribution of the time shift from the top of the hour for hourly gauge observations ingested by the MRMS system. The green highlighted area represents the time window considered by the MRMS gauge QC algorithm.

radar-only -QPE (Zhang et al. 2016) used in the IOC version of the gauge QC algorithm. Version 3 of the 1-h High Resolution Rapid Refresh (HRRR; Benjamin et al. 2016) model quantitative precipitation forecast (QPF) was included to improve the QC of gauges located outside of adequate radar coverage. Significant biases in the HRRR QPF precluded a direct value comparison (e.g., Martinaitis et al. 2020), yet it provided a means to delineate areas receiving precipitation versus no precipitation.

One proposed change tested within the new QC algorithm was increasing the number of gauge QC flags to 16 unique designations (Table 1). The QC flag structure allowed for more detailed characterizations of how a gauge observation passed or failed each hour. The addition of the HRRR QPFs along with new QC flags and other QC algorithm updates described in the following subsections resulted in a more complex MRMS gauge QC decision tree with 103 total nodes.

1) INITIAL GAUGE QC CHECKS AND PAIRINGS

The updated gauge QC logic retained the initial checks for missing gauge values (denoted as a negative gauge observation), ensuring the gauge resides within the declared MRMS domain based on latitude and longitude bounds, and significant deviations of observation times to prevent incorrect bias adjustments during ongoing precipitation events (Fig. 3). The acceptable time window criteria utilized in this study was increased to $\pm 7$ min from the top of the hour, which allowed for the inclusion of the ASOS gauge network that generally reported at the time HH:53 (e.g., the 1300 UTC observation reported at 1253 UTC). Increasing the acceptable time window by 2 min from the IOC-based $\pm 5$-min window allowed for 4.4% more gauges to be available for further QC inspection. The $\pm 7$-min time window encompass 92.3% of all ingested gauge observations (Fig. 4). Gauge observations with a time stamp outside of the $\pm 7$-min time window were flagged for nonuse by the MRMS system.

All remaining gauge observations were spatially matched to a MRMS grid cell. The IOC MRMS gauge QC logic employed a
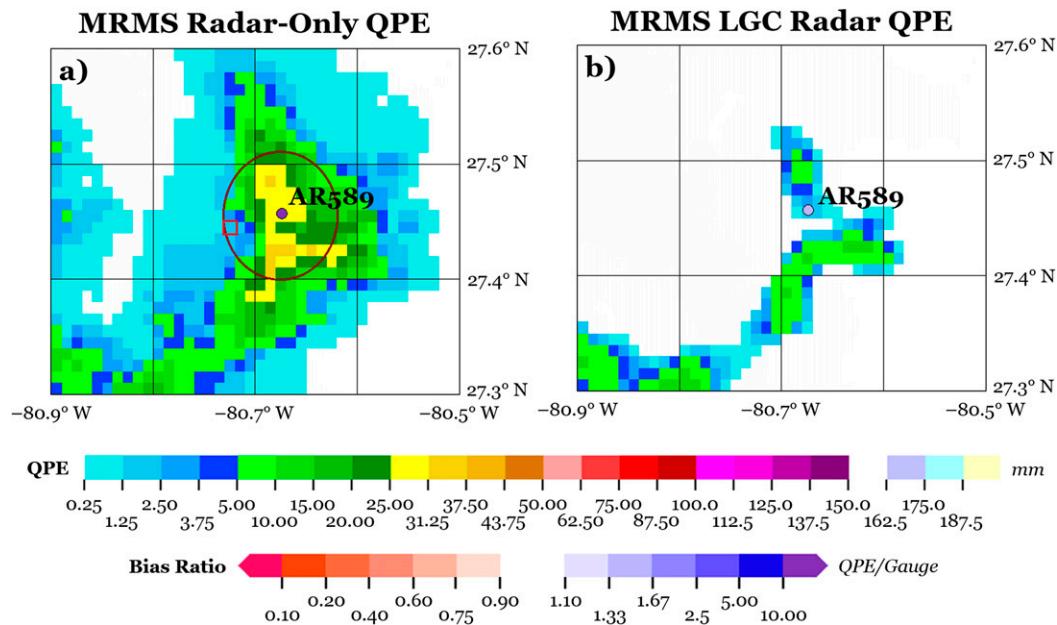


FIG. 5. One-hour accumulation of (a) the MRMS radar-only QPE and (b) the MRMS locally gauge corrected radar QPE that was influenced by the gauge AR589 northeast of Okeechobee, Florida, for the 1-h period ending 2200 UTC 27 Sep 2017. The gauge AR589 is represented by a bubble plot that is color filled with its bias ratio compared to its collocated grid cell of MRMS-generated QPE. Depicted around AR589 in (a) is the 5-km search radius (dark red circle) to best match the gauge with a MRMS radar-only QPE value and the MRMS grid cell (red square) that best matched the gauge AR589.
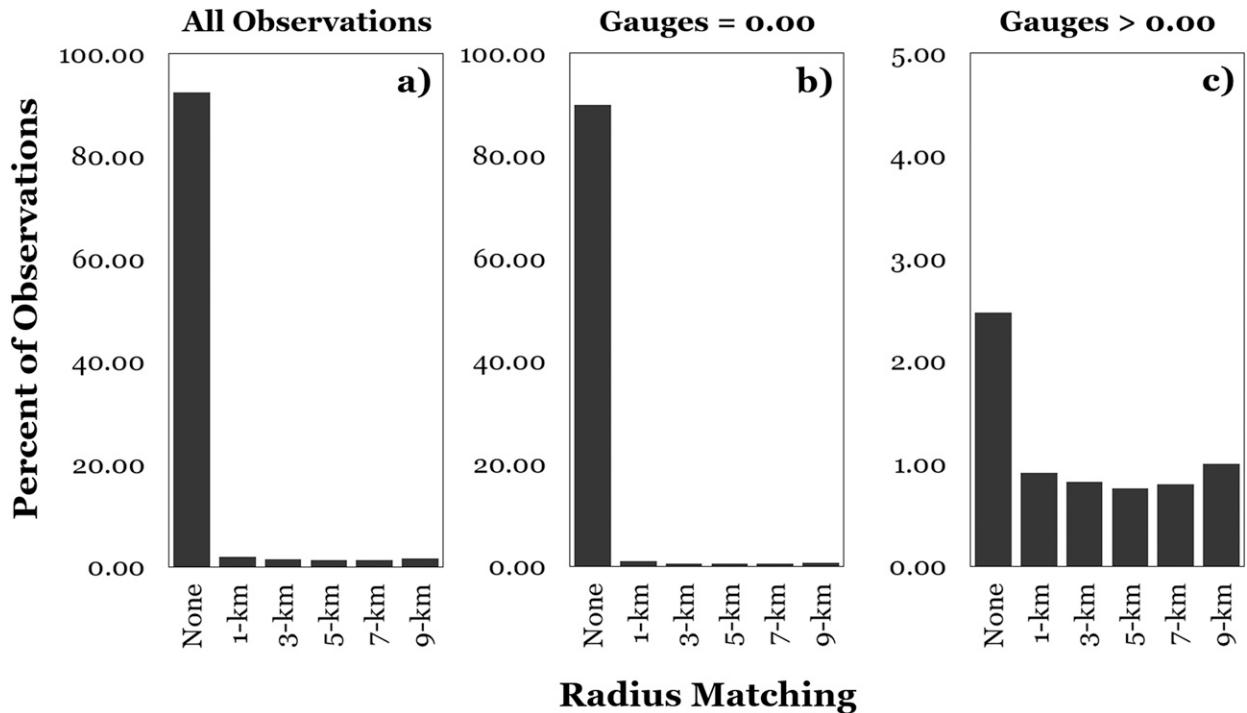
FIG. 6. Radius-matching of gauge observations to the MRMS $Q_{DP}$ product for (a) all observations, (b) for gauge values that were zero, and (c) gauge values that were nonzero for the period 1 Aug 2018–29 Feb 2019. The percent of observations were broken down by those that the best-matching MRMS grid cell did not radially match beyond its collocated grid cell, and radial matching up to 1, 3, 5, 7, and 9 km, respectively.

5-km radius around the grid cell containing the gauge observation to pair the gauge observation to the best-matching gridded hourly radar QPE value (Qi et al. 2016); however, the best-matching radius check presented challenges with erroneous gauge observations and the subsequent interpolation of the erroneous values into a bias-correction radar QPE. The example convective rainfall event northeast of Okeechobee, Florida, demonstrated the best-matching search radius challenges. Convective rainfall produced localized hourly accumulations of 25–37 mm ending 2200 UTC 27 September 2017 (Fig. 5a). The gauge AR589 for the same 1-h accumulation period recorded 1.27 mm, yet the MRMS hourly radar QPE was 28.7 mm at its collocated grid cell. The 5-km radius for spatial matching paired the gauge AR589 with a grid cell having a best-matched value of 2.03 mm at the edge of the convective rainfall footprint, and the gauge was subsequently passed. The local gauge-correction scheme in MRMS (Zhang et al. 2016) would then calculate the radar–gauge difference at the grid cell containing the gauge and interpolate that difference using an inverse distance weighting scheme. The interpolation of the significant radar–gauge difference at the collocated grid cell falsely reduced the spatial coverage and magnitude of the accumulated precipitation for this event (Fig. 5b).

A sensitivity analysis was conducted to determine the appropriate spatial matching search radius to still account for some horizontal advection of hydrometeors. Hourly gauge observations ingested by the MRMS system from 1 August 2018 to 29 February 2019 were compared against the MRMS $Q_{DP}$ product at each collocated grid cell and over various radii from 1 to 9 km (Fig. 6). Approximately 92.4% of all gauge observations matched with the $Q_{DP}$ value at the grid cell collocated with the gauge. All radius-based matching had <2% of the observations each for each radius range with 1 km recording the greatest percentage at 1.91%. Half of the nonzero gauge values matched with their collocated grid cell or grid cells within 1 km of the gauge site. The percent of gauge observation best matched with $Q_{DP}$ grid cells decreased out to the 5-km radius and then increased up to the 9-km radius. This was likely attributed to the greater number of gridded $Q_{DP}$ estimations that the gauge can be paired against. The proposed gauge QC scheme reduced the radius configuration from 5 to 1 km for spatial matching. Application of a 1-km radial check for best matching the gauge to the gridded QPE would have properly resulted in the failure of the gauge AR589.

The gauge QC decision tree separated the matched observations into four categories based on whether the gauge recorded any liquid or liquid-equivalent precipitation and the surface wet-bulb temperature $T_W$ as defined by the HRRR model (Fig. 3). A surface $T_W \le 0.00°C$ was used to define winter precipitation environments (Martinaitis et al. 2015). Utilization of the surface $T_W$ allowed for the identification of winter precipitation in above-freezing ambient temperatures and nonsaturated relative humidity values (e.g., Matsuo and Sasyo 1981). The separation
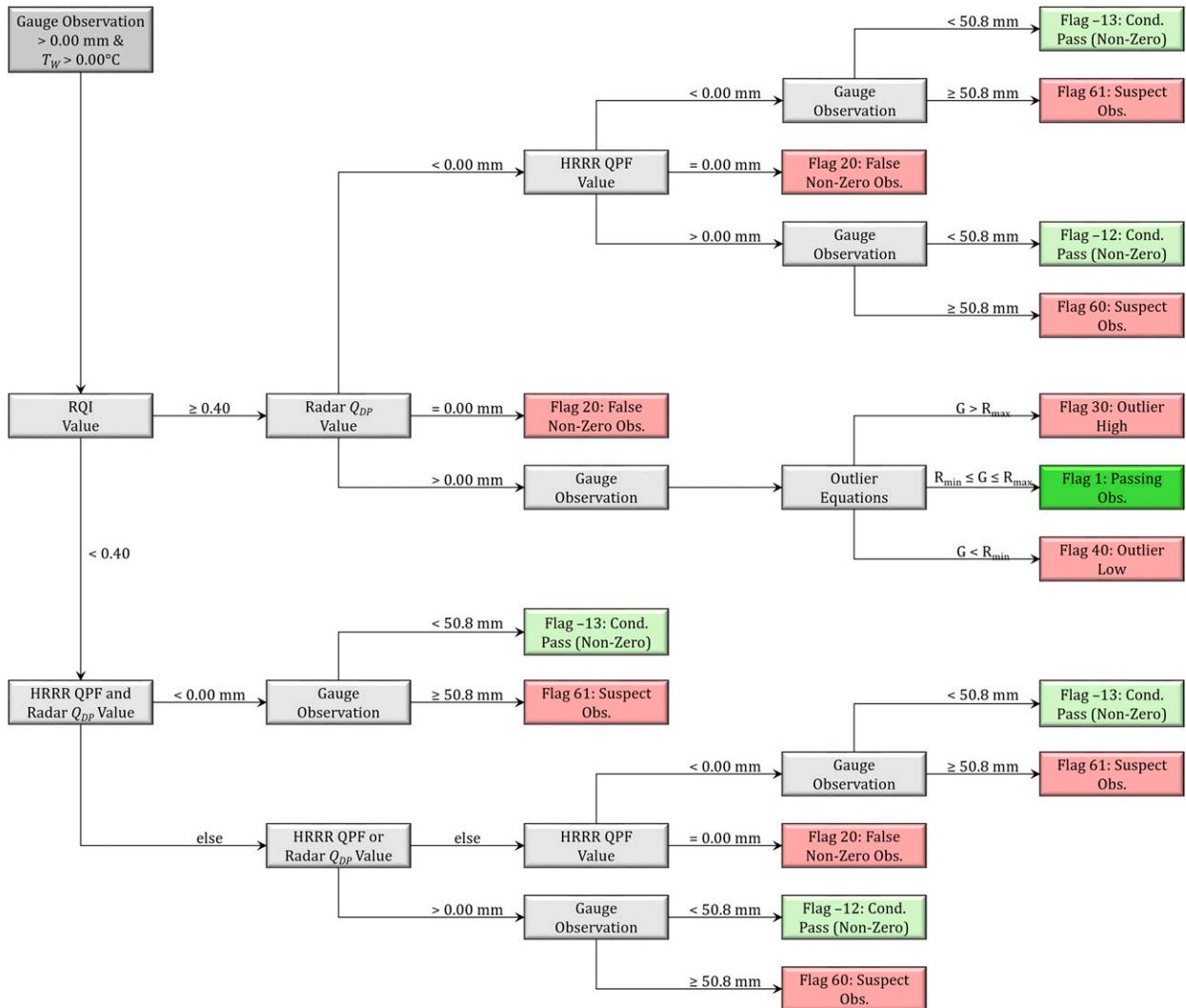
FIG. 7. MRMS gauge QC decision tree for hourly gauge observations reporting nonzero accumulations in environments characterized by $T_W > 0.00°C$.

of gauge observations based on their values and on the surface $T_W$ permitted the use of unique QC logic and QC flags for each category.

2) NONZERO GAUGE VALUES IN NONWINTER ENVIRONMENTS

The QC of hourly nonzero gauge observations in nonwinter environments began with the comparison of the gauge locations with respect to radar coverage via the Radar Quality Index (RQI) product (Fig. 7) defined by Zhang et al. (2012) and updated by Martinaitis et al. (2020). RQI quantifies the adequacy of radar coverage based upon radar beam blockage and radar beam height characteristics related to the freezing-level height and other reference height variables. RQI values ranged from 1.00 to 0.00, where RQI = 1.00 represented a completely unblocked radar beam below the melting layer. A decision point in the QC methodology occurred at RQI = 0.40

where it has been shown that radar coverage begins to significantly overshoot shallow precipitation features.

Gauge sites that resided in regions defined by RQI ≥ 0.40 were compared directly to the $Q_{DP}$ product (Fig. 7). Nonzero gauge observations failed QC when the collocated region had no radar-detected accumulated precipitation. The gauge underwent outlier checks if there was accumulated precipitation with $Q_{DP}$. The gauge value was compared against an expected maximum value $R_{max}$ and minimum value $R_{min}$ based on the best matched $Q_{DP}$ hourly accumulation and RQI via

$$R_{\max} = (9.00 - 6.50x)Q_{DP}^{(0.15x+0.76)} + (0.43x + 0.86), \quad (1)$$

and

$$R_{\min} = (0.26x - 0.06)Q_{DP}^{(1.50-0.30x)}, \quad (2)$$
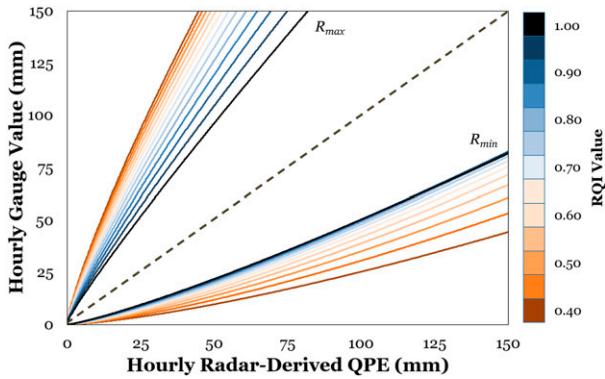
FIG. 8. Dynamic outlier checks based on the equations for $R_{max}$ and $R_{min}$ for hourly nonzero gauge observations located in adequate radar coverage areas based on the RQI range 0.40–1.00 and in environments characterized by $T_W > 0.00°C$.

where $x$ is the RQI value at the grid cell containing the best spatially matched $Q_{DP}$ value. The addition of RQI to the outlier check equations allowed for threshold $R_{max}$ and $R_{min}$ values to be based dynamically on the uncertainty in radar quality (Fig. 8) and were subjectively derived using a control dataset of approximately 95 000 nonzero gauge observations across the RQI spectrum. Gauge observations beyond the $R_{max}$ and $R_{min}$ values were flagged as outlier high and outlier low observations, respectively.

Gauge sites that resided in areas where $Q_{DP}$ was missing or where RQI < 0.40 due to the radar coverage overshooting precipitation features at farther distances utilized a different logic set that included precipitation data from the 1-h HRRR model QPF (Fig. 7). Radar-based $Q_{DP}$ was included in the nonzero comparison given the potential for the radar to detect precipitation when RQI < 0.40. $Q_{DP}$ was not included in the identification of false nonzero observations when RQI < 0.40
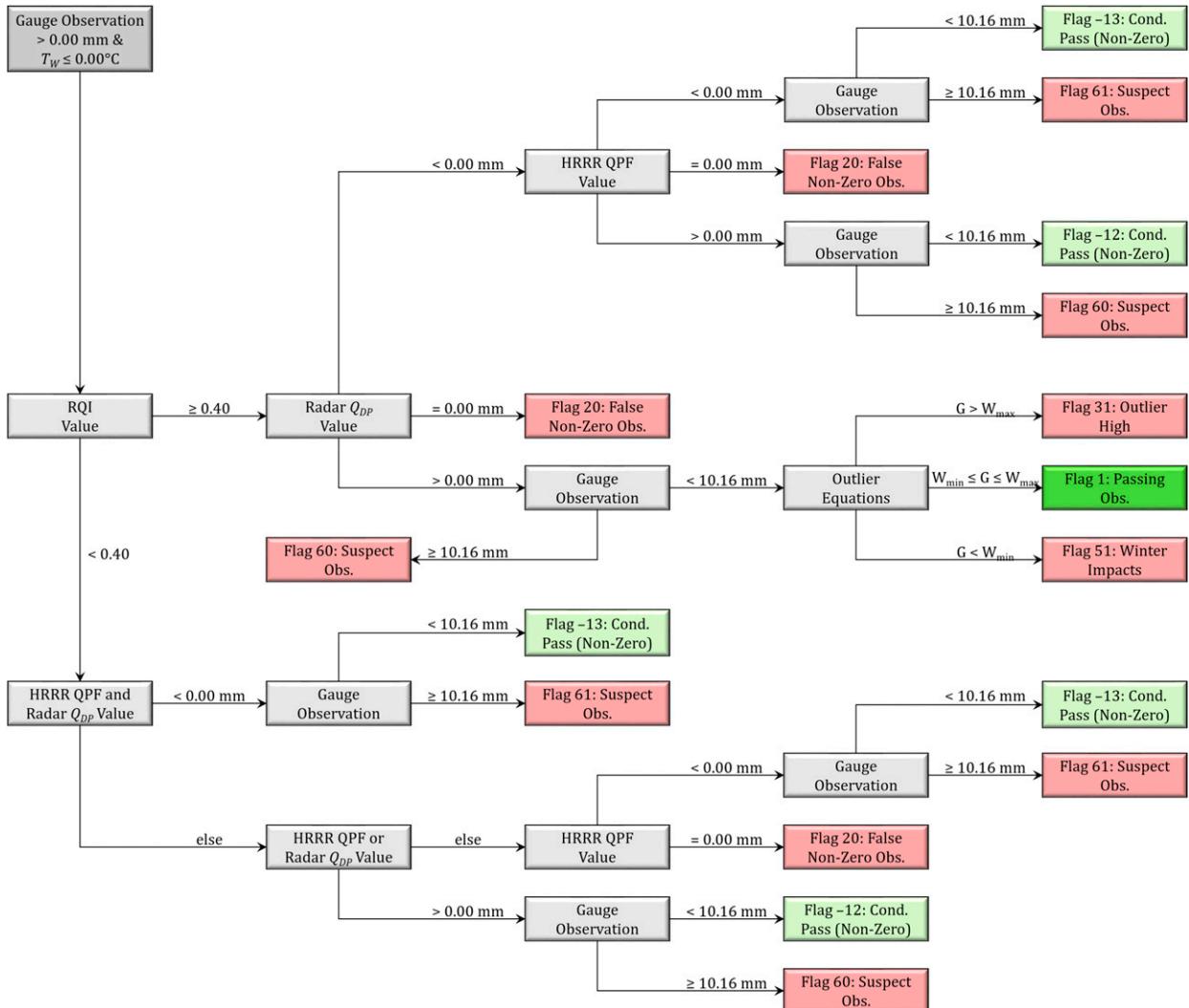


FIG. 9. As in Fig. 7, but for hourly gauge observations reporting nonzero accumulations in environments characterized by $T_W \leq 0.00°C$.
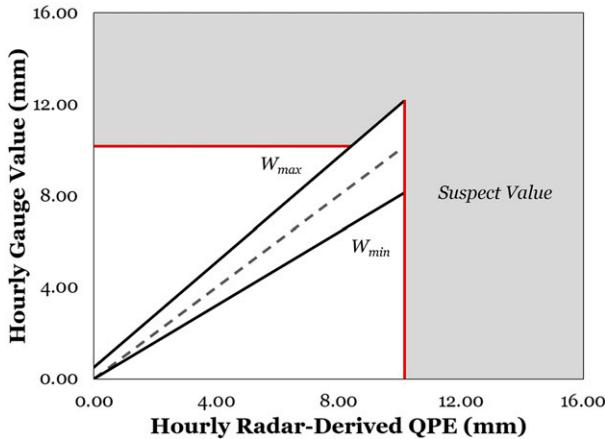
FIG. 10. Outlier checks based on the equations for $W_{max}$ and $W_{min}$ for hourly nonzero gauge observations located in adequate radar coverage areas based on the RQI range 0.40–1.00 and in environments characterized by $T_W \leq 0.00°$C. Also depicted are the suspect value constraints that influence the outlier check per the gauge QC logic of Fig. 9.

given the overshooting of shallow precipitation features, which could erroneously fail a nonzero gauge observations. The overall coverage of the HRRR allowed for a more accurate representation of where precipitation did not occur in the areas of reduced radar coverage.

A best matched nonzero value from the 1-h HRRR model QPF or $Q_{DP}$ in regions where RQI < 0.40 resulted in a suspect observation check. The suspect check was applied given the nonuse of direct value comparisons of the gauge to $Q_{DP}$ and the HRRR QPF due to uncertainties in $Q_{DP}$ when RQI < 0.40

and biases in the HRRR model. Hourly observations $\geq 50.8$ mm were flagged as a suspect value, given the rarity of hourly rainfall accumulations $\geq 50.8$ mm in data-sparse regions. Gauge values < 50.8 mm were conditionally passed given the challenges of making direct numerical comparisons between the gridded product and the gauge. Different conditional flag and suspect flag designations were utilized based on the availability of comparative data.

### 3) NONZERO GAUGE VALUES IN WINTER ENVIRONMENTS

Martinaitis et al. (2015) identified various challenges with automated hourly gauge observations during winter precipitation events and the widespread occurrence of winter impacts on the liquid-equivalent values. Results from the study showed that utilizing gauges in winter precipitation regimes without a stringent QC scheme generated worse results for locally bias-corrected QPEs than a radar-only approach in winter precipitation. The study showed it was more advantageous to not bias correct radar-derived QPE with gauges; however, biases in radar-generated QPE still exist and would still need to be corrected for. The updated gauge QC algorithm in this study retained gauge observations in winter environments for QPE correction and verification in MRMS through stricter QC measures than the measures employed for gauges in nonwinter environments.

The logic for nonzero hourly gauge accumulations in winter environments (i.e., when model surface $T_W \leq 0.00°$C) was identical to that of the nonwinter environment logic apart from changes relating to defining outlier values and the suspect check threshold (Fig. 9). The suspect check threshold value was set to 10.16 mm for hourly liquid-equivalent gauge values when $T_W \leq 0.00°$C. The suspect value check was applied regardless
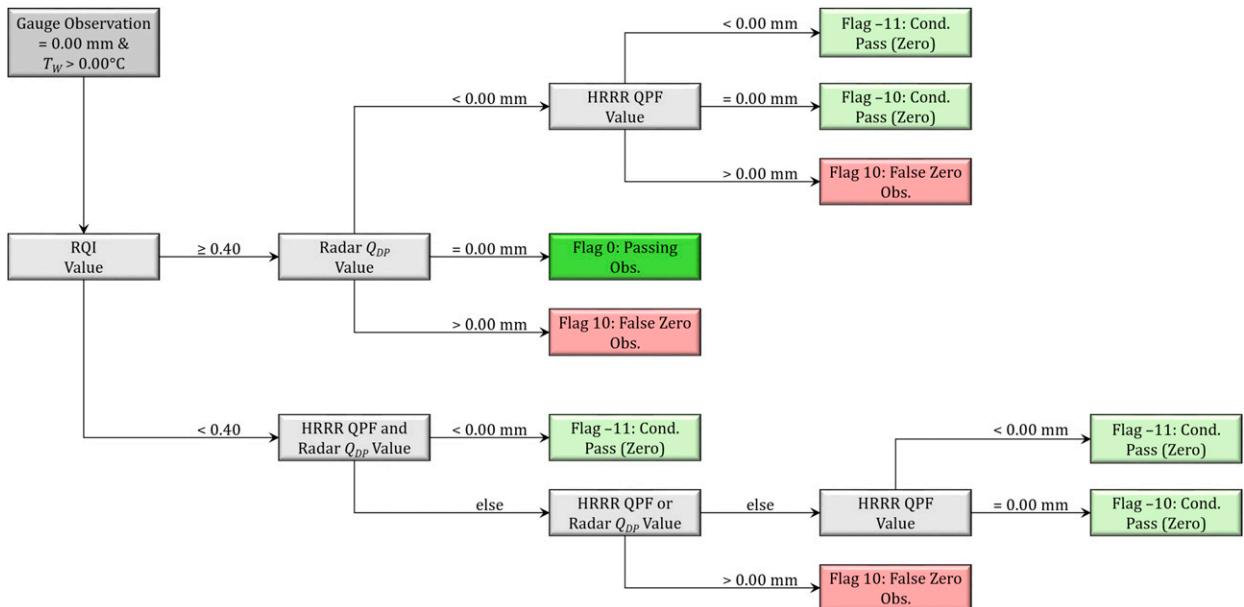


FIG. 11. As in Fig. 7, but for hourly gauge observations reporting zero accumulations in environments characterized by $T_W > 0.00°$C.
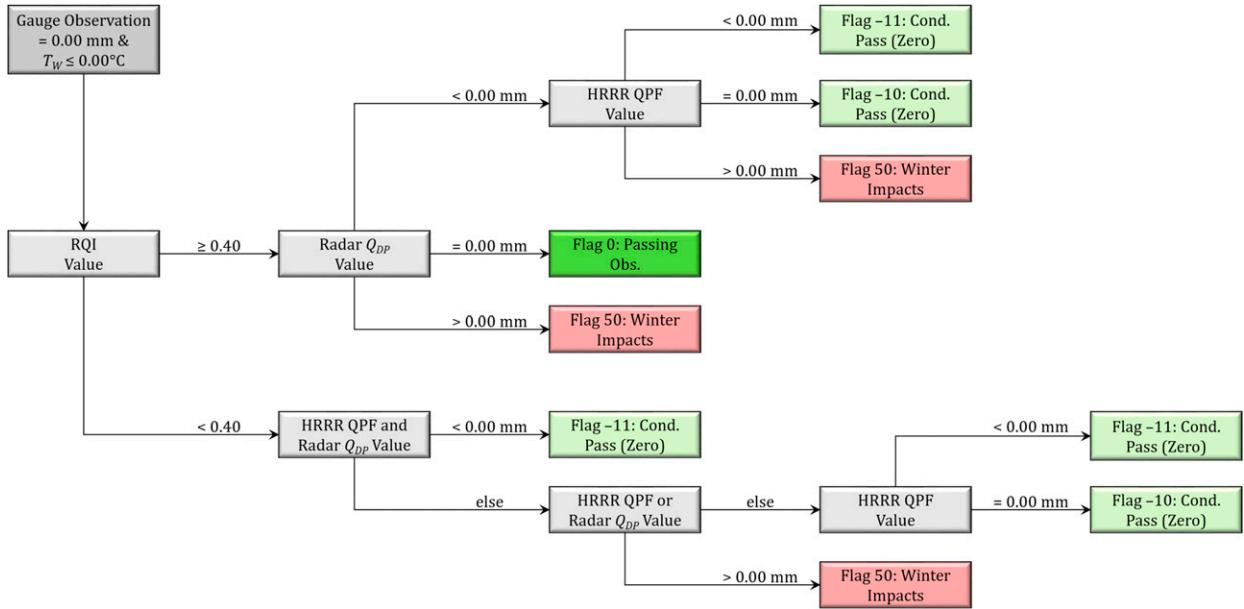
FIG. 12. As in Fig. 7, but for hourly gauge observations reporting zero accumulations in environments characterized by $T_W \leq 0.00°C$.

of radar coverage or radar data availability to remove erroneously high values, since hourly liquid accumulations exceeding 10.16 mm are rare (e.g., Kirkham et al. 2019). Gauge observations in regions where RQI $\geq$ 0.40 were analyzed with outlier equations using an expected winter environment maximum value $W_{max}$ and minimum value $W_{min}$ based on best matched $Q_{DP}$ hourly accumulation and were defined as the following:

$$W_{max} = (1.15 \times Q_{DP}) + 0.5, \qquad (3)$$

and

$$W_{min} = 0.8 \times Q_{DP}. \qquad (4)$$

These equations were more aggressive in removing lower gauge values compared to $Q_{DP}$ (Fig. 10) to account for findings by Martinaitis et al. (2015) that characterized the impacts of winter precipitation on gauges that produced a nonzero value. The restrictive nature of Eqs. (3) and (4) limit the passing rate of nonzero gauge observations in winter precipitation to approximately 39% based on analysis of 180 000 hourly gauge observations. Gauge observations found to be less than $W_{min}$ were designated with a winter impacts flag for nonzero observation.

#### 4) ZERO GAUGE VALUE IN NONWINTER ENVIRONMENTS

Logic for gauge observations that record no hourly precipitation depended upon the availability of radar-derived QPE and HRRR model QPF to determine the quality of the observation (Fig. 11). Matching the zero gauge value to the gridded $Q_{DP}$ data yielded a direct pass or fail decision when radar coverage was deemed to be adequate (i.e., RQI $\geq$ 0.40) and available. Instances of poor radar coverage (RQI < 0.40)

or when radar data was not available relied upon the HRRR model QPF as the QC check. A zero gauge observation would fail if the HRRR model QPF or the $Q_{DP}$ product had a best-matched nonzero value, while it would conditionally pass if the HRRR model QPF was zero or missing.

#### 5) ZERO GAUGE VALUE IN WINTER ENVIRONMENTS

The logic for gauge observations that record no hourly precipitation in winter environments was identical to that of the nonwinter environment logic except for how the gauge was flagged for nonuse in the MRMS system (Fig. 12). Gauges that were best matched with a grid cell containing $Q_{DP}$ or 1-h HRRR model QPF values > 0 were flagged as having winter weather impacts instead of a false zero observation. This allowed for identification of gauge observations that could have been compromised by winter precipitation that either blocked the gauge orifice or accumulated on the gauge orifice walls and prevented the recording of a nonzero observation (Goodison et al. 1998).

### b. Analysis of gauge QC

A 21-month dataset of hourly gauge observations ingested by the MRMS system from 0000 UTC 1 November 2017 to

TABLE 2. The average number and percent of gauge observations per hour that were passed, conditionally passed, and failed by the MRMS gauge QC algorithm presented in this study.

| Gauge pass/fail group | Number of observations per hour | Percent of observations per hour |
|---|---|---|
| Passed QC | 14 435 | 78.89% |
| Conditionally passed QC | 1416 | 7.74% |
| Failed QC | 2447 | 13.37% |
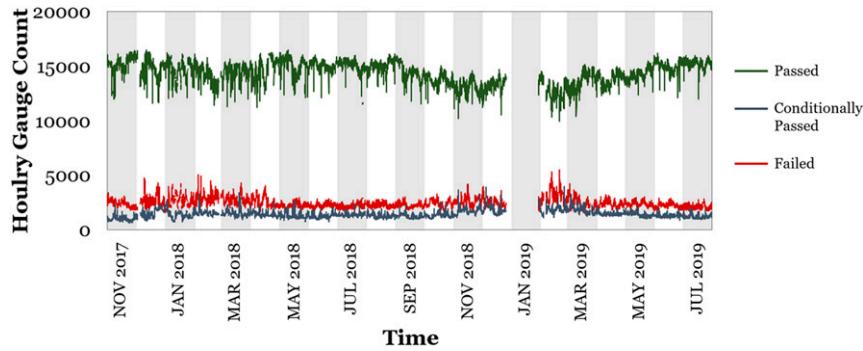
FIG. 13. Number of gauge observations per hour that were passed (green line), conditionally passed (blue line), and failed (red line) by the MRMS gauge QC algorithm for the 21-month study period from November 2017 to July 2019.

2300 UTC 31 July 2019 were processed by the new gauge QC logic. Hours that had at least 75% of total available hourly observations present were included in the analysis. This encompassed 13 616 h that totaled approximately $2.492 \times 10^8$ hourly gauge observations.

The 13 616 h evaluated contained an average of 18 298 hourly gauge observations per hour. The presented gauge QC algorithm would pass approximately 14 435 observations per hour and conditionally pass approximately 1416 observations per hour (Table 2). The combination of these two QC classifications represented 86.63% of the observations retained for use within the MRMS system. The remaining 13.37% (approximately 2447) observations per hour were flagged as failing the gauge QC algorithm and not utilized by the MRMS system.

Seasonal variabilities existed in the proportion of gauges that were passed, conditionally passed, and failed (Fig. 13). Quantified seasonal statistics (Table 3) showed the December–February (DJF) cool season period had a greater percentage of gauges that failed QC (14.9%) than in the June–August (JJA) warm season period (11.8%). The proportion of hourly gauge observations retained for use (i.e., observations flagged as passed or conditionally passed) had variability as well. Observations during the DJF months had a smaller percentage of passing flags (76.2%) than the JJA months (81.4%), while more gauges were conditionally passed in the DJF months (8.9%) than the JJA months (6.8%). This was a result of reduced RQI values in the DJF months from lower freezing-level heights, a prominent factor in calculating RQI

(Zhang et al. 2012; Martinaitis et al. 2020). There was also greater variability in the average gauge counts for the overall pass/fail conditions during the DJF months with standard deviation values that were more than double that during the JJA months.

Percent contributions of the individual QC flags highlighted the seasonal variabilities throughout the different pass/fail conditions derived from the gauge QC scheme. Gauges that were passed were predominantly labeled with QC Flag = 0 (passed with a zero value; Fig. 14), which accounted for about 75.55% of all observations (Table 4). An average of 3.34% of gauge observations were passed with nonzero values (QC Flag = 1). Some seasonal oscillations were noted in the data. Increased passing nonzero values and decreased passing zero values were observed during the DJF periods due to large-scale synoptic systems moving across the CONUS. Seasonal variations were more remarkable within the observations labeled with QC Flag = −10 (Fig. 15), the predominant conditionally passed flag for zero gauge observations that accounted for an average of 7.30% of all observations (Table 4). All other flags designated as conditionally passed were on average ≤0.33% of all observations, with more nonzero observations that were conditionally passed occurring during the non-JJA months.

The 10 flags that classified failed gauge observations had varying impacts across the 21-month study period. Gauges flagged as being outside the ±7-min time window (QC Flag = −20) were approximately 8.56% of all observations (Fig. 16a; Table 4). This was on average the second most utilized flag designation behind QC Flag = 0 (passed gauge

TABLE 3. Breakdown of the average number, percent, and standard deviation of gauges that were passed, conditionally (cond.) passed, and failed by the MRMS gauge QC algorithm for each season. Also listed were the number of hours of observations available per season across the 21-month study period.

| Season | Available hours | Average gauge count | | | Percentage | | | Standard deviation | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Pass | Cond. pass | Fail | Pass | Cond. pass | Fail | Pass | Cond. pass | Fail |
| DJF | 3040 | 13 989 | 1627 | 2736 | 76.2% | 8.9% | 14.9% | 1193 | 405 | 605 |
| MAM | 4206 | 14 356 | 1461 | 2475 | 78.5% | 8.0% | 13.5% | 899 | 293 | 414 |
| JJA | 3572 | 15 136 | 1267 | 2189 | 81.4% | 6.8% | 11.8% | 501 | 161 | 221 |
| SON | 2798 | 14 148 | 1312 | 2421 | 79.1% | 7.3% | 13.6% | 1041 | 335 | 352 |
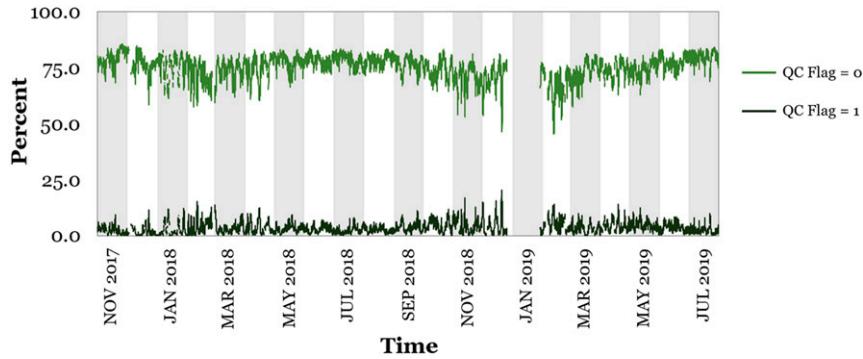
FIG. 14. Percent of gauge observations per hour that were designated with QC flags representing a passed gauge.

observations with zero values). The two dominant QC flags that failed gauges in rain regimes were those that were false zero (QC Flag = 10) and false nonzero (QC Flag = 20) observations (Fig. 16b). Those two classifications represented 1.85% and 1.75% of all observations, respectively (Table 4). There were no discernible seasonal patterns for false zero observations; however, false nonzero observations were flagged more often in the October–April period. Periods of increased flagging of false precipitation classifications correlated with the winter impacts flag for zero observations (QC Flag = 50; Fig. 16c). The winter precipitation flag for zero observations represented an average of 1.02% of all observations (Table 4), yet the gauge QC algorithm classified >2.0% of observations with this flag for 2474 h, >5.0% of observations for 447 h, and >10.0% of observations for 43 h. A similar trend was found in the false nonzero observations for hours following the gauges flagged with winter precipitation impacts (Fig. 16b), which matched the findings of Martinaitis et al. (2015) regarding widespread blockage of gauge observations and post-event thaw from winter precipitation. Gauge observations flagged as suspect observations were <0.01% of all observations (Table 4) and had some increase in use during the October–April period (Fig. 16d).

Overall statistical results were similar to Qi et al. (2016), yet the presented gauge QC algorithm generated improvements in three key aspects: the ability to perform a comprehensive QC on gauge in winter precipitation regimes, the use of dynamic outlier power curves to improve the removal of erroneous nonzero gauge observations, and the additional QC of gauges outside of adequate radar coverage (i.e., when RQI < 0.40). The authors consider the QC of gauges in poor or nonexistent radar coverage to be the most consequential. The ability to analyze gauges with nonradar sources prevented the creation of false precipitation areas despite not conducted direct value comparisons. This was most notable during the melting of winter precipitation. Previous QC would allow reasonable nonzero gauge values when RQI < 0.10 to be conditionally passed, and would subsequently create areas of false precipitation (e.g., Fig. 17a). If the HRRR model identified no precipitation features in the area, then the QC logic would correctly flag the gauges as

false precipitation values, and no hourly precipitation would be generated (e.g., Fig. 17b).

### c. Comparison to manual QC

Forecasters at NWS River Forecast Centers (RFCs) can manually analyze and flag erroneous gauge observations to create accurate Multisensor Precipitation Estimates (MPE; Young et al. 2000). One method for testing the skill of the fully automated MRMS gauge QC algorithm was to compare its results to that of manual human intervention. The NWS Lower Mississippi RFC (LMRFC) retained records of gauges that were removed by forecasters for the generation of MPE. This study compared gauges that were manually flagged by LMRFC to the gauges flagged as failed by the fully automated MRMS gauge QC algorithm. Each gauge observation manually flagged by LMRFC was classified with one of five designations (Table 5). The LMRFC QC flags were directly comparable to the proposed

TABLE 4. The percent of gauge observations per proposed gauge QC flag for the 21-month study period. The table is separated into the three pass/fail group of flags from the QC algorithm.

| Gauge pass/fail group | Proposed gauge QC flag | Percent of observations |
|---|---|---|
| Passed QC | 0 | 75.55% |
| | 1 | 3.34% |
| Conditionally passed QC | −10 | 7.30% |
| | −11 | 0.11% |
| | −12 | 0.33% |
| | −13 | <0.01% |
| Failed QC | −20 | 8.56% |
| | 10 | 1.85% |
| | 20 | 1.75% |
| | 30 | 0.04% |
| | 31 | 0.02% |
| | 40 | 0.07% |
| | 50 | 1.02% |
| | 51 | 0.05% |
| | 60 | <0.01% |
| | 61 | <0.01% |
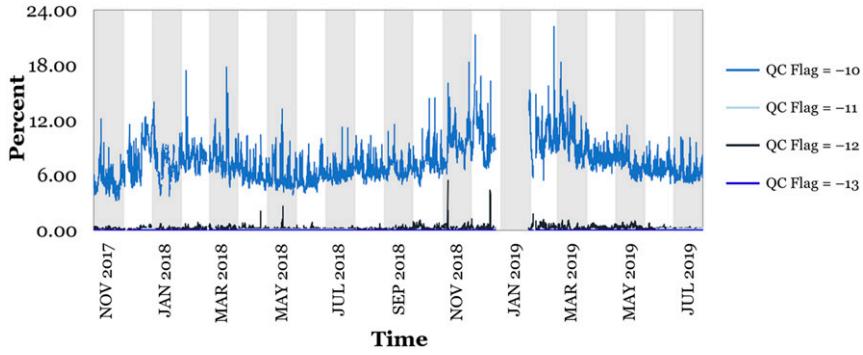
FIG. 15. Percent of gauge observations per hour that were designated with QC flags representing a conditionally passed gauge.
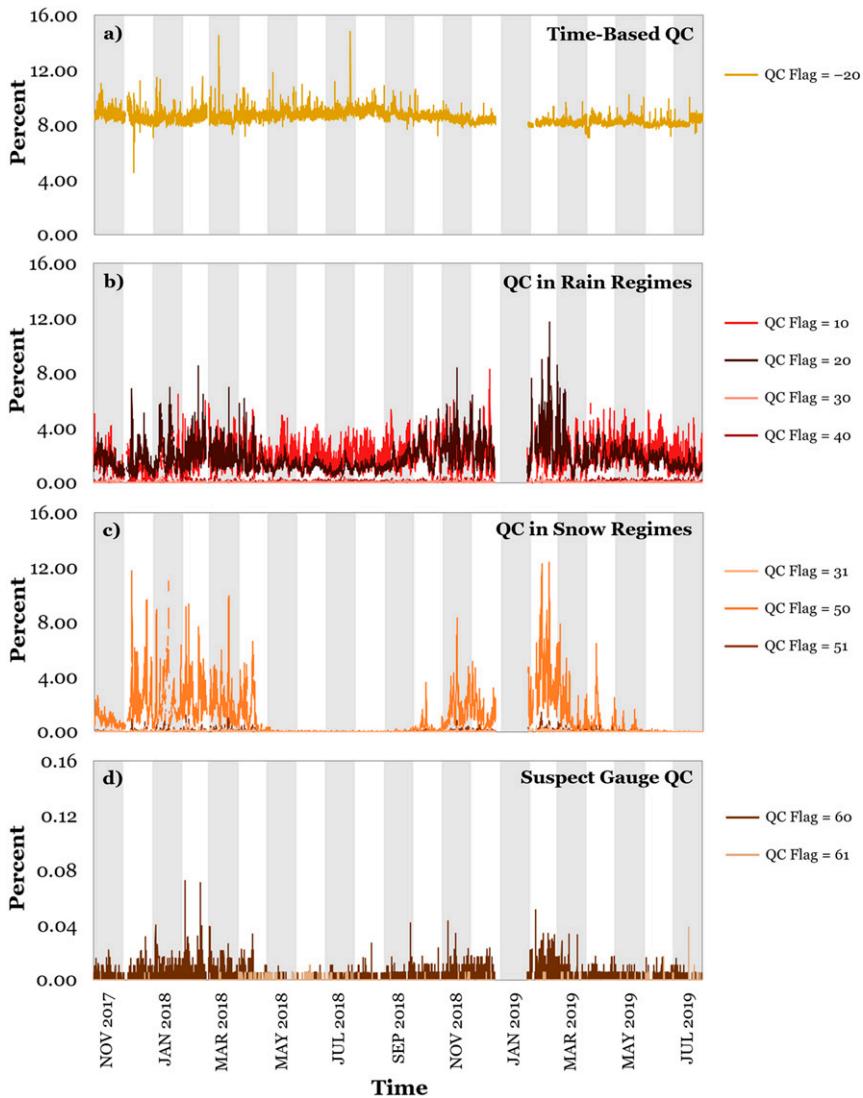


FIG. 16. Percent of gauge observations per hour that were designated with QC flags representing a failed gauge broken down by the following themes: (a) time-based QC, (b) QC conducted in rain regimes (i.e., $T_W > 0.00°C$), (c) QC conducted in snow regimes (i.e., $T_W \leq 0.00°C$), and (d) QC for suspect observations.
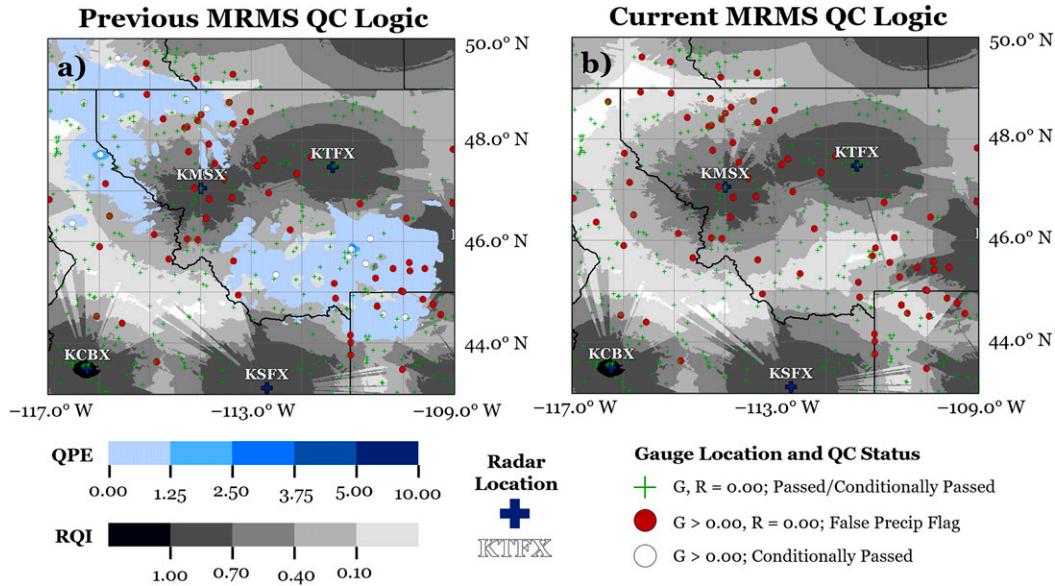
FIG. 17. Comparison of MRMS Mountain Mapper QPE over western Montana and northern Idaho for the 1-h accumulation ending 1900 UTC 3 Apr 2018 using (a) the previous gauge QC logic defined by Qi et al. (2016) and (b) the gauge QC logic presented in this study. Hourly HADS and MADIS gauges were classified by three different categories: when both gauge and gridded QPE values were zero ($G$, $R = 0.00$; green crosses), when the gauge was nonzero and the gridded QPE was zero ($G > 0.00$, $R = 0.00$; red circles), and when the gauge was nonzero ($G > 0.00$) and was conditionally passed in regions of inadequate radar coverage (white circles). The hourly RQI values (gray shading) and radar locations (blue crosses) were provided to characterize the radar coverage in the region.

MRMS QC flag classifications. Some MRMS flag classifications were applicable for multiple LMRFC QC flags.

Comparisons between the MRMS and LMRFC gauge QC results over the LMRFC domain (Fig. 18) were conducted over a 21-month period from November 2017 to July 2019. A total of 6461 hourly gauge observations were matched between the LMRFC and MRMS gauge datasets. Approximately 85.53% of the observations were both flagged by MRMS and LMRFC (Table 6). The 14.47% that did not match (i.e., when the MRMS gauge QC passed the observation) could be attributed to the fact that LMRFC flagging a gauge does not mean a failed observation and its removal from precipitation generation. LMRFC forecasters could identify a potentially erroneous gauge observation during a given hour when analyzing data for operational QPE generations. Subsequent accumulations and temporal analysis of said gauge could later deem it accurate for use in LMRFC QPE production. This would encompass situations where LMRFC forecasters could initially deem a gauge observation as questionable for a given hour, yet 24-h evaluations found the gauge to be performing nominally. The questioned observation would then be retained for use in LMRFC QPE generation.

The reasoning for the observation being flagged when both MRMS and LMRFC flagged a gauge matched 95.82% of the time. The breakdown of the gauge observations that had mismatching QC flags between LMRFC and MRMS showed that 84.42% of mismatches were when the gauge was outside of the ±7-min window as designated by the MRMS QC logic (i.e., classified with QC Flag = −20; Fig. 19). The outside time

window QC flag in MRMS was not considered within the LMRFC classifications. If the gauges flagged by the MRMS time constraint condition were excluded from the analysis, then the matching of QC flags between MRMS and LMRFC were correctly paired 99.32% of the time. This demonstrated that the fully automated QC system within MRMS was comparable to the manual QC conducted by forecasters each hour.

## 4. Observational characteristics

### a. Duplication of observations

The ingest and combination of multiple gauge network feeds into the MRMS system introduced the opportunity for duplicate gauge observations, since observations overlapped between the HADS and MADIS datasets. Duplications were

TABLE 5. The QC flags from LMRFC. Listed are the description of each flag and the corresponding MRMS gauge QC flag(s).

| LMRFC flag | LMRFC flag description | Corresponding MRMS flags |
|---|---|---|
| L | Light gauge value when no precipitation is present | 20 |
| O | Overestimation gauge value | 30, 31 |
| U | Underestimated gauge value | 40, 51 |
| Z | Zero gauge value when precipitation is present | 10, 50 |
| E | Extreme or false gauge value | 20, 30, 31, 60, 61 |
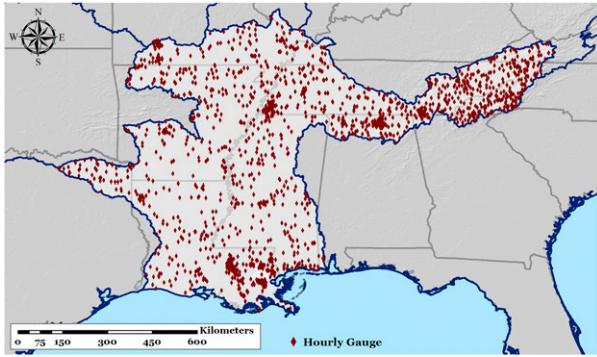
FIG. 18. Locations of available hourly gauge observations in the MRMS system within the LMRFC area of interest.

TABLE 6. Analysis of comparing for gauges that were flagged by LMRFC vs the MRMS gauge QC algorithm. Three categories were analyzed in the QC comparisons based on if both MRMS and LRMFC flagged the gauge and if the QC flags matched (per Table 5). Listed are both the number of and percent of gauges for each category.

| Both MRMS and LMRFC flag gauge | Matching flags | Count | Percent |
|---|---|---|---|
| Yes | Yes | 5295 | 81.95% |
| Yes | No | 231 | 3.58% |
| No | N/A | 935 | 14.47% |

mitigated in the MRMS system based on matching gauge identifications. Priority was given to the HADS network when duplicate observations existed. Analysis of the HADS and MADIS network feeds ingested into the MRMS system from June to October 2017 found 1494 unique gauge IDs that were duplicated but had different latitude and/or longitude coordinates. This accounted for 21.4% of the HADS network sites. Approximately 92.24% of the gauge sites that were identified had location variances ≤ 5.00 km between the HADS and MADIS site location (Table 7). A total of 75.97% of all observations had a difference of ≤1.00 km (the spatial resolution of the MRMS grid). Most differences in the site coordinates for spatial differences ≤ 5.00 km were attributed to rounding and precision differences, including when converting coordinates from degrees, minutes, seconds to decimal degrees (Table 8).

The remaining 7.76% of site coordinates differences identified from June to October 2017 resulted in spatial variances > 5.00 km. There were no discernable primary causes for the differences in site coordinates between the two networks. Large spatial discrepancies were either a result of one or both of the latitude and longitude coordinates being significantly different between the two providing networks (Table 8).

### b. Location errors

Location accuracy of a gauge site is critical for its utility within a high spatial resolution grid. Any spatial error in the application of a gauge can provide a mischaracterization of the rainfall quality during the verification process. Spatial mismatching through locational errors can also introduce irregular bias corrections of a gridded QPE product. The creation of a single gauge list through the combination of various gauge networks and the subsequent identification of location differences via site duplications in the MRMS system necessitated a review of the accuracy of the latitude and longitude coordinates and the performance impact it had on the MRMS system.

A total of 838 ASOS sites were evaluated for their location accuracy and compared to the MRMS radar-only QPE from 0000 UTC 1 January 2015 to 2300 UTC 31 December 2015 over the CONUS. The latitude and longitude coordinates were obtained from the NWS Location Identifiers (NWSLI) database. Visual identification of the ASOS sites were conducted using satellite imagery from the Google Earth system, a platform image and mapping dataset using locational information and extrapolated satellite imagery. Studies of various cities and regions across the world had found the horizontal positional accuracy of Google Earth to have a mean offset distance of 4.0–7.0 m with an average horizontal RMSE of 1.8–5.0 m (e.g., Paredes-Hernández et al. 2013; Mohammed et al. 2013;



FIG. 19. Distribution of flag mismatches between the MRMS gauge QC algorithm and the LMRFC flag designations.

TABLE 7. The percent of site coordinate distance differences (km) from duplicate observations between the HADS and MADIS network collections as seen within the MRMS system.

| Distance difference in site coordinates (km) | Percent of duplicate sites with different coordinates |
|---|---|
| $D \leq 0.1$ km | 64.32% |
| $0.1$ km $< D \leq 1.0$ km | 11.65% |
| $1.0$ km $< D \leq 5.0$ km | 16.27% |
| $5.0$ km $< D \leq 10.0$ km | 1.81% |
| $10.0$ km $< D \leq 50.0$ km | 2.07% |
| $D < 50.0$ km | 3.88% |

Farah and Algarni 2014). The imaging quality of the satellite data varied at each location; thus, the identification of the ASOS power supply box was used as the verification coordinate. A total of 817 ASOS stations were identified and recorded through visual inspection within Google Earth, while 21 stations were unable to be located through visual inspection.

The distance between the ASOS site locations from the NWSLI database and the Google Earth satellite site identifications were calculated using the Haversine formula. Approximately 83.57% of ASOS NWSLI coordinates were within 1.00 km of the Google Earth identified location (Table 9). The majority of these differences were between 0.25 and 1.00 km (e.g., Fig. 20a). The remaining 16.43% of horizontal positional differences were classified as >1.00 km (e.g., Fig. 20b). Only five ASOS sites (0.61%) had locational differences between Google Earth and NWSLI that were within the upper bound of mean offset distance differences (i.e., <7.0 m) found in the horizontal positional studies.

The Google Earth defined locations resulted in 79.31% of identified ASOS sites residing in a different MRMS grid cell compared to the NWSLI coordinates. The impacts of having gauge sites located outside of their true location were demonstrated through a QPE collocation comparison between the NWSLI and Google Earth coordinates for only the ASOS sites where the different coordinates resided in different grid cells. The MRMS radar-only QPE from 0000 UTC 1 January to 2300 UTC 31 December 2015 was utilized as the gridded QPE. Situations where the gridded radar-only QPE was zero at both the NWSLI and Google

Earth coordinates were excluded from the analysis. This resulted in a sample size of 246 504 observations.

Comparisons between the hourly gauge observations and the MRMS radar-only QPE values based on the NWSLI coordinates showed a number of outlier pairings (Fig. 21a). The most notable outlier data pairs were instances when the gauge value was nonzero while the collocated grid point based on the NWSLI coordinates had a zero QPE value. Applying the new gauge coordinates based on Google Earth mitigated the gauge pairings with zero QPE values and reduced the overall scatter of the data pairings (Fig. 21b). Statistical evaluations of the gauge versus MRMS radar-only QPE values when utilizing the two different coordinate locations depicted a 7.8% improvement of the root-mean-square error from 2.418 to 2.230 mm using the Google Earth coordinates; moreover, the correlation coefficient improved from 0.776 to 0.810 (Table 10). There was an increase in the underestimation signal of the MRMS radar-only QPE through the mean bias ratio value using the Google Earth coordinates; however, this finding was better aligned with other studies that analyzed the performance of MRMS radar-only QPE, particularly in cool season environments (Cocks et al. 2016, 2017).

### c. Instrumentation characteristics

Numerous studies regarding the identification and quantification of various rain gauge observational limitations included methodologies for adjusting these observations. Extensive analyses derived from atmospheric conditions (e.g., Goodison et al. 1998; Yang et al. 1998) or from mechanical limitations (e.g., Parsons 1941) highlighted the potential biases that can occur from wetting losses, evaporation, wind undercatch, and tipping limitations (e.g., maximum tipping rate, double tips, etc.). Variances in the percent of losses were shown to be dependent upon the precipitation type and atmospheric conditions; however, Goodison et al. (1998) demonstrated how catch ratios were also dependent on the design and instrumentation of the gauges themselves.

Research that examined just the effects of wind undercatch and attempted to adjust for it had varying adjustment equations for both liquid precipitation and solid winter precipitation (Fig. 22; Table 11). There was a significant distinction between equations applied for rainfall events versus solid winter precipitation events where adjustment equations for

TABLE 8. Example gauge observations of distance differences (km) between sites listed in both HADS and MADIS. Listed are the latitude and longitude values for each example gauge site.

| Site ID | HADS | | MADIS | | Distance difference (km) |
|---|---|---|---|---|---|
| | Latitude (°) | Longitude (°) | Latitude (°) | Longitude (°) | |
| CBBN5 | 35.9419 | −107.0772 | 35.9419 | −107.0773 | 0.009 |
| RNDC1 | 41.4269 | −121.4625 | 41.4270 | −121.4626 | 0.014 |
| STBM3 | 42.0828 | −72.0575 | 42.0825 | −72.0583 | 0.074 |
| VLCC1 | 33.2369 | −117.0142 | 33.2370 | −117.0086 | 0.521 |
| BEFM8 | 47.4889 | −112.8800 | 47.5000 | −112.9000 | 1.945 |
| GOSC1 | 33.0844 | −116.6714 | 33.0843 | −116.8769 | 19.146 |
| CTLN2 | 39.4750 | −114.9861 | 38.9038 | −114.8142 | 65.219 |
| BDAN2 | 32.8211 | −106.8819 | 33.8210 | −106.8819 | 111.184 |

TABLE 9. The percent of site coordinate distance differences (km) from the coordinates of ASOS sites provided by NWSLI and those identified within Google Earth.

| Distance difference in site coordinates (km) | Percent of ASOS sites |
|---|---|
| $D \leq 0.10$ km | 13.11% |
| $0.10$ km $< D \leq 0.25$ km | 15.56% |
| $0.25$ km $< D \leq 0.50$ km | 25.12% |
| $0.50$ km $< D \leq 1.00$ km | 29.78% |
| $1.00$ km $< D \leq 2.00$ km | 12.99% |
| $2.00$ km $< D \leq 10.00$ km | 3.07% |
| $D < 10.0$ km | 0.37% |

solid winter precipitation were much more aggressive compared to rainfall-based equations. Previous studies pertaining to snowfall were capped at $5$–$7$ m s$^{-1}$ due to challenges in generating a correction factor in winds exceeding $7$ m s$^{-1}$; moreover, blowing snow can also attribute to inaccuracies in a gauge observation (e.g., Goodison et al. 1998). Other meteorological factors were also shown to influence some adjustment equations. Some rainfall-based equations considered the instantaneous rain rate when adjusting for wind undercatch (Table 11). Førland et al. (1996) utilized the ambient temperature to adjust for wind undercatch in snow (Table 11). Application of a temperature component to the Førland et al. (1996) adjustment equation showed a 2% difference per degree Celsius at a wind speed of $5$ m s$^{-1}$.

Differences in the adjustment equations were also noted between those applied to shielded gauges versus unshielded gauges (Fig. 22). Variations in the rainfall-based equations with respect to shield availability were shown to be minimal. Previous studies by Duchon and Essenberg (2001) and Yang et al. (1998) demonstrated that the use of a wind shield during a rainfall event would reduce wind undercatch by 1%–3%; however, those studies examined events where the winds were $\leq 15$ m s$^{-1}$. It is unknown if a larger disparity exists between shielded and unshielded gauges during more significant wind events (e.g., tropical cyclones). The application of a wind shield was shown to be significant for solid winter precipitation, particularly snowfall. Adjustment ratios for shielded gauges ranged from 1.50 to 3.45 at a wind magnitude of $7$ m s$^{-1}$, while unshielded gauges had adjustment ratios of 5.20–6.65 at the same wind magnitude.

Each of the aforementioned studies that examined observational adjustments focused on a singular gauge type where gauge instrumentation characteristics were known, but this information is not available in large collections of gauge networks. Publicly available metadata provided with gauge collections generally consisted of the site identifier, the latitude and longitude coordinates, and occasionally other information such as the elevation of the station above mean sea level and the source network of the gauge. The metadata information needed to provide accurate adjustments for overcoming observational limitations are not freely available; moreover, the lack of gauge instrumentation characteristics would not permit for accurate adjustment equations.

The work by Martinaitis et al. (2021) highlighted the shortcomings of adjusting for wind undercatch in rainfall through a
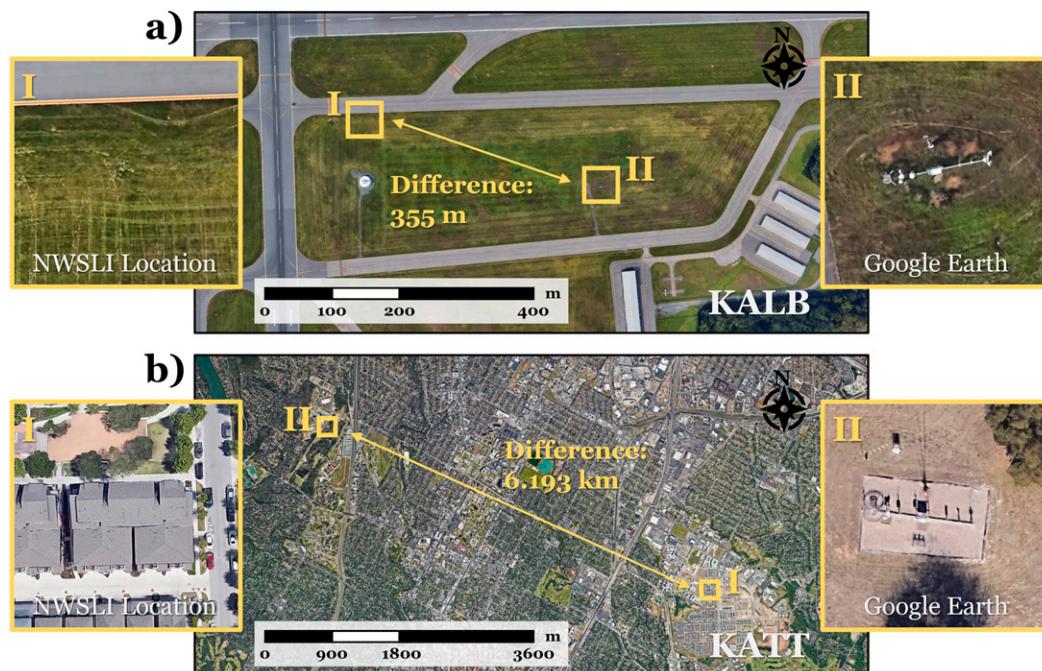


FIG. 20. Example illustrations of the difference between the location of gauges in Google Earth using the NWSLI latitude and longitude coordinates vs the locations of ASOS gauges in the Google Earth system for the gauges (a) Albany, NY (KALB), and (b) Austin, TX (KATT).
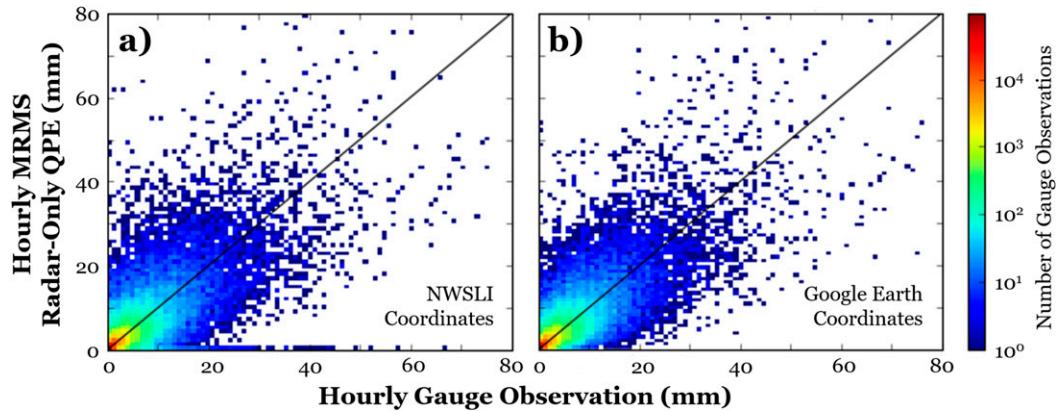
FIG. 21. Density scatterplots comparing hourly gauge observations (mm) to hourly MRMS radar-only QPE (mm) at the collocated MRMS grid cell based on gauge locations from (a) the NWSLI coordinates and (b) Google Earth for the study period 0000 UTC 1 Jan–2300 UTC 31 Dec 2015. ASOS sites considered in the analysis were those where the different coordinates between NWSLI and Google Earth resided in different MRMS grid cells. Hours where the gridded MRMS radar-only QPE was zero at both the NWSLI and Google Earth coordinates were also excluded.

blanket methodology given the lack of metadata on gauge instrumentation characteristics. Assumptions were made to tailor the adjustment equation to the gauge properties. A singular adjustment equation was applied, since it was unknown what gauges were shielded or not given the lack of wind shield metadata; moreover, the difference in wind undercatch errors between shielded and unshielded gauges was shown to be negligible in previous research (Duchon and Essenberg 2001; Yang et al. 1998). The height of all gauges was assumed to be 1.00 m given the need to interpolate 10-m winds down to a gauge height. It was also presumed that all hourly gauge observations have an 8-in. (20.3 cm) diameter orifice, yet Nešpor and Sevruk (1999) demonstrated that wind-induced errors varied based on the orifice diameter. While the results of applying a single conceptual wind undercatch correction to all observations yielded positive results, it is unknown how accurate those adjustments were. Applying a similar single methodology to winter precipitation would likely lead to reduced confidence in the correction for wind undercatch given the importance of knowing if a wind shield exists when adjusting the gauge observations.

## 5. Summary

The MRMS v12.0 system update afforded the opportunity to evaluate all aspects of the ingest and QC of hourly gauge observations within the MRMS system. The addition of new feeds increased the number of hourly gauge observations that were available pre-QC from approximately 7000 to 18 000–20 000 per hour; moreover, analysis of the data latency characteristics showed that about 65% of all observations would be available for a quicker development of gauge-based and gauge-adjusted products in the new two-pass paradigm of product delivery.

New advancements were introduced in the fully automated MRMS gauge QC algorithm, including the use of HRRR model QPF as a means to QC gauges outside of adequate radar

coverage, the use of dynamic outlier curves in regions of adequate radar coverage, and new logic to handle gauges in winter precipitation environments. Updates were also made to the best-matching radial check and the time window check. Analysis of a 21-month dataset showed that approximately 86.63% of gauge observations per hour were passed or conditionally passed for use in the MRMS system. The majority of gauges that were failed fell outside of the ±7-min time window designed to prevent the mismatching of observations to top-of-the-hour gridded QPEs. Gauges that were failed by the MRMS system were compared to manual gauge QC conducted by the NWS LMRFC. The reasons for flagging a gauge matched for >95% of observations.

Other characteristics of the gauge ingest process were identified throughout the study. The utilization of HADS and MADIS networks introduced the opportunity for duplicate observations. Most discrepancies in the duplicate observations were a result of precision and rounding differences in gauge latitude and longitude coordinates. Investigations of the ASOS gauge network showed that >16% of ASOS coordinates

TABLE 10. Statistical analysis of ASOS gauge sites based on their NWSLI site coordinates vs the coordinates identified via Google Earth. Listed are the mean bias ratio (radar vs gauge), mean error (mm), root-mean-square error (RMSE; mm), and correlation coefficient (CC). ASOS sites considered in the analysis were those where the different coordinates between NWSLI and Google Earth resided in different MRMS grid cells. Hours where the gridded MRMS radar-only QPE was zero at both the NWSLI and Google Earth coordinates were also excluded.

| Coordinate source | Mean bias ratio | Mean error (mm) | RMSE (mm) | CC |
|---|---|---|---|---|
| NWSLI | 0.917 | 0.169 | 2.418 | 0.776 |
| Google Earth | 0.892 | 0.225 | 2.230 | 0.810 |
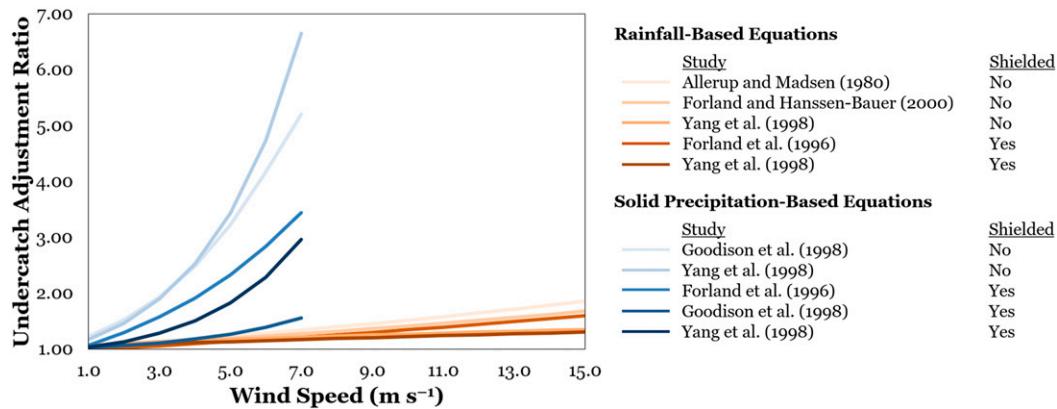
FIG. 22. Illustration of the various adjustment equations for wind undercatch from previous studies. The equations are color coded based on those related to rainfall (orange) vs solid winter precipitation (blue) as well as studies that utilized gauges that were shielded (dark shading) and unshielded (light shading).

differed by $>1.0$ km when identified through Google Earth; moreover, over 79% of differences in the ASOS latitude and longitude coordinates resulted in the gauge being located in a different MRMS grid cell. The utilization of new Google Earth-defined coordinates improved the comparisons of gauge observations to gridded MRMS radar-only QPEs. These findings were delivered to the NWS and resulted in changes to the ASOS metadata.

Previous studies made efforts to adjust for limitations in recording precise gauge observations, yet the challenge presented to applications and systems that ingest multiple gauge networks is the lack of available metadata to properly adjust these observations. Gauge networks ingested by MRMS were not accompanied with the metadata needed to effectively modify the gauge value to account for wind undercatch or any other observational limitations of gauges. Not having such gauge instrumentation characteristics freely available coupled

with any potential blanket methodology to correct for gauge limitations (e.g., wind undercatch, wetting loss, evaporation, tipping limitations, etc.) would likely produce inaccurate adjustments to the gauge observation, especially for solid winter precipitation where greater variations in adjustment factors based on instrumentation exist.

The authors believe that continued studies on the limitations of and the ability to adjust gauge observations coupled with accessibility to complete and accurate gauge metadata would improve the ability to create QPEs needed for various hydrometeorological applications. Efforts should be made to better document and make available the exact characteristics of gauge instrumentation at each site. Having the knowledge of such characteristics as the gauge type and its measurement limitations, gauge height, orifice diameter, and the shielding and heating element configurations across all networks ingested by large gauge collection platforms would further increase

TABLE 11. Adjustment equations for wind undercatch from recent studies. Listed are the precipitation type studied, the gauge type, if a windshield was present, the adjustment equation, and the accumulation period.

| Study | Precipitation type | Gauge type | Shield | Adjustment ratio equation | Accumulation |
|---|---|---|---|---|---|
| Allerup and Madsen (1980) | Rain | Hellman 8 in. | No | $\exp\{[-0.0010 \times \ln(R)] - [0.0082 \times u \times \ln(R)] + (0.0420 \times u) + 0.0100\}$ | Instantaneous |
| Førland et al. (1996) | Rain | Hellman 8 in. | Yes | $\exp\{[-0.001\,01 \times \ln(R)] - [0.012\,177 \times u \times \ln(R)] + (0.034\,331 \times u) + 0.007\,697 - 0.05\}$ | Instantaneous |
| Førland and Hanssen-Bauer (2000) | Rain | Hellman 8 in. | No | $\exp\{[-0.001\,01 \times \ln(R)] - [0.012\,177 \times u \times \ln(R)] + (0.034\,331 \times u) + 0.007\,697\}$ | Instantaneous |
| Yang et al. (1998) | Rain | NWS 8 in. | No | $\exp[4.605 - (0.062 \times u^{0.58})]$ | Daily |
| Yang et al. (1998) | Rain | NWS 8 in. | Yes | $\exp[4.606 - (0.041 \times u^{0.69})]$ | Daily |
| Førland et al. (1996) | Snow | Geonor 8 in. | Yes | $\exp[-0.121\,59 + (0.185\,46 \times u) + (0.006\,918 \times T) - (0.005\,254 \times T \times u)]$ | Hourly |
| Goodison et al. (1998) | Snow | Nipher | Yes | $100.00 - (0.44 \times u^2) - (1.98 \times u)$ | Daily |
| Yang et al. (1998) | Snow | NWS 8 in. | Yes | $\exp[4.606 - (0.036 \times u^{1.75})]$ | Daily |
| Goodison et al. (1998) | Snow | Hellman 8 in. | No | $100.00 + (1.13 \times u^2) - (19.45 \times u)$ | Daily |
| Yang et al. (1998) | Snow | NWS 8 in. | No | $\exp[4.606 - (0.157 \times u^{1.28})]$ | Daily |

the ability to tailor any adjustment factors to each gauge configuration. Only then could the hydrometeorological community gain a greater understanding of how much precipitation is occurring during any given event.

## REFERENCES

Allerup, P., and H. Madsen, 1980: Accuracy of point precipitation measurements. *Nord. Hydrol.*, **11**, 57–70, https://doi.org/10.2166/nh.1980.0005.

Benjamin, S. G., and Coauthors, 2016: A North American hourly assimilation and model forecast cycle: The Rapid Refresh. *Mon. Wea. Rev.*, **144**, 1669–1694, https://doi.org/10.1175/MWR-D-15-0242.1.

Chen, M., and P. Xie, 2008: Quality control of daily precipitation reports at NOAA/CPC. *12th Conf. on IOAS-AOLS*, New Orleans, LA, Amer. Meteor. Soc., 3.3, https://ams.confex.com/ams/88Annual/techprogram/paper_131381.htm.

Cocks, S. B., S. Martinaitis, B. Kaney, J. Zhang, and K. Howard, 2016: MRMS QPE performance during the 2013/14 cool season. *J. Hydrometeor.*, **17**, 791–810, https://doi.org/10.1175/JHM-D-15-0095.1.

——, J. Zhang, S. Martinaitis, Y. Qi, B. Kaney, and K. Howard, 2017: MRMS QPE performance east of the Rockies during the 2014 warm season. *J. Hydrometeor.*, **18**, 761–775, https://doi.org/10.1175/JHM-D-16-0179.1.

——, and Coauthors, 2019: A prototype quantitative precipitation estimation algorithm for operational S-band polarimetric radar utilizing specific attenuation and specific differential phase. Part II: Performance verification and case study analysis. *J. Hydrometeor.*, **20**, 999–1014, https://doi.org/10.1175/JHM-D-18-0070.1.

Duchon, C. E., and G. R. Essenberg, 2001: Comparative rainfall observations from pit and aboveground rain gauges with and without wind shields. *Water Resour. Res.*, **37**, 3253–3263, https://doi.org/10.1029/2001WR000541.

Farah, A., and D. Algarni, 2014: Positional accuracy assessment of GoogleEarth in Riyadh. *Artif. Satell.*, **49**, 101–106, https://doi.org/10.2478/arsa-2014-0008.

Førland, E. J., and I. Hanssen-Bauer, 2000: Increased precipitation in the Norwegian Arctic: True or false? *Climatic Change*, **46**, 485–509, https://doi.org/10.1023/A:1005613304674.

——, and Coauthors, 1996: Manual for operational correction of Nordic precipitation data. DNMI KLIMA Rep. 24/96, 66 pp.

Goodison, B. E., P. Y. T. Louie, and D. Yang, 1998: WMO solid precipitation measurement intercomparison. Instruments and Observing Methods Rep. 67, WMO/TD-872, 212 pp., 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, https://doi.org/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, https://doi.org/10.1029/1999JD900228.

Helms, D., P. Miller, M. Barth, D. Starosta, B. Gordon, S. Schofield, F. Kelly, and S. Koch, 2019: 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, 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, https://doi.org/10.1175/2009WAF2222227.1.

Kirkham, J. D., and Coauthors, 2019: Near real-time measurements of snow-water equivalent in the Nepal Himalayas. *Front. Earth Sci.*, **7**, 177, https://doi.org/10.3389/feart.2019.00177.

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, https://ams.confex.com/ams/Annual2006/techprogram/paper_102834.htm.

Lewis, E., and Coauthors, 2018: A rule based quality control method for hourly rainfall data and a 1 km resolution gridded hourly rainfall dataset for Great Britain: CEH-GEAR1hr. *J. Hydrol.*, **564**, 930–943, https://doi.org/10.1016/j.jhydrol.2018.07.034.

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

——, H. M. Grams, C. Langston, J. Zhang, and K. Howard, 2018: A real-time evaporation correction scheme for radar-derived mosaicked precipitation estimations. *J. Hydrometeor.*, **19**, 87–111, https://doi.org/10.1175/JHM-D-17-0093.1.

——, and Coauthors, 2020: A physically based multisensor quantitative precipitation estimation approach for gap-filling radar coverage. *J. Hydrometeor.*, **21**, 1485–1511, https://doi.org/10.1175/JHM-D-19-0264.1.

——, S. B. Cocks, A. P. Osborne, M. J. Simpson, L. Tang, J. Zhang, and K. W. Howard, 2021: The historic rainfalls of Hurricanes Harvey and Florence: A perspective from the Multi-Radar Multi-Sensor system. *J. Hydrometeor.*, **22**, 721–738, https://doi.org/10.1175/JHM-D-20-0199.1.

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, https://ams.confex.com/ams/Annual2005/techprogram/paper_83718.htm.

Mascaro, G., 2017: Multiscale spatial and temporal statistical properties of rainfall in central Arizona. *J. Hydrometeor.*, **18**, 227–245, https://doi.org/10.1175/JHM-D-16-0167.1.

Matsuo, T., and Y. Sasyo, 1981: Non-melting phenomena of snowflakes observed in subsaturated air below freezing level. *J. Meteor. Soc. Japan*, **59**, 26–32, https://doi.org/10.2151/jmsj1965.59.1_26.

Mohammed, N. Z., A. Ghazi, and H. E. Mustafa, 2013: Positional accuracy testing of Google Earth. *Int. J. Multidiscip. Sci. Eng.*, **4**, 6–9.

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, https://doi.org/10.1175/1520-0426(1999)016<0450:EOWIEO>2.0.CO;2.

Paredes-Hernández, C. U., W. E. Salinas-Castillo, F. Guevara-Cortina, and X. Martínez-Becerra, 2013: Horizontal positional accuracy of Google Earth's imagery over rural areas: A study case in Tamaulipas, Mexico. *Bull. Geod. Sci.*, **19**, 588–601, https://doi.org/10.1590/S1982-21702013000400005.

Parsons, D. A., 1941: Calibration of a Weather Bureau tipping-bucket gage. *Mon. Wea. Rev.*, **69**, 205–208, https://doi.org/10.1175/1520-0493(1941)069<0205:COAWBT>2.0.CO;2.

Qi, Y., S. Martinaitis, J. Zhang, and S. Cocks, 2016: A real-time automated quality control of hourly rain gauge data based on multiple sensors in MRMS system. *J. Hydrometeor.*, **17**, 1675–1691, https://doi.org/10.1175/JHM-D-15-0188.1.

Rasmussen, R., and Coauthors, 2012: How well are we measuring snow: The NOAA/FAA/NCAR winter precipitation test bed. *Bull. Amer. Meteor. Soc.*, **93**, 811–829, https://doi.org/10.1175/BAMS-D-11-00052.1.

Seo, D. J., and J. P. Breidenbach, 2002: Real-time correction of spatially nonuniform bias in radar rainfall data using rain gauge measurements. *J. Hydrometeor.*, **3**, 93–111, https://doi.org/10.1175/1525-7541(2002)003<0093:RTCOSN>2.0.CO;2.

Sevruk, B., 1989: Wind induced measurement error for high-intensity rains. WMO Tech. Doc., 328 pp.

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

——, J.-A. Hertig, and R. Spiess, 1991: The effect of precipitation gauge orifice rim on the wind field deformation as investigated in a wind tunnel. *Atmos. Environ.*, **25A**, 1173–1179, https://doi.org/10.1016/0960-1686(91)90228-Y.

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

Smith, J. A., and W. F. Krajewski, 1991: Estimation of the mean field bias of radar rainfall estimates. *J. Appl. Meteor.*, **30**, 397–412, https://doi.org/10.1175/1520-0450(1991)030<0397:EOTMFB>2.0.CO;2.

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, https://doi.org/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, https://ams.confex.com/ams/pdfpapers/95173.pdf.

Wang, Y., S. Cocks, L. Tang, A. Ryzhkov, P. Zhang, J. Zhang, and K. Howard, 2019: A prototype quantitative precipitation estimation algorithm for operational S-band polarimetric radar utilizing specific attenuation and specific differential phase: Part I: Algorithm description. *J. Hydrometeor.*, **20**, 985–997, https://doi.org/10.1175/JHM-D-18-0071.1.

Wilson, J. W., and E. A. Brandes, 1979: Radar measurement of rainfall: A summary. *Bull. Amer. Meteor. Soc.*, **60**, 1048–1058, https://doi.org/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, https://doi.org/10.1175/1520-0426(1998)015<0054:AONSNP>2.0.CO;2.

Young, C. B., A. A. Bradley, W. F. Krajewski, A. Kruger, and M. L. Morrissey, 2000: Evaluating NEXRAD multisensor precipitation estimates for operational hydrologic forecasting. *J. Hydrometeor.*, **1**, 241–254, https://doi.org/10.1175/1525-7541(2000)001<0241:ENMPEF>2.0.CO;2.

Zhang, J., 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., International Association of Hydrological Science, 388–393.

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

——, L. Tang, S. Cocks, P. Zhang, A. Ryzhkov, K. Howard, C. Langston, and B. Kaney, 2020: A dual-polarization radar synthetic QPE for operations. *J. Hydrometeor.*, **21**, 2507–2521, https://doi.org/10.1175/JHM-D-19-0194.1.