MRMS QPE Performance during the 2013/14 Cool Season

STEPHEN B. COCKS, STEVEN M. MARTINAITIS, AND BRIAN KANEY

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

JIAN ZHANG AND KENNETH HOWARD

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

(Manuscript received 15 June 2015, in final form 1 December 2015)

ABSTRACT

A recently implemented operational quantitative precipitation estimation (QPE) product, the Multi-Radar Multi-Sensor (MRMS) radar-only QPE (Q3RAD), mosaicked dual-polarization QPE, and National Centers for Environmental Prediction (NCEP) stage II QPE were evaluated for nine cool season precipitation events east of the Rockies. These automated, radar-only products were compared with the forecaster quality-controlled NCEP stage IV product, which was considered as the benchmark for QPE. Community Collaborative Rain, Hail and Snow Network (CoCoRaHS) 24-h accumulation data were used to evaluate product performance while hourly automated gauge data (quality controlled) were used for spatial and time series analysis. Statistical analysis indicated all three radar-only products had a distinct underestimate bias, likely due to the radar beam partially or completely overshooting the predominantly shallow winter precipitation systems. While the forecaster quality-controlled NCEP stage IV estimates had the best overall performance, Q3RAD had the next best performance, which was significant as Q3RAD is available in real time whereas NCEP stage IV estimates are not. Stage II estimates exhibited a distinct tendency to underestimate gauge totals while dual-polarization estimates exhibited significant errors related to melting layer challenges.

1. Introduction

Multi-Radar Multi-Sensor (MRMS) quantitative precipitation estimation (QPE) products have been transitioned into the National Weather Service (NWS) operations at the National Centers for Environmental Predictions (NCEP; Zhang et al. 2014). As part of this transition, a systematic validation and verification effort is underway to characterize the MRMS performance in meteorological, aviation, and hydrological applications. For hydrological applications, this includes evaluating product performance for precipitation events during the warm and cool season as well as in mountainous and less challenging terrain. While other studies are currently in progress, this report documents MRMS QPE performance east of the Rockies during the 2013/14 cool season. While the cool season represents a distinct challenge for evaluating radar estimates because of radar beam overshoot (Smith et al. 1996), frozen precipitation impacts on gauges (Martinaitis et al. 2015), and brightband contamination (Smith et al. 1996) within the melting layer, there is, nonetheless, a need to see how well the operational precipitation estimate products handle these challenges.

Evaluations of QPE often involve an intercomparison of radar precipitation estimates to rain gauge accumulations (e.g., Steiner et al. 1999; Martinaitis et al. 2015); hence, there are a number of limitations that must be considered (Wilson and Brandes 1979; Krajewski et al. 2010). Ground clutter, blockage, and nonmeteorological echoes can contaminate the lower-elevation scans; however, the extra information provided by dual-polarization (Dual-Pol) radar data has been used by MRMS to mitigate these effects (Tang et al. 2014). Increased sampling volume at greater distances (Steiner et al. 1999; Zhang et al. 2012), beam overshoot and brightbanding in the melting layer (Smith et al. 1996; Zhang and Qi 2010), improper calibration, and use of improper reflectivity–rain rate (Z–R) relationships (Wilson and Brandes 1979; Steiner et al. 1999) can significantly affect radar-derived

Corresponding author address: Stephen B. Cocks, National Severe Storms Laboratory, 120 David L. Boren Blvd., Norman, OK 73072.
E-mail: stephen.cocks@noaa.gov

DOI: 10.1175/JHM-D-15-0095.1

TABLE 1. Dates, precipitation types and reported totals, and short event summary for the nine precipitation events evaluated in the study. The precipitation totals (in parentheses; pertaining to the type of precipitation listed before the parentheses) are for the 24-h period ending at 1200 UTC for the given date.

Date	Precipitation types and notable totals	Synoptic summary
6 Dec 2013	Rain (25–100 mm), freezing rain, sleet, and snow (25–100 mm)	Precipitation developed along/behind a strong cold front
7 Dec 2013	Rain (25–75 mm), freezing rain, sleet, and snow (50–250 mm)	Precipitation developed along/behind stationary front
22 Dec 2013	Rain (100–250 mm), freezing rain, sleet, and snow (50–220 mm)	Precipitation developed along/behind cold front
23 Dec 2013	Primarily rain (70–170 mm)	Precipitation developed along a slow-moving cold front
6 Jan 2014	Primarily snow (100–380 mm)	Precipitation developed along a cold front
29 Jan 2014	Rain (12–25 mm), freezing rain (5–12 mm), sleet, and snow (25–75 mm)	Precipitation, associated with upper trough, developed over cold surface dome of high pressure
3 Feb 2014	Rain (50–125 mm), freezing rain, sleet, and snow	Precipitation developed along and behind a cold front
5 Feb 2014	Rain (30–80 mm), freezing rain, sleet, and snow (100–300 mm)	Precipitation developed along and behind a stationary front
13 Feb 2014	Rain (25–60 mm), freezing rain, sleet, and snow	Precipitation, associated with upper trough, developed over cold surface dome of high pressure

precipitation estimates. Conversely, blockages and poor site placement (Sieck et al. 2007; Fiebrich et al. 2010), gauge undercatch due to strong winds (Wilson and Brandes 1979; Sieck et al. 2007), power outages preventing data transmission (Martinaitis 2008), mechanical malfunctions, and telemetry and transmission problems (Groisman and Legates 1994; Marzen and Fuelberg 2005; Kim et al. 2009) can contribute to gauge errors.

In this study, it was clear some of these error factors were present. Analysis indicated a number of the automated reporting gauges were becoming clogged, or "stuck," because of frozen precipitation (Martinaitis et al. 2015) adversely affecting precipitation products that used these gauges to adjust radar estimates. Therefore, this paper documented the performance of the MRMS radar-only QPE (Q3RAD; Zhang et al. 2014), the mosaicked NWS operational Dual-Pol QPE (Giangrande and Ryzhkov 2008), and the NCEP stage II (Lin and Mitchell 2005) during the 2013/14 cool season over the United States east of the Rockies. In turn, the product performance was compared to a benchmark estimate, the forecaster quality-controlled NCEP stage IV product.

2. Data and methodology

The precipitation events evaluated (Table 1) occurred between December 2013 and February 2014 and had significant areas of precipitation totals ≥50.8 mm (2.00 in.) and/or significant impacts due to frozen precipitation. Upper-air, numerical model, and radar data combined with radar precipitation and gauge totals were evaluated for 24-h periods ending at 1200 UTC. Hourly and 24-h radar-derived estimates R and gauge accumulations G were compared at corresponding grid points (henceforth called R/G pairs). Approximately 10 000 rain gauges from a variety of national and regional networks are ingested by the MRMS system, including 24-h precipitation data from the Community Collaborative Rain, Hail and Snow Network (CoCoRaHS) and hourly data from the Hydrometeorological Automated Data System (HADS). CoCoRaHS gauge totals are reported by volunteer observers trained to monitor and report at 0700 local standard time each day the liquid and frozen precipitation types measured by catchment rain gauges. HADS gauges are automated, with the primary gauge type being the heated tipping-bucket variety.

In this study, gauge performance challenges during frozen precipitation were prevalent. Figure 1a shows the Q3RAD 24-h precipitation estimate with Meteorological Phenomena Identification Near the Ground (mPING) crowd-sourcing reports (Elmore et al. 2014) and model surface freezing lines (dashed lines) superimposed. The purple bias ratio dots denote locations where gauge performance was suspect. Approximately 80% of the suspect gauges were located where frozen precipitation occurred (see pink, white, and blue mPING symbols and model surface temperature analysis); hence, the gauges likely were stuck (Martinaitis et al. 2015). These gauges impacted the MRMS Q3RAD gauge-corrected (Q3GC) product, which was calculated by interpolating hourly radar–gauge differences onto a grid (using inverse distance weighting; see Zhang et al. 2016), which in turn are subtracted from the hourly Q3RAD fields. Hence, Q3GC is quite dependent on gauge accuracy, and stuck/suspect gauges significantly impacted the products estimate, as shown in Fig. 2. NWS analysis (not shown) indicated that a swath of moderate to heavy snow fell from the Texas Panhandle into Iowa, which can be seen in the Q3RAD estimate (Fig. 2a,

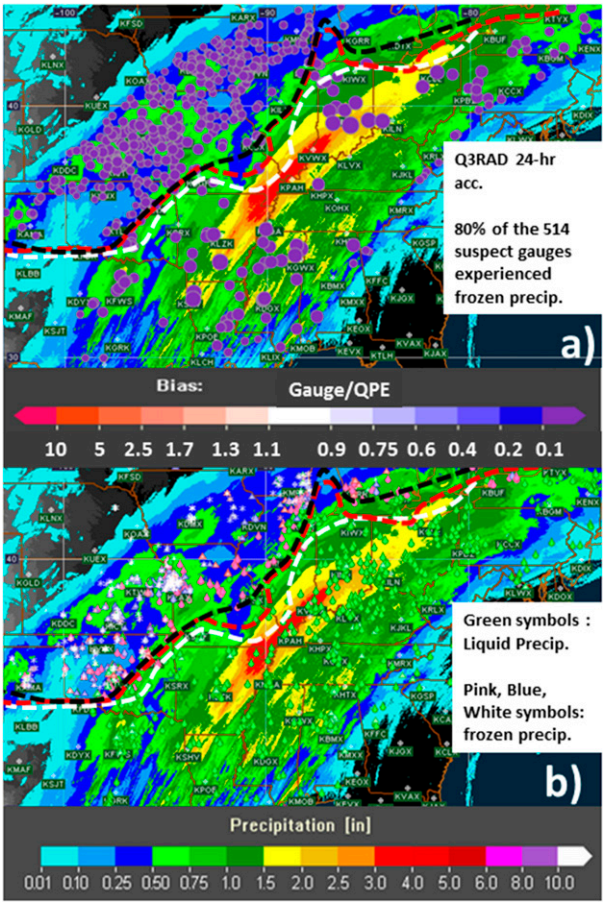

FIG. 1. (a),(b) Q3RAD 24-h QPE ending at 1200 UTC 22 Dec 2013. Suspect gauges, gauge locations with totals ≤0.25 mm (0.01 in.) and Q3RAD ≥6.4 mm (0.25 in.), denoted by purple bias ratio bubbles. Dashed lines denote the RAP model analysis 0°C surface temperature at 1800 UTC 21 Dec (black), 0000 UTC 22 Dec (red), and 1200 UTC 22 Dec 2013 (white).

white dashed oval) but was totally removed in the Q3GC estimate (Fig. 2b) because of the stuck gauges. This type of scenario was seen repeatedly during the winter, which led to the development of new MRMS gauge quality-control (QC) procedures (Martinaitis et al. 2015).

Hence, the Q3GC product was not evaluated in this study. Instead, the following radar-only products were evaluated: MRMS Q3RAD, mosaicked, Dual-Pol QPE, and the NCEP stage II radar-based precipitation estimates. A brief description of each product follows. MRMS (Zhang et al. 2016) is a centralized radar data processing system that integrates radar, rain gauge, and numerical model data to generate a range of products for operational use (Zhang et al. 2016). As part of the process, reflectivity from the contiguous U.S. Weather Surveillance Radar-1988 Doppler (WSR-88D) and Environment Canada radars are integrated into a reflectivity mosaic. If there are multiple radar inputs for a given grid point, a weighted mean based on the beam position within the melting layer, then its height and distance from radar is taken (Zhang et al. 2011, 2014). Using the mosaicked reflectivity and numerical model data, MRMS classifies radar echoes as one of seven precipitation types. For each type, a unique Z–R relation is used to calculate a rate that can be summed to create precipitation estimates (see Table 2). MRMS precipitation estimates are displayed on a 1 km × 1 km Cartesian coordinate grid.

The operational Dual-Pol QPE algorithm not only uses WSR-88D reflectivity, differential reflectivity Z_{DR}, and specific differential phase K_{DP} data to estimate precipitation, but also input from two other WSR-88D algorithms: the Hybrid Hydrometeor Classification (HHC) and the Melting Layer Detection Algorithm (MLDA; Berkowitz et al. 2013). The HHC uses a fuzzy logic approach, based on a set of processed interest fields, to determine the most likely classification for an echo at a given height above the ground and distance from the radar (Park et al. 2009). The MLDA uses initial input from sounding or numerical model data to determine the freezing level height, for example, the top of the melting layer. If radar echoes are of sufficient coverage and vertical depth, the MLDA utilizes reflectivity, differential reflectivity, and correlation coefficient (CC) to determine the likely top and bottom of the melting layer (Giangrande and Ryzhkov 2008; Park et al. 2009). A list of the precipitation classes and the relationships used to calculate the precipitation rate are shown in Table 3. Within MRMS, Dual-Pol QPE mosaics are created by summing hourly Dual-Pol QPE accumulations from the WSR-88D level III data. A "nearest neighbor" approach is used to determine which Dual-Pol quantitative precipitation estimates to assign to a grid point. There was no attempt made to smooth the Dual-Pol QPE discontinuities that resulted from such a mosaic, as boundaries between radars highlighted any radar-to-radar estimate inconsistencies that may be related to reflectivity, differential reflectivity, or HHC differences. The Dual-Pol QPE mosaic product is displayed on a 1 km × 1 km Cartesian grid. Finally, Giangrande and Ryzhkov (2008) recognized that Dual-Pol QPE would likely have challenges within the melting layer, as was the case in this study.

The NCEP stage II raw radar precipitation estimates are essentially a non-bias-adjusted mosaicked form of the legacy Precipitation Processing System (PPS) estimate that is available with WSR-88Ds. PPS uses the lowest unblocked/ground-clutter-free elevation of WSR-88D reflectivity to estimate precipitation rates via a Z–R relationship chosen by the forecaster, according to the synoptic- or mesoscale situation, that is uniformly applied to the entire radar field of view. Once PPS precipitation

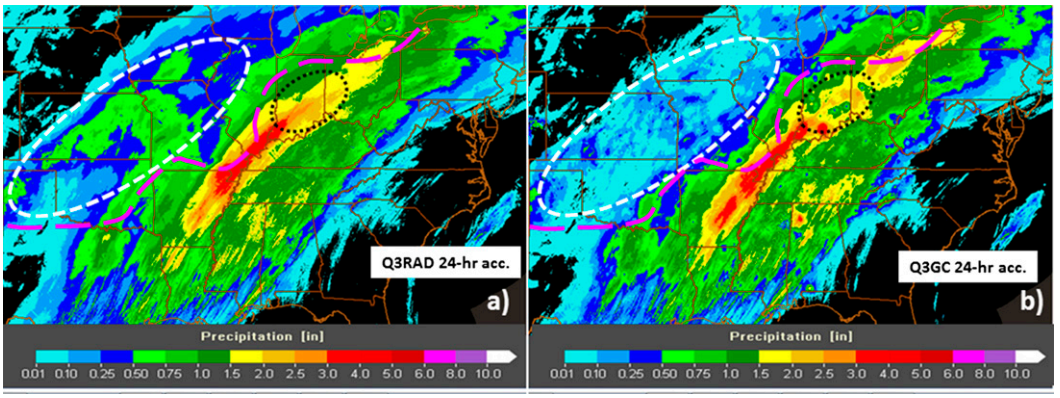

FIG. 2. (a) Q3RAD and (b) Q3GC 24-h QPE ending at 1200 UTC 22 Dec 2013. White dashed (black dotted) ovals indicate significant widespread (localized) differences between the two products caused by incorporation of suspect gauges into Q3GC. Pink long dashed line indicates RAP model surface freezing line at 1200 UTC.

estimates are created at a local WSR-88D, the data are transmitted, in real time, to NCEP. Individual WSR-88D rainfall estimate fields are merged onto a national 4 km × 4 km grid. Inputs from multiple radars for a given grid point are averaged using an inverse-distance-weighting formula (further details on the stage II mosaic process can be found at http://www.emc.ncep.noaa.gov/mmb/ylin/pcpanl/QandA/). When ingested within MRMS, stage II data are remapped to 1 km × 1 km grids. Table 4 lists Z–R relationships typically used by forecasters to estimate precipitation rates via PPS.

The performance of these three radar-only products was compared to the NCEP stage IV product, considered the benchmark for precipitation estimate studies within the hydrological community. Stage IV estimates use a combination of forecaster quality-controlled WSR-88D, satellite, and rain gauge data to create a refined rainfall estimate analysis. It is important to note that the stage IV estimates are not considered a real-time product, as there typically is a delay before it is available to ensure all relevant gauges are included in the analysis. The NCEP stage IV estimates were remapped to 1 km × 1 km grids to match the resolutions of the other products. As only 1200 UTC stage IV 24-h estimates were available within MRMS; 24-h estimates from the radar-only products were matched to 1200 UTC for comparison purposes.

Each precipitation estimate was compared to 1200 UTC 24-h accumulations from CoCoRaHS gauges that, over a number of assessments, have been found to be more consistent and suitable. As a minimal QC measure to reduce erroneous totals, both the radar estimate and the CoCoRaHS gauges were required to be ≥2.5 mm (0.10 in.) before including the R/G pair into the analysis. As most of the evaluated precipitation events spanned central (CST) and eastern standard time (EST) zones, there was a 1-h difference between the 1300 UTC (0700 CST) CoCoRaHS gauge totals and the 1200 UTC 24-h estimates. However, time offset errors affected each of the evaluated products in the same manner.

Performance assessment statistics were generated based on all available R/G pairs and by pairs stratified by 24-h gauge totals ≤6.4 mm (0.25 in.), ≤12.7 mm (0.50 in.), >12.7 mm (0.50 in.), >25.4 mm (1.0 in.), and >50.8 mm

TABLE 2. Class type and relations used to estimate precipitation for MRMS surface precipitation types east of the Rocky Mountains. Note that currently, warm and cool stratiform rain types use the same relationships. For further details on precipitation type and Z–R relationships, please see Zhang et al. (2014).

Surface precipitation type	Z–R relationships used
No echo	—
Warm stratiform rain	Max between $Z = 75(R^2)$ and $Z = 200(R^{1.6})$, capped at 50 dBZ
Cool stratiform rain	Max between $Z = 75(R^2)$ and $Z = 200(R^{1.6})$, capped at 50 dBZ
Convective rain	$Z = 300(R^{1.4})$, capped at 53 dBZ
Convective rain with hail	$Z = 300(R^{1.4})$, capped at 49 dBZ
Snow	$Z = 75(R^{2.0})$
Tropical convective	$Z = 250(R^{1.2})$ if high probability of warm rain processes; if moderate probability, a weighted mean using $Z = 250(R^{1.2})$ and $Z = 300(R^{1.4})$
Tropical stratiform	$Z = 250(R^{1.2})$ if high probability of warm rain processes; if moderate probability, a weighted mean between $Z = 250(R^{1.2})$ and $Z = 200(R^{1.6})$

TABLE 3. Class type, rain-rate factor, and complete relationship used to estimate precipitation for Dual-Pol HHC classes.

HHC class	Factor	Precipitation calculated via
Ground clutter	—	—
Unknown	—	—
No echo	—	—
Biological	—	—
Light to moderate rain	1.0	$R(Z, Z_{DR})$
Heavy rain	1.0	$R(Z, Z_{DR})$
Big drops	1.0	$R(Z, Z_{DR})$
Hail/rain below melting layer top	1.0	$R(K_{DP})$
Hail/rain above melting layer top	0.8	$0.8R(Z)$
Graupel	0.8	$0.8R(Z)$
Wet snow	0.6	$0.6R(Z)$
Dry snow (below melting layer top)	1.0	$1.0R(Z)$
Dry snow (above melting layer top)	2.8	$2.8R(Z)$
Ice crystals	2.8	$2.8R(Z)$

(2.0 in.). Only matched *R*/*G* pairs were used for the statistical assessment. For statistical measures we used the mean bias ratio, defined as the ratio of the gauge total to radar estimate, root-mean-square error (RMSE), mean absolute error (MAE), and CC to evaluate product performance. For time series analysis and diagnosing error trends, HADS hourly data were used with rigorous QC measures. Quality control was primarily applied via a set of power-law equations (Fig. 3) used to determine the maximum and minimum hourly gauge values likely for a given hourly radar QPE (Zhang et al. 2016). As an example from Fig. 3, the maximum hourly gauge total allowed for an hourly radar QPE total of 10 mm (0.39 in.) would be 35 mm (1.38 in.); a gauge total higher than 35 mm would be labeled as suspect. Hence, gauge values greater (less) than the maximum (minimum) thresholds were flagged as outliers and not used in the analysis. While not perfect, the technique was effective in identifying and removing suspect gauges in large datasets.

3. Statistical analysis and results

Figure 4 shows scatterplots and statistical results for the radar estimate gauge comparisons. As expected, the forecaster quality-controlled stage IV (Fig. 4d) estimates had the best overall performance with the lowest RMSE [8.6 mm (0.3 in.)] and MAE [5.6 mm (0.2 in.)], highest CC (0.91) and a bias ratio (1.03) closest to 1.0. Of the radar-only products, the Q3RAD estimates (Fig. 4a) had the overall next best performance with an RMSE of 13.7 mm (0.5 in.), an MAE of 9.7 mm (0.4 in.), a CC of 0.77, and a bias ratio of 1.17. While Dual-Pol QPE (Fig. 4b) had a better bias ratio (1.09) than Q3RAD, it also had a lower CC (0.62) and higher RMSE [17.8 mm (0.7 in.)] and MAE [12.8 mm (0.5 in.)] measures. Dual-Pol also exhibited the most variability, as was easily seen

TABLE 4. The *Z*–*R* relationships and recommended usage for the legacy PPS (adapted from the WSR-88D Operator Handbook; Radar Operations Center 2015).

Z–*R* relationships	Recommended use
$Z = 300(R^{1.4})$	Deep convection
$Z = 250(R^{1.2})$	Tropical convective systems
$Z = 200(R^{1.6})$	General stratiform rain
$Z = 130(R^{2.0})$	Winter stratiform and orographic rain east of Rockies
$Z = 75(R^{2.0})$	Winter stratiform and orographic rain west of Rockies

when compared to the other scatterplots. Some of this variability was likely due to differential reflectivity calibration errors, a challenge that should continue to improve as new procedures are developed to more precisely calibrate differential reflectivity (Cunningham et al. 2013; Hoban et al. 2014). However, Dual-Pol also exhibited a number of overestimated *R*/*G* pairs, particularly for gauge totals between 10 and 45 mm, which were not as prominent or absent in the other products. Subsequent examination indicated these overestimates to be related to brightband contamination of reflectivity (further discussed in section 4). MRMS applied a brightband correction to reflectivity (Zhang et al. 2016) prior to making precipitation estimates while stage IV was carefully quality controlled, both actions of which mitigated the melting layer effect for these products.

While stage II estimates (Fig. 4c) exhibited a CC (0.75) similar to Q3RAD, they also had the highest RMSE [22.4 mm (0.9 in.)] and MAE [17.0 mm (0.7 in.)] and the most distinct underestimate bias ratio (2.67) of all the products. The authors noted that stage II estimates may be different than the single-radar PPS estimates from which stage II was generated. A comparison between the two products (not shown) indicated stage II estimates often exhibited a dampened brightband contamination in the melting layer but also displayed a much stronger underestimate bias ratio. The difference is possibly a result of the stage II mosaicking scheme that takes an inverse-distance-weighted mean of multiple radar PPS estimates in overlapping regions. However, a high QPE from the nearest radar may still be dampened by relatively lower QPE values from data taken higher above ground or from farther-away radars. This would also act to dampen any brightband contamination that was consistently observed in this study.

Table 5 indicated that all the radar-only products underestimated (bias ratio >1.0) precipitation for gauge totals >12.7 mm (0.5 in.). This trend was more pronounced for the higher [>25.4 mm (1.0 in.) and >50.8 mm (2.0 in.)] gauge totals, with a corresponding increase in

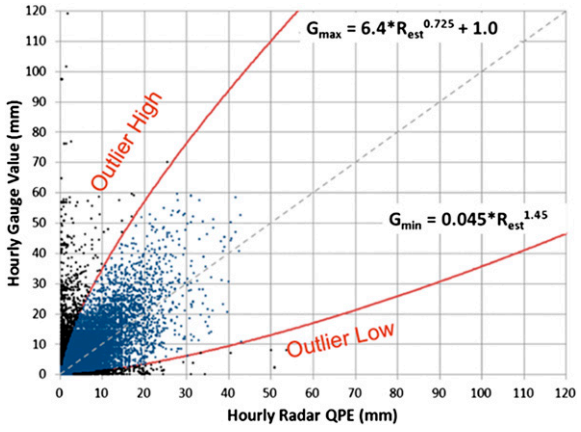

FIG. 3. Power laws used to assist in QC of *R*/*G* pairs. The upper (lower) curve represents the upper (lower) bound for gauge values for a given Q3RAD hourly total.

RMSE and MAE. This was not surprising, as radar beam overshoot is more common during the cool season because of shallower precipitation systems and lower cloud bases. This effect likely was partially mitigated by the different mosaic processes used by Q3RAD, mosaicked Dual-Pol, and the stage II products. While Q3RAD exhibited a higher tendency to underestimate than Dual-Pol, it generally had lower RMSE and MAE measures for all but the highest gauge totals [>50.8 mm (2.0 in.)] where the errors were similar (differences <1.0 mm). Overall, both Q3RAD and Dual-Pol exhibited lower RMSE and MAE than stage II for totals >12.7 mm (0.5 in.). As a contrast, the benchmark stage IV estimate exhibited RMSE that ranged from 38% to 67% lower than that of Q3RAD.

For precipitation totals ≤12.7 mm (0.5 in.), Q3RAD, Dual-Pol, and stage IV exhibited a distinct overestimate bias ratio while all of the products generally had low correlation coefficients, most below 0.30 (not shown). A significant portion of these errors was likely due to precipitation evaporating/sublimating before reaching the ground and/or gauge undercatch in snow and light rain when significant wind was present. An advancement of the seamless hybrid scan reflectivity mosaic algorithm was installed in MRMS during the spring of 2014 to help mitigate errors due to evaporation and sublimation. It compares multiple radar observations at an overlapping point and ensures the lowest radar bin has significant echoes present before coding a geographical point as having precipitation. However, it will still be unable to

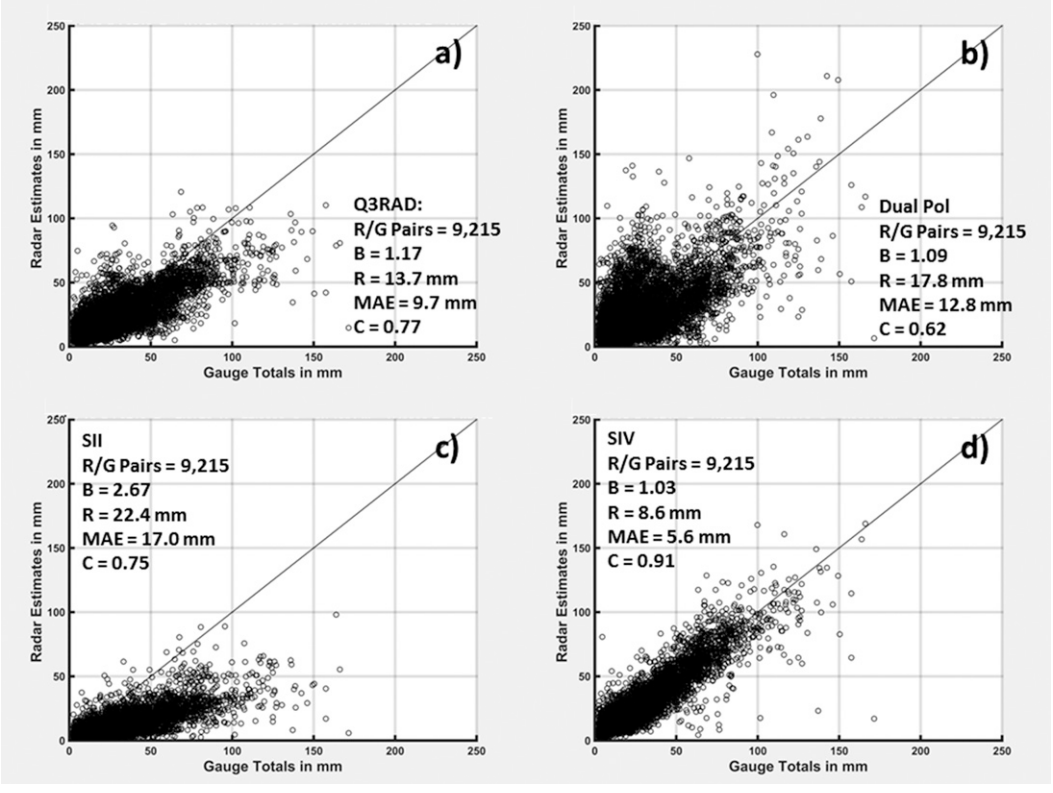

FIG. 4. The 24-h QPE from (a) Q3RAD, (b) Dual-Pol QPE, (c) stage II, and (d) stage IV vs CoCoRaHS gauges for all cool season cases. In the legend, B, R, and C denote bias ratio, RMSE, and CC, respectively. Note that the results shown used matched *R*/*G* pairs for each product.

TABLE 5. The number of *R*/*G* pairs, gauge-to-radar estimate bias ratio, RMSE, and MAE for each precipitation estimate stratified by gauge amount.

Product	24-h gauge amount	No. *R*/*G* pairs	Bias ratio	RMSE (mm)	MAE (mm)
Q3RAD	$G \leq 6.4$ mm (0.25 in.)	597	0.38	10.3	8.0
	$G \leq 12.7$ mm (0.50 in.)	2323	0.64	8.4	6.2
	$G > 12.7$ mm (0.50 in.)	6892	1.27	15.1	10.8
	$G > 25.4$ mm (1.00 in.)	3566	1.38	19.4	15.0
	$G > 50.8$ mm (2.00 in.)	1050	1.48	28.9	24.2
Dual-Pol QPE	$G \leq 6.4$ mm (0.25 in.)	597	0.36	13.3	8.9
	$G \leq 12.7$ mm (0.50 in.)	2323	0.60	12.3	7.9
	$G > 12.7$ mm (0.50 in.)	6892	1.17	19.3	14.4
	$G > 25.4$ mm (1.00 in.)	3566	1.30	22.8	18.3
	$G > 50.8$ mm (2.00 in.)	1050	1.37	29.8	24.9
Stage II	$G \leq 6.4$ mm (0.25 in.)	597	1.02	3.0	2.1
	$G \leq 12.7$ mm (0.50 in.)	2323	1.71	5.3	4.5
	$G > 12.7$ mm (0.50 in.)	6892	2.81	25.7	21.3
	$G > 25.4$ mm (1.00 in.)	3566	2.84	33.5	29.9
	$G > 50.8$ mm (2.00 in.)	1050	2.75	48.6	45.5
Stage IV	$G \leq 6.4$ mm (0.25 in.)	597	0.50	7.7	5.1
	$G \leq 12.7$ mm (0.50 in.)	2323	0.76	6.0	3.8
	$G > 12.7$ mm (0.50 in.)	6892	1.06	9.4	6.2
	$G > 25.4$ mm (1.00 in.)	3566	1.07	11.9	8.2
	$G > 50.8$ mm (2.00 in.)	1050	1.06	17.3	11.8

determine whether echoes seen at the lowest radar bin actually reach the ground, especially at farther ranges.

4. Precipitation event statistics

a. Mean bias, RMSE, and correlation for each event

Figure 5 showed the mean bias ratio for each evaluated product for each cool season precipitation event. As expected because of the forecaster quality control, stage IV exhibited a bias ratio consistently closer to one than observed with Q3RAD and Dual-Pol QPE. Stage II exhibited a distinct tendency to underestimate precipitation with mean bias ratios >2.00 for all but one event. Figure 6 showed the RMSE for each product for each precipitation event. Stage IV exhibited the lowest RMSE, less than 9.0 mm for all but two events, while stage II had the highest RMSE, generally more than 15.0 mm for all but three events. While stage IV had the lowest RMSE per event, Q3RAD had the next lowest, ~25% better than Dual-Pol. For six of the nine events, Q3RAD RMSE was comparable to the stage IV RMSE, that is, the difference between the RMSE measures was ≤6.4 mm (0.25 in.). This was significant as Q3RAD is a real-time product updated every 2 min while stage IV estimates are available after a significant time delay. Stage IV exhibited the highest CC values (not shown) for each event; Q3RAD generally exhibited the next best values followed by stage II and Dual-Pol. The generally lower correlation coefficients observed with Dual-Pol were due to the aforementioned calibration and melting layer challenges of which the following examples are notable.

b. Midwestern U.S. precipitation event of 5–6 January 2014

The combination of an intense, highly amplified upper-level storm system and relatively moist air flowing up and over an arctic cold front produced moderate to very heavy snow, with some totals as high as 380 mm (15 in.), across much of Illinois, Indiana, southern Michigan, and northwestern Ohio. However, over extreme southeastern (southwestern) sections of Illinois (Indiana),

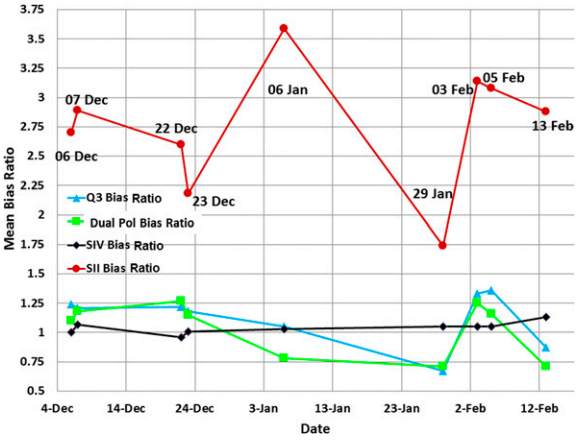

FIG. 5. Q3RAD (blue line with triangles), Dual-Pol QPE (green line with squares), stage II (red line with circles), and stage IV (black line with diamonds) mean bias ratio statistics for each cool season case evaluated.

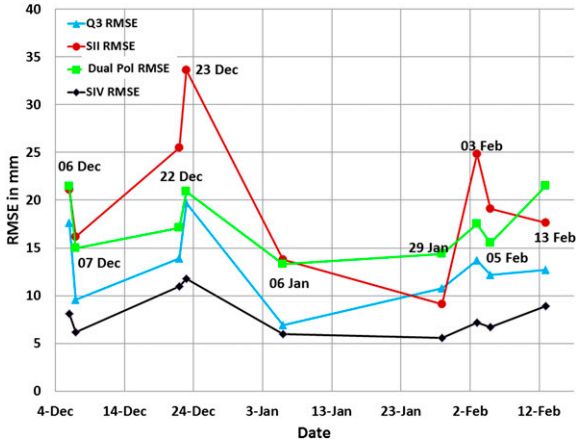

FIG. 6. As in Fig. 5, but for RMSE.

snow totals were much lower as synoptic observations indicated temperatures ranged from 0.0° to 7°C (32°–44°F) between 1200 and 1700 UTC that resulted in a variety of precipitation types across the area. The authors evaluated radar reflectivity and differential reflectivity calibration in the region of interest using the radar reflectivity comparison tool (RRCT). Gourley et al. (2003) described the tool's function, which is to match grid points between two WSR-88Ds meeting minimal horizontal, vertical, and temporal displacement criteria. Comparisons between matching grid points are made if reflectivity meets minimal QC and sample size requirements. More recently, the tool was upgraded to include Z_{DR} comparisons between two radars using the same methodology. The results give a measure of the relative Z and Z_{DR} calibration between a single radar and its neighbors. For a robust performance of Dual-Pol QPE, Z and Z_{DR} would need to be calibrated to within 1 dBZ and 0.1 dB, respectively, although for moderate to heavy rain the Z_{DR} requirement could be relaxed to 0.2 dB (Ryzhkov et al. 2005). For this event, RRCT indicated Z values were within ±1 dBZ of each other, the exception being the Evansville, Indiana, WSR-88D (KVWX), which was ~1.5 dBZ higher than neighboring radars. For Z_{DR}, most radars, where precipitation occurred, were within ±0.15 dB of each other. The exceptions were the Indianapolis, Indiana, WSR-88D (KIND) and the Wilmington, Ohio, WSR-88D (KILN), which were 0.16 dB higher and ~0.5 dB lower, respectively, than neighboring radars. So, for example, any $R(Z)$-based precipitation estimates from KVWX would likely be biased high because of the higher reflectivity. The $R(Z, Z_{DR})$ base estimates for KILN would be biased high because of the low biased Z_{DR}. Hence, Z calibration challenges contributed to the error of all the evaluated products, while Z_{DR} calibration contributed to the Dual-Pol QPE errors alone.

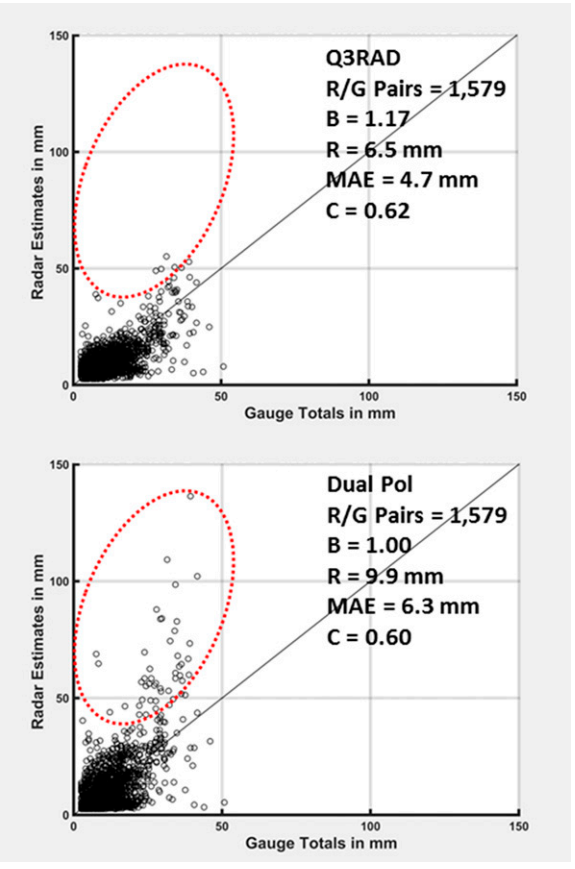

FIG. 7. (top) Q3RAD and (bottom) Dual-Pol quantitative precipitation estimates vs CoCoRaHS gauge totals for the 24-h period ending at 1200 UTC 6 Jan 2014 (legend as in Fig. 4). Dashed red oval represents region of overestimates associated with melting layer.

Figure 7 showed 24-h accumulations of Q3RAD and Dual-Pol QPE versus CoCoRaHS gauge totals. Q3RAD exhibited an underestimate bias ratio of 1.17, while the Dual-Pol QPE bias ratio was 1.0. However, Dual-Pol QPE exhibited a distinct group of overestimate R/G pairs for gauge totals between 10 and 45 mm (dashed red in Fig. 7), a pattern similar to that seen in Fig. 4b. Q3RAD exhibited much fewer overestimate R/G pairs in the same region. The evidence suggested the possibility that brightband contamination in the reflectivity data affected the Dual-Pol estimates for these R/G pairs. To determine if this was the case, 24-h accumulations of Dual-Pol QPE, Q3RAD estimates derived from reflectivity not corrected for brightband contamination ($Q3RAD_{noncor}$), and Q3RAD estimates derived from brightband-corrected reflectivity were examined and were shown in Fig. 8. The pattern seen in the Dual-Pol QPE and $Q3RAD_{noncor}$ estimates, a semicurved arc, indicated brightband contamination was present in the noncorrected reflectivity fields. Figure 9 further confirmed the presence of a bright band, as a reflectivity cross section indicated the highest

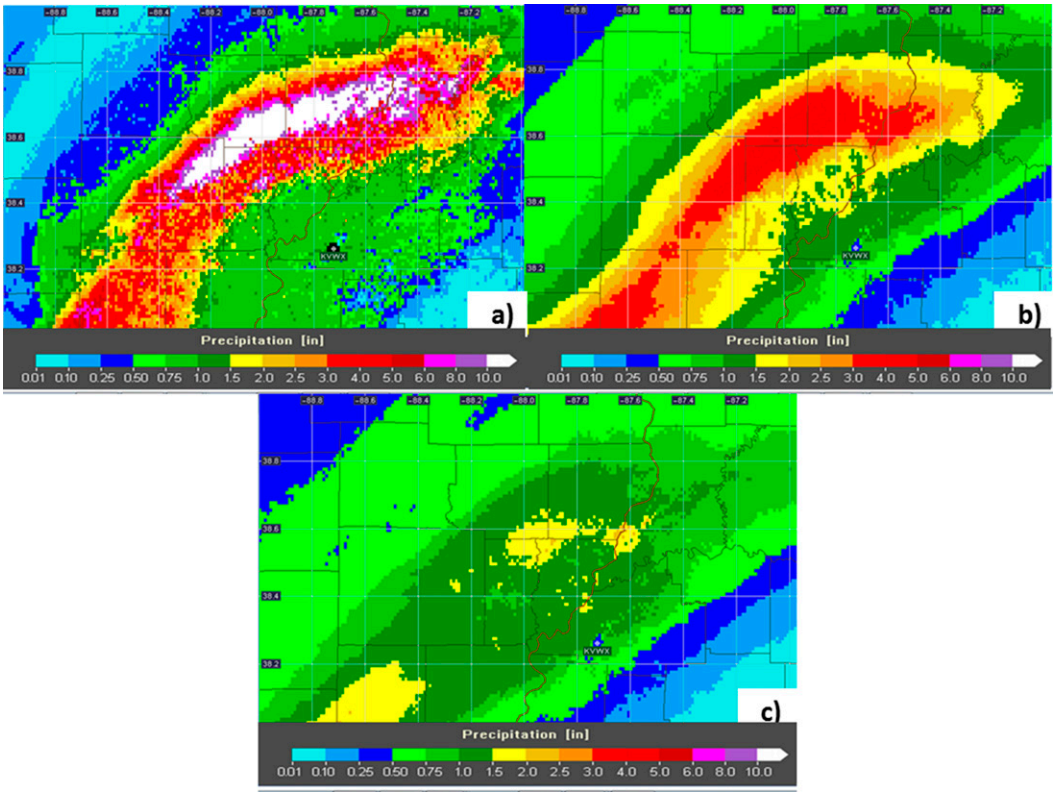

FIG. 8. (a) Dual-Pol QPE, (b) Q3RAD not corrected for brightband contamination, and (c) Q3RAD estimates for the 24-h period ending at 1200 UTC 6 Jan 2014.

Z values were generally horizontally distributed and mostly confined below 2.0 km, too low for intense convection. The band of lower CC values near the radar and the large variety (snow, sleet, and rain) of precipitation types seen in surface observations (not shown) indicated the melting level nearly extended to the ground.

Figure 10 showed the digital hybrid reflectivity (DHR), CC, HHC algorithm, and the 1-h Dual-Pol QPE for KVWX during a period where the highest 1-h Dual-Pol QPE accumulations were generated. The 1-h Dual-Pol QPE accumulations >15.2 mm (0.6 in.) were generally collocated with the region of higher reflectivity and CC values <0.96. However, Dual-Pol QPE accumulations >50.8 mm (2.0 in.) were offset north of the maximum Z/minimum CC values and lay across a region Z_{DR} values ≤0.5 dB (not shown) between 1600 and 1700 UTC. The Z_{DR} values in this region generally were in the range from 0.5 to −2.0 dB, the lower values likely being questionable. During the same period, the HHC primarily coded radar echoes as light to moderate rain, graupel, or rain/hail. Time series analysis (not shown) indicated the highest rain rates, up to 200 mm h^{-1} (7.9 in. h^{-1}), occurred for the light to moderate rain classification. The appearance of this HHC class at an altitude at the top and above the melting layer is surprising in and of itself. The available data indicated the HHC was erroneous because of an incorrectly identified melting layer from the WSR-88D MLDA. The Dual-Pol QPE equation used to estimate precipitation rates in light to moderate rain was

$$R(Z, Z_{DR}) = 0.0067(Z^{0.927})(Z_{DR}^{-3.43}), \quad (1)$$

with $R(Z, Z_{DR})$ (mm h^{-1}) and where Z (Z_{DR}) is reflectivity (differential reflectivity) in linear units. Higher Z and lower Z_{DR} values both act to raise the rain rates in Eq. (1), which is particularly sensitive for $Z_{DR} \leq 0$ dB and $Z \geq 40$ dBZ, of which both were present over the highest Dual-Pol QPE accumulations. Therefore, the much higher Dual-Pol QPE accumulations were most likely due to a combination of 1) incorrect precipitation classification around the melting layer, 2) brightband contamination within the melting layer, 3) low Z_{DR} values near the top of the melting layer, and 4) the sensitivity of Eq. (1) to factors 1–3.

Even though there appeared to be more rain than sleet and snow for the brightband region during the

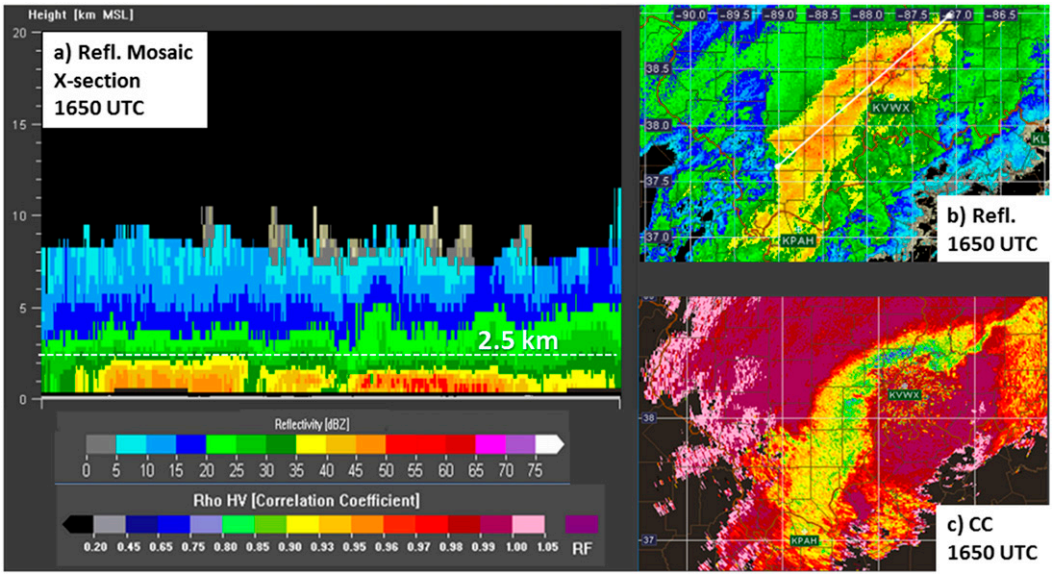

FIG. 9. (a) Reflectivity mosaic cross section, (b) 0.5° elevation angle reflectivity with white line indicating location of cross section, and (c) CC at 1650 UTC 6 Jan 2014.

hour, both automatic and CoCoRaHS gauges were at least partially impacted, particularly after 2100 UTC, when all precipitation was snow. Automatic reporting gauges were impacted the most, but CoCoRaHS data were collected by volunteers trained to handle such impacts and represent the best estimate of ground truth in the region. There were only four CoCoRaHS gauge locations over the easternmost portion of the higher

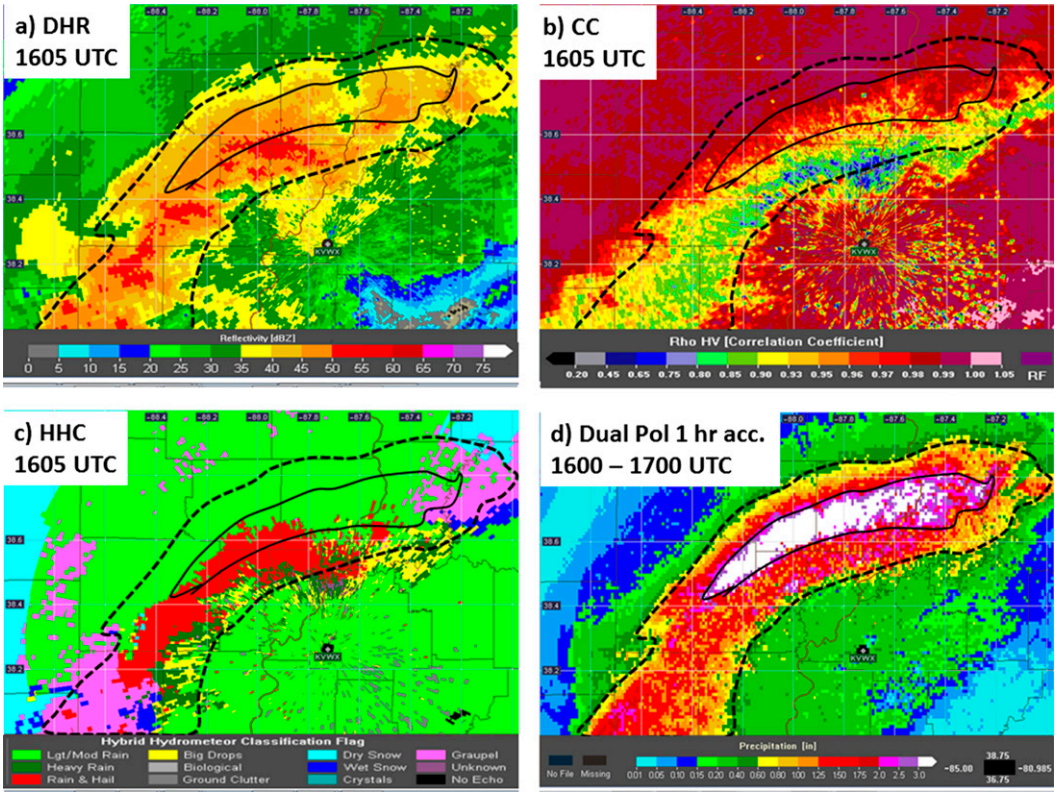

FIG. 10. (a) DHR, (b) 0.5° elevation angle CC, and (c) HHC algorithm images from KVWX at 1605 UTC 5 Jan 2014. (d) The Dual-Pol precipitation estimate is also shown for the 1-h period ending at 1700 UTC of the same date. The dashed (solid) black line approximately outlines 1-h precipitation totals ≥15.2 mm (0.6 in.) [50.8 mm (2.0 in.)].

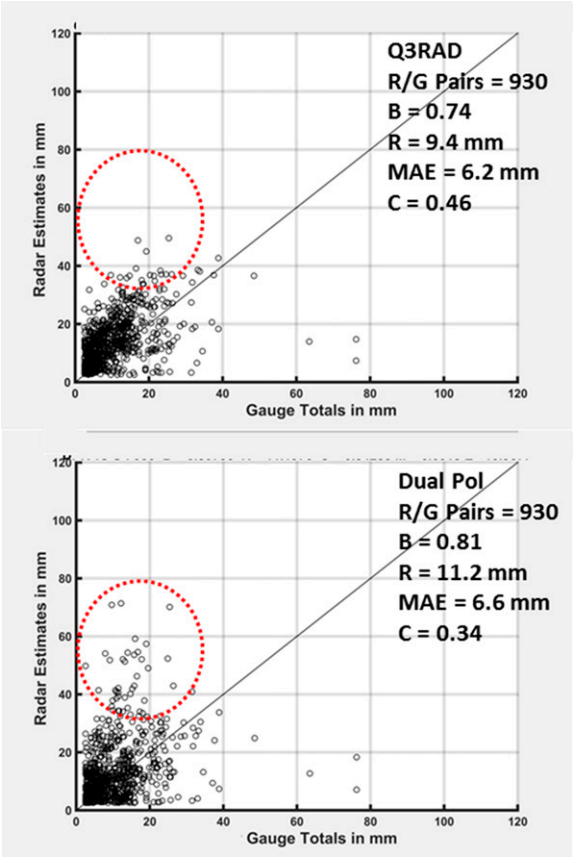

FIG. 11. As in Fig. 7, but for the 24-h period ending at 1200 UTC 29 Jan 2014.

Dual-Pol QPE accumulations, where 24-h estimates ranged from 76.0 to 137.0 mm (3.0–5.4 in.). Using these gauges, the absolute value of the 24-h Dual-Pol QPE minus gauge error averaged 69.1 mm (2.72 in.); in contrast, the absolute value of the Q3RAD minus gauge error was 5.5 mm (0.22 in.), representing a substantial reduction in the error.

c. *Southeastern U.S. precipitation event of 28–29 January 2014*

An arctic air mass was in place over much of the southeastern United States on 28 and 29 January 2014. An upper trough, jet stream dynamics, and moisture flowing over the colder air resulted in the development of freezing rain, sleet, and snow over large sections of the region. Using the RRCT, most radar Z values appeared to be within ± 1 dBZ of each other; Z_{DR} clearly indicated some sites with biases above 0.30 dB while others were between 0.15 and 0.30 dB. Ordinarily, the Z_{DR} calibration would likely affect some of the precipitation estimates; however, much of the area was below freezing at the surface, and as a result most Dual-Pol quantitative precipitation estimates did not use Eq. (1). Instead, when a radar echo was classified as dry or wet snow, ice crystals, graupel, or rain/hail (above the melting layer only), precipitation rates were calculated via a factor, dependent on the echo classification (see Table 3), multiplied by the convective Z–R relationship given by

$$R(Z) = 0.0171(Z^{0.714}), \tag{2}$$

with $R(Z)$ (mm h^{-1}) and Z (linear dBZ units). This equation is more often recognized in the following form:

$$Z = 300(R^{1.4}), \tag{3}$$

where the variables and units are the same as in Eq. (2).

Figure 11 showed the scatterplots for Q3RAD and Dual-Pol QPE; both estimates exhibited large variability, an overestimated bias ratio, and poor correlation. Similar to the previous case, Dual-Pol QPE exhibited a number of overestimates for gauge totals in the 10–35 mm range (dashed red in Fig. 11) that was not as prevalent in Q3RAD. Once again, this suggested brightband contamination affected the Dual-Pol estimates for these R/G pairs. Figure 12 showed the 24-h mosaicked estimates around the Jackson, Mississippi, WSR-88D (KDGX). The observed boundaries are from the nearest-neighbor method used to create the mosaic, with the higher totals from KDGX. The Dual-Pol quantitative precipitation estimates exhibited maximum liquid precipitation totals of 50.8–76.2 mm (2.0–3.0 in.), substantially more than Q3RAD$_{noncor}$ and Q3RAD. For this case, the Q3RAD$_{noncor}$ was likely significantly lower than the Dual-Pol quantitative precipitation estimate for two reasons. The first is related to the MRMS mosaic process used to develop the Z field that is in turn used to create the precipitation estimate: if more than one radar was available for a grid point, a weighted mean was taken (Zhang et al. 2011), which may act to smooth out spuriously high Z. Second, once the mosaic Z field is created, a unique Z–R relation is applied dependent on the radar echoes' classification. This is in contrast to what is used by the Dual-Pol quantitative precipitation estimates, which were calculated using the convective Z–R relationship [Eq. (2)] multiplied by a coefficient (Table 3) that is dependent on the HHC.

Figure 13 showed the DHR, CC, HHC algorithm, and the 1-h Dual-Pol quantitative precipitation estimate for KDGX during a period where the highest 1-h Dual-Pol QPE accumulations were generated. A calibration check of the KDGX indicated it appeared reasonably calibrated in Z, although it showed a significantly high (at least 0.4 dB) bias in Z_{DR}. The CC data clearly indicated a melting layer to the south of the radar, but precipitation had refrozen by the time it reached the

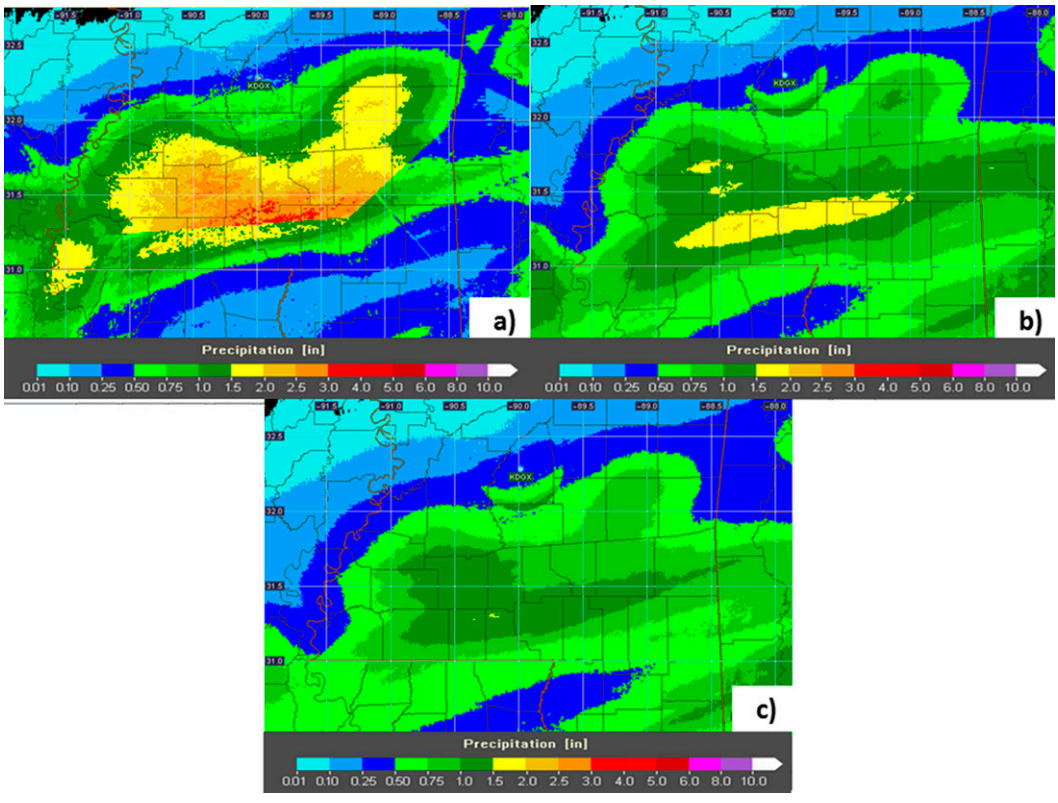

FIG. 12. As in Fig. 8, but for the 24-h period ending at 1200 UTC 29 Jan 2014.

ground, as indicated by surface observations (not shown). The HHC between 1600 and 1700 UTC classified radar echoes as either dry snow, the most common classification, or as graupel (in areas with higher reflectivity). The Dual-Pol QPE 24-h totals are quite high throughout the melting layer, likely as a result of brightband contamination and the use of the empirical relationships (Table 3).

There were five CoCoRaHS gauges within the brightband region that were available for comparison. The absolute value of the Dual-Pol QPE minus gauge 24-h accumulation error within the melting layer region averaged 33.3 mm (1.31 in.); in contrast, the absolute value of the Q3RAD minus gauge error was 14.2 mm (0.56 in.), representing a significant reduction in the error. While it is true that the CoCoRaHS gauges may have suffered some impacts such as gauge undercatch in snow and/or ice accumulation, within the catchment these are the best estimates to what actually reached the ground in the area of interest.

d. Ohio Valley/southeastern U.S. precipitation event of 2–3 February 2014

A cold front, associated with arctic air, was slowly moving eastward across the Ohio Valley and the southeastern United States on 2 and 3 February 2014. Frozen precipitation was in progress over portions of Indiana and northwestern Ohio while rain was occurring southward as moisture was isentropically lifted up and over the colder air. Another area of precipitation, farther southwest of the Ohio Valley region, developed in advance of a 500-hPa short wave. This area of precipitation moved east-northeastward into the region by the evening of 2 February, resulting in moderate to heavy rainfall in the Southeast and a variety of precipitation types over the Ohio Valley and the Appalachian Mountains. A check of radar calibration across the area indicated the Paducah, Kentucky, WSR-88D (KPAH); KILN; KVWX; and the Knoxville, Tennessee, WSR-88D (KMRX) were at least 1.0 dBZ too warm or too cool than neighboring radars. Otherwise, the rest of the radars in the region were within ±1 dBZ. With a couple of exceptions, most radar Z_{DR} values were within 0.15–0.30 dB of each other. Therefore, Z_{DR} (Z) calibration errors likely affected Dual-Pol QPE (all the estimates). Figure 14 showed the Q3RAD and Dual-Pol QPE scatterplots for this event; Dual-Pol QPE exhibited more overestimated R/G pairs for gauge totals between 10 and 45 mm (dashed red in Fig. 14) than Q3RAD for gauge totals. As with the other

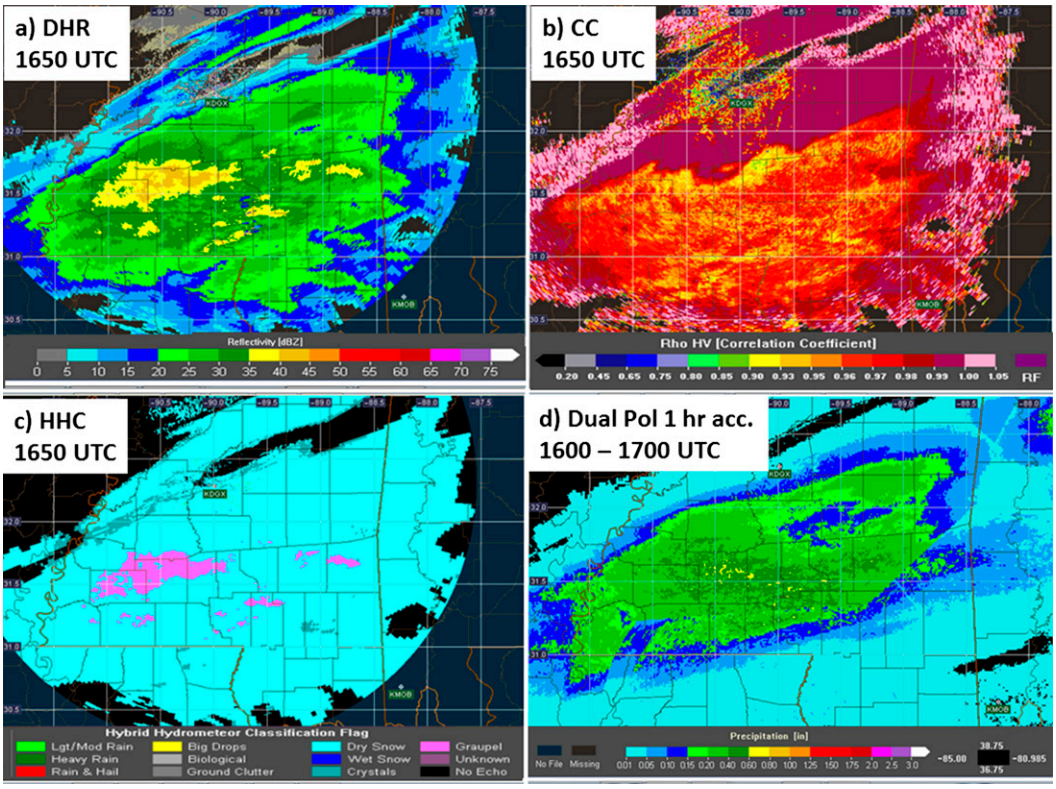

FIG. 13. (a) DHR, (b) 0.5° elevation angle CC, and (c) HHC algorithm images from KDGX at 1650 UTC 28 Jan 2014. (d) The Dual-Pol precipitation estimate is also shown for the 1-h period ending at 1700 UTC of the same date. The dashed black line approximately outlines 1-h precipitation totals ≥15.2 mm (0.6 in.).

examples, this suggested brightband contamination was the reason for the overestimated R/G pairs.

Figure 15 shows DHR, CC, HHC data, and the 1-h Dual-Pol quantitative precipitation estimate for the Jackson, Kentucky, WSR-88D (KJKL) during a time where the highest 1-h Dual-Pol QPE accumulations were being generated. This radar was reasonably calibrated with its neighbors with regards to Z with the mean Z_{DR} bias around 0.15 dB. The CC data between 0400 and 0500 UTC clearly indicated a melting layer, with values below 0.97, that is nearer to the ground in the northwest half of the radar field of view. This is corroborated by surface observations that indicated precipitation was mostly frozen in this region but mainly liquid elsewhere. The melting layer in the northwest half also coincided with higher reflectivity values; the HHC, in the region of higher Dual-Pol QPE accumulations, generally classified echoes as light to moderate rain. As mentioned previously, Eq. (1) was used to calculate precipitation estimates for this HHC. It appears the overestimates are primarily a result of brightband-contaminated reflectivity, although lower Z_{DR} values, in the 0.0–1.25 dB range, and the sensitivity of Eq. (1) to higher Z and lower Z_{DR} likely contributed as well.

A comparison was made between Dual-Pol quantitative precipitation estimates and seven available CoCoRaHS gauges within the region of greatest 24-h accumulations (the arc region). The absolute value of the Dual-Pol QPE minus gauge error within the area of interest averaged 33.8 mm (1.33 in.); in contrast, the magnitude of the Q3RAD minus gauge error was 11.5 mm (0.46 in.), a significant reduction in error. It should be noted that there were a couple of gauge sites where Q3RAD actually underestimated when compared to the CoCoRaHS 24-h totals, indicating the brightband correction applied to the reflectivity data may have been too aggressive. However, the error magnitudes observed were substantially lower than that seen with Dual-Pol QPE for each gauge.

5. Q3RAD precipitation type analysis

Using hourly gauge data, an analysis was conducted of the MRMS precipitation type contributions to the Q3RAD totals for over- and underestimated R/G pairs to better understand what may be causing some of the Q3RAD error trends seen in the statistics. MRMS uses a ''surface precipitation type'' algorithm to classify radar

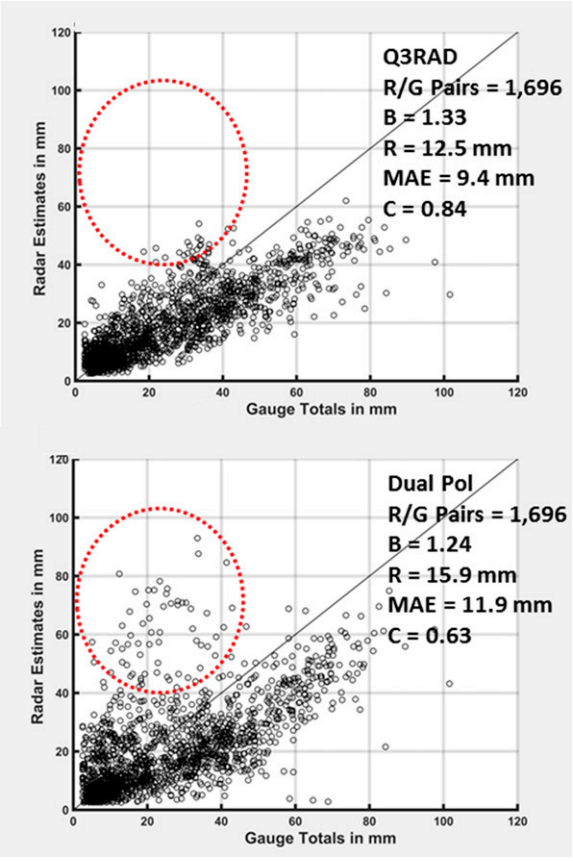

FIG. 14. As in Fig. 7, but for the 24-h period ending at 1200 UTC 3 Feb 2014.

data based on a combination of echo characteristics and model data in order to assign a unique Z–R relationship for each class (Zhang et al. 2011, 2014). There are seven possible precipitation classifications: 1) warm stratiform (WS), 2) cool stratiform (CS), 3) tropical stratiform (TS), 4) convective (CO), 5) hail (HL), 6) tropical convective (TC), and 7) snow (SN). If no radar echoes are present for a given time step, then the pixel in question is assigned the designation "no echo" (NE). To determine the importance of the stratiform and convective precipitation types to R/G pair over- and underestimated values, the various precipitation classifications were combined into three categories: stratiform (WS, CS, and TS), convective (CO, HL, and TC), and snow (SN). While most of the SN classifications were generally stratiform-like radar echoes with model temperatures indicating the surface was at or below freezing, the precipitation type was still separated out to identify trends related to gauge challenges with measuring frozen precipitation.

To determine what precipitation categories contributed most to the Q3RAD totals for under- and overestimated R/G pairs, the Q3RAD estimate (the rate divided by the number of time steps per hour) per time step for each category was calculated for all hourly R/G pairs. From this, the total Q3RAD total for each precipitation category and the percent contribution to the Q3RAD total for over- and underestimated R/G pairs could be calculated. Overall, there were more than 3.5 times as many first Standard Deviation of Radar Estimate—Gauge Errors (SDE) underestimates than overestimates in the data, confirming the Q3RAD underestimate tendency discussed earlier using CoCoRaHS data. The average percent Q3RAD contribution to the total for each precipitation category and the standard deviation of the average for the nine evaluated cases were calculated and graphed for all R/G pairs and first SDE over- and underestimates (Fig. 16).

For all R/G pairs, most of the Q3RAD contribution came from the stratiform rain categories followed by snow and convection, which is not surprising for cool season precipitation events. For first SDE overestimate error R/G pairs, most of the Q3RAD contribution came from the stratiform (~56%) and snow (~37%) categories; convection classifications contributed the least (<7%) to the Q3RAD totals. The Q3RAD tendency was generally observed for the lighter 24-h gauge totals [e.g., ≤6.4 mm (0.25 in.)] as discussed in section 3. A closer examination of 24-h accumulations for each precipitation event showed gauge totals ≤6.4 mm (0.25 in.) were generally located along the periphery of precipitation accumulations and on the cold air side of winter cyclones. An examination of hourly synoptic observations revealed drier air and significant wind speeds were often present in these regions. Surface dewpoint depressions (DD) typically ranged from 2.2° to 5.6°C (4°–10°F) and sustained wind speeds typically ranged from 5 to 10 m s^{-1} [10–20 knots (kt; 1 kt = 0.51 m s^{-1})]. Under those conditions, not only would there be some evaporation/sublimation of falling precipitation, but also significant gauge undercatch effects (Rasmussen et al. 2012; Goodison and Yang 1996) due to the wind, both of which would cause Q3RAD overestimates. In some of the overestimate cases examined, precipitation developed and fell into a layer of cold, dry air, allowing substantial evaporation or sublimation to occur. Figure 17 shows an example from a winter storm that developed over the Southeast on 28 and 29 January 2014. The 1307 UTC surface observations clearly showed dewpoint depressions of 11.1°–16.7°C (20°–30°F) over the region of interest. As precipitation began to develop, some of it evaporated before reaching the ground. The echoes seen in the radar cross section are relatively weak, with some spots suggesting little reaching the ground. The time series for two gauges along the cross section indicated over 2.5 mm (0.10 in.) of Q3RAD precipitation were generated from the radar

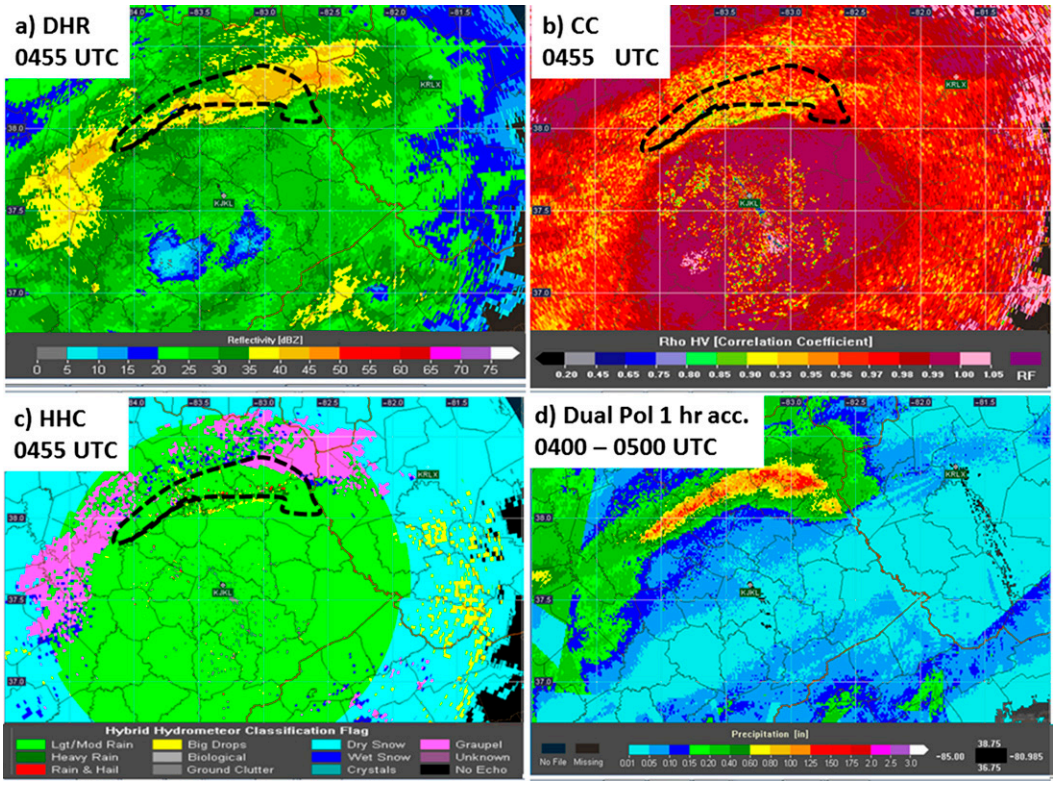

FIG. 15. (a) DHR, (b) 0.5° elevation angle CC, and (c) HHC algorithm images from KJKL at 0455 UTC 03 Feb 2014. (d) The Dual-Pol precipitation estimate is also shown for the 1-h period ending at 0500 UTC of the same date. The dashed black line approximately outlines 1-h precipitation totals ≥15.2 mm (0.6 in.).

data, but none were recorded by the gauges. Observations indicated temperatures remained above freezing at Upatoi Creek at Fort Benning (UPAG1) [Macon (MCN)] until approximately 1800 UTC 28 February (0000 UTC 29 February), so gauge impacts due to frozen precipitation were unlikely. Note the Rapid Refresh (RAP) analysis temperature and dewpoint spread decreased until the gauge reported precipitation indicating atmospheric moistening via evaporation (sublimation) of rain drops (snow) was likely taking place. In this same case, there were other locations where the wind was at or above 5 m s^{-1} (10 kt) with light rain, freezing rain, or snow occurring, so gauge undercatch likely played a significant role in the overestimates as well. It is important to note that with the cases examined in this study, overestimates were found mainly in the cold air side of winter cyclones. This will not always be the case, as radar quantitative precipitation overestimates are always possible whenever layers of relatively dry air are present between weak radar echoes and the ground and/or significant surface wind speeds are present. Other possible causes for Q3RAD overestimates would be *Z*–*R* limitations, not enough correction applied to brightband contamination, and improper radar calibration.

For first SDE underestimated *R*/*G* pairs, the chief contribution to the Q3RAD totals were from the stratiform category (82.8%), followed by snow (8.6%) and convection (8.6%). Further, the CS precipitation type contributed the most to the underestimate errors. A significant amount of the error was likely due to the radar beam partially overshooting the generally lower cloud bases and shallower precipitation systems found during the wintertime, an example of which is illustrated in Fig. 18. There were few underestimated *R*/*G* pairs (Fig. 18, black dots) from northwestern Arkansas to northwestern Tennessee and north of the Ohio River, where gauges were impacted by frozen precipitation. In the warmer air, underestimate *R*/*G* pairs were quite prevalent, a number of which were clustered in regions where the bottom of the radar beam is at least 1 km above ground level. Analyses for the events on 3 and 13 February 2014 indicated similar tendencies (not shown). To get a more quantitative look at radar beam overshoot, the percentage of $R - G < 0$ (underestimates) *R*/*G* pairs was plotted as a function of the seamless hybrid scan reflectivity beam bottom height (H-SHSR) above the ground (Fig. 19). The percentage of underestimates increased linearly from approximately

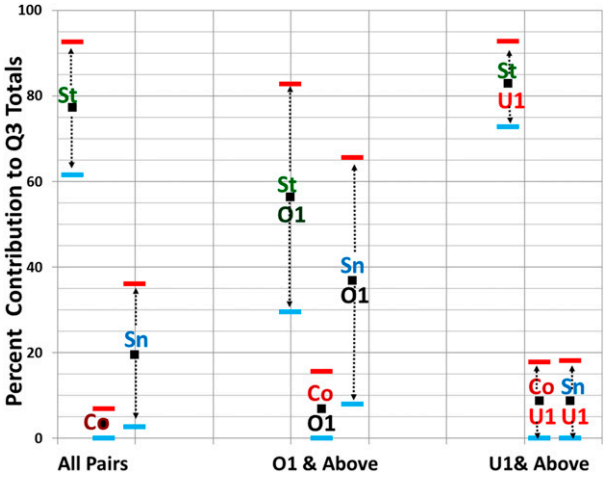

FIG. 16. Percent contribution to Q3RAD totals for all R/G pairs and the first SDE over- (O_1) and underestimated (U_1) R/G pairs. Red and light blue horizontal hashes mark the first SDE uncertainty of the nine-case average. St, Co, and Sn denote stratiform, convective, and snow categories, respectively.

45% at 0.25 km to near 70% at 2.0 km, showing a clear dependency on increasing beam height/distance from radar.

While some underestimates may be a result of partial beam blockage, there are also clear examples where unrepresentative Z–R relationships likely play a part in the error. Once again, an examination of each precipitation event revealed that underestimates ≥25.4 mm (1.0 in.) primarily occurred over a baroclinic zone along and behind a frontal region. For six of the events, surface observations indicated moderate to heavy rain generally occurred in temperatures ranging from 2.0° to 12.0°C (34.0°–55°F) under saturated conditions (relative humidity >90%). Figures 20a–d showed time series of four gauges, all of which were located in a region of heavy precipitation that fell over portions of Missouri, Kentucky, and Indiana on 21–22 December 2013. The time series showed hourly gauge, Q3RAD, Q3RAD not corrected for brightband contamination ($Q3RAD_{noncor}$), and the surface precipitation type. The gauges chosen each had over 85 mm (3.3 in.) of precipitation measured over the 24-h period. Radar (not shown) indicated primarily stratiform rain with brief periods of convection. While brightband corrections were made for Q3RAD in Figs. 20b and 20d, the $Q3RAD_{noncor}$ was still substantially less than the hourly gauge totals. Overall, the Q3RAD and $Q3RAD_{noncor}$ estimates were substantially less than the hourly gauge totals and illustrated that the Z–R relationships used to estimate rain

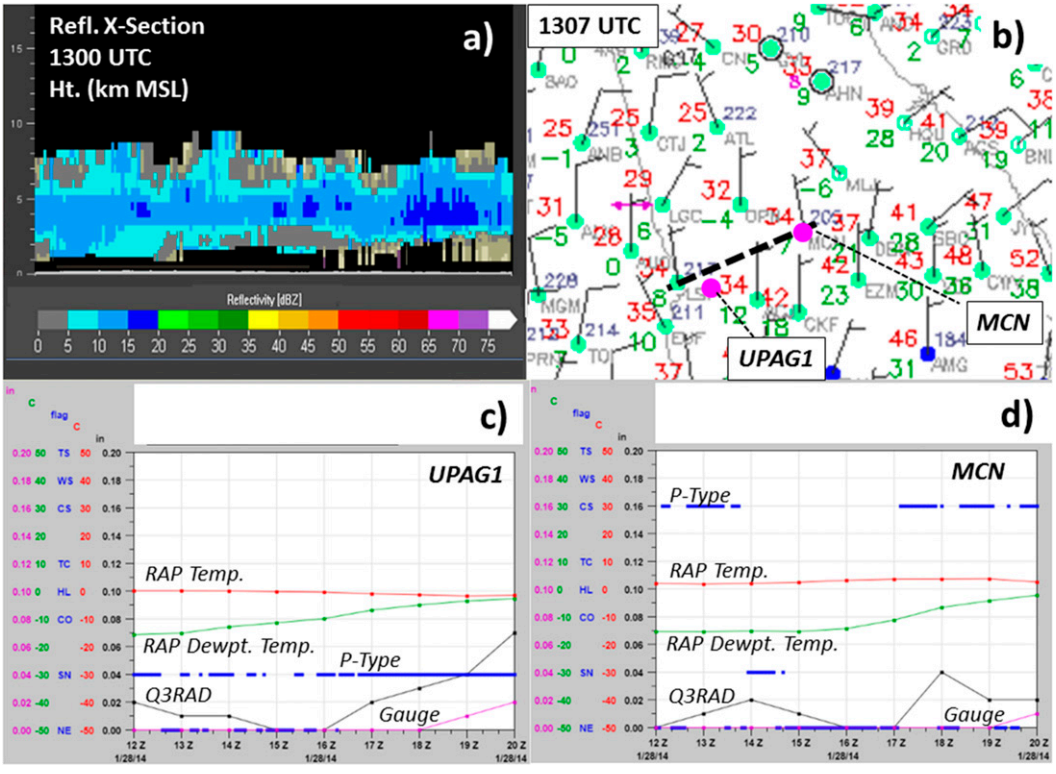

FIG. 17. (a) Reflectivity cross section, (b) 1307 UTC surface synoptic observations, and time series for the gauges (c) UPAG1 and (d) MCN for the 24-h period ending at 1200 UTC 29 Jan 2014. Precipitation type is given by blue dots, gauge totals in purple, and Q3RAD in black. The RAP analysis surface temperature and dewpoint are given in red and green, respectively. Approximate location of reflectivity cross section in (a) is given by dashed black line in (b).

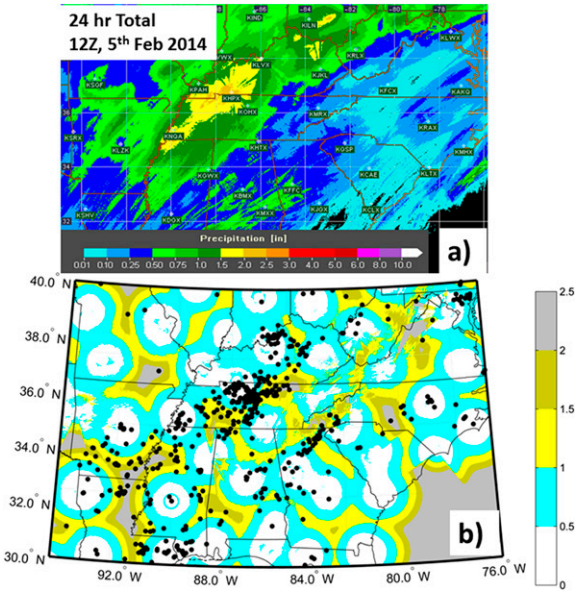

FIG. 18. (a) The 24-h Q3RAD accumulation and (b) the height of the bottom of the radar beam (km) with locations of underestimates of at least the first SDE (black dots) for the period ending at 1200 UTC 5 Feb 2014.

rates in this region were not representative of the microphysical environment present. In addition to beam overshoot and Z–R limitations, overcorrection of bright-band reflectivity and/or improper radar calibration can cause underestimates. Finally, thawing gauges can cause underestimates via gauges previously stuck or clogged with ice and snow within the gauge orifice beginning to melt as temperatures rise and/or new precipitation beginning to fall. The thawing frozen precipitation causes tipping-bucket gauges to report a precipitation total more than what actually fell [see Martinaitis et al. (2015) for an example].

6. Conclusions

Examination of nine weather events east of the Rocky Mountains quickly revealed challenges in evaluating radar precipitation estimates during the cool season. Analysis showed a large number of automatic gauges likely became stuck in freezing temperatures when frozen precipitation was present. Since stuck gauges adversely impacted local gauge bias-adjusted radar QPE, this study only evaluated radar-only precipitation estimates. Using quality-controlled gauges and the NCEP stage IV analysis as a benchmark for comparison purposes, Q3RAD, Dual-Pol QPE, and stage II exhibited a tendency to underestimate precipitation, a trend more pronounced for higher gauge totals. A significant portion of this error could be attributed to radar beam overshoot. Stage II had the highest RMSE and MAE and the strongest underestimate tendency of all the products. Dual-Pol QPE had a slightly better bias ratio than Q3RAD, but it also had a higher RMSE and MAE

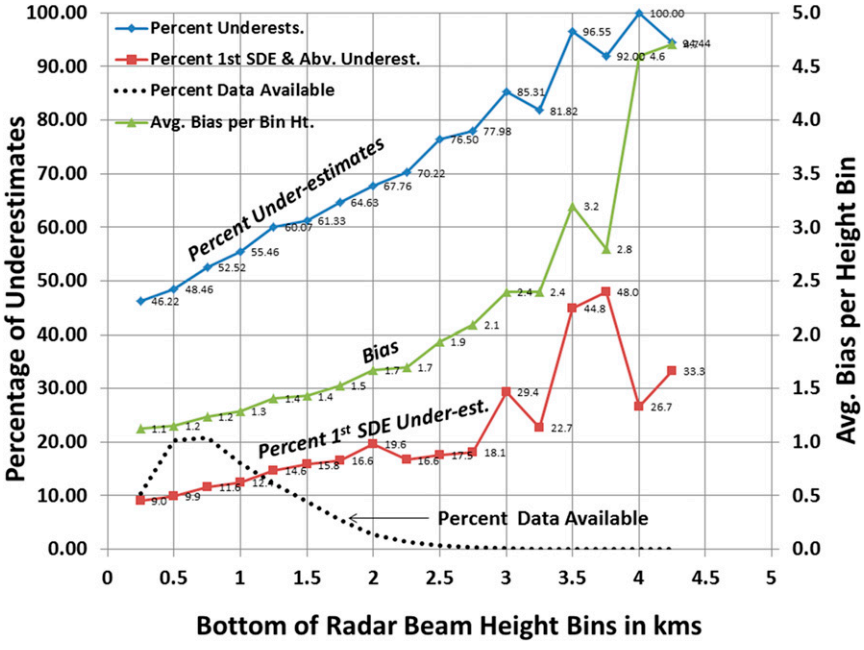

FIG. 19. Percentage of Q3RAD underestimates (blue), first SDE underestimates (red), and gauge-to-radar estimate mean bias ratio (green) as a function of radar beam bottom height. Dotted black line indicates percentage of the data available for a given height bin.

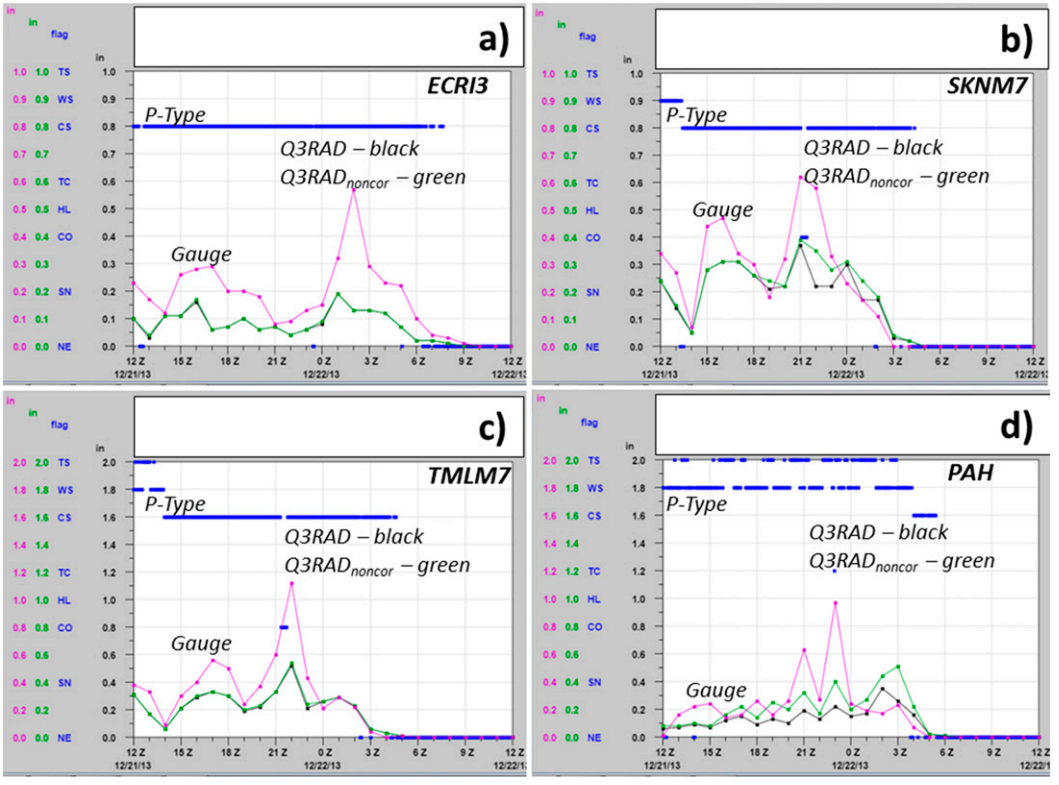

FIG. 20. Time series of the gauges (a) ECRI3, (b) SKNM7, (c) TMLM7, and (d) PAH for the 24-h period ending at 1200 UTC 22 Dec 2013. Precipitation type is given by blue dots, gauge totals in purple, Q3RAD in black, and $Q3RAD_{noncor}$ in green. Note that there is essentially no difference between Q3RAD and $Q3RAD_{noncor}$ in (a) and (c).

as well as lower correlation coefficient values. Q3RAD had the overall best performance of the three radar-only products. For lighter 24-h precipitation totals Q3RAD, Dual-Pol, and even stage IV exhibited a distinct overestimate bias, likely related to evaporation/sublimation of precipitation before reaching the ground and/or gauge undercatch.

Event by event statistics showed both Dual-Pol and Q3RAD bias ratios stayed fairly close, generally within 0.25, to the stage IV benchmark. Q3RAD tended to have lower RMSE and MAE than Dual-Pol and stage II. Further, Q3RAD RMSE was comparable to the stage IV RMSE [e.g., the difference $\leq$6.4 mm (0.25 in.)] for six of the events evaluated. Dual-Pol RMSE and MAE and lower correlation coefficient values were partly related to both Z and Z_{DR} calibration errors; however, a significant portion of the error was related to challenges regarding brightband contamination in the melting layer, of which some notable examples were shown. These challenges were anticipated, and work continues today on developing new techniques to further improve the Dual-Pol QPE system (Ryzhkov et al. 2014).

Analysis of MRMS precipitation classification contributions to Q3RAD totals for overestimate R/G pairs indicated the stratiform (chiefly the cool stratiform type) and snow categories contributed the highest percentages. A significant portion of this error was related to precipitation evaporating or sublimating prior to reaching the ground and to gauge undercatch. Other factors such as Z–R limitations and improper radar calibration also played a role. Analysis indicated the cool stratiform type also was associated with underestimate errors, the primary factors being radar beam overshoot and Z–R limitations, although other factors are significant as well.

There are some future investigations that are worthy of mentioning as they are related to improvement of precipitation estimates. First, an analysis of radar-only QPE performance east of the Rockies during the warm season as well as analysis for west of the Rockies is currently in progress. Second, work is underway to integrate Dual-Pol information into MRMS Q3RAD precipitation estimates via the use of specific attenuation (Ryzhkov et al. 2014; Wang et al. 2014). Third, a future project will likely examine the feasibility of improving Dual-Pol quantitative precipitation estimates in the

melting layer by incorporating a brightband correction to WSR-88D reflectivity, similar to what has been implemented in MRMS (Zhang and Qi 2010; Zhang et al. 2011). Finally, a probabilistic approach to QPE, similar to Kirstetter et al. (2015), will be further explored in the effort to further improve MRMS precipitation estimates.

Acknowledgments. We thank Alexander Ryzhkov and Dusan Zrnić and the anonymous reviewers for providing many useful comments that improved the manuscript. Funding was provided by NOAA/Office of Oceanic and Atmospheric Research under NOAA–University of Oklahoma Cooperative Agreement NA11OAR4320072, U.S. Department of Commerce, as well as the Radar Operations Center technical transfer memorandum of agreement.

REFERENCES

Berkowitz, D. S., J. A. Schultz, S. Vasiloff, K. L. Elmore, C. D. Payne, and J. B. Boettcher, 2013: Status of dual pol QPE in the WSR-88D network. *27th Conf. on Hydrology*, Austin, TX, Amer. Meteor. Soc., 2.2. [Available online at https://ams.confex.com/ams/93Annual/webprogram/Paper221525.html.]

Cunningham, J. G., W. D. Zittel, R. R. Lee, and R. L. Ice, 2013: Methods for identifying systematic differential reflectivity biases on the operational WSR-88D network. *36th Conf. on Radar Meteorology*, Breckenridge, CO, Amer. Meteor. Soc., 9B.5. [Available online at https://ams.confex.com/ams/36Radar/webprogram/Paper228792.html.]

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.

Fiebrich, C. A., C. R. Morgan, A. G. McCombs, P. K. Hall Jr., and R. A. McPherson, 2010: Quality assurance procedures for mesoscale meteorological data. *J. Atmos. Oceanic Technol.*, **27**, 1565–1582, doi:10.1175/2010JTECHA1433.1.

Giangrande, S., and A. Ryzhkov, 2008: Estimation of rainfall based on the results of polarimetric echo classification. *J. Appl. Meteor.*, **47**, 2445–2462, doi:10.1175/2008JAMC1753.1.

Goodison, B. E., and D. Yang, 1996: In situ measurements of solid precipitation in high latitudes: The need for correction. *Proc. Workshop on the ACSYS Solid Precipitation Climatology Project*, WMO/TD-739, WCRP-93, Reston, VA, World Climate Research Programme, 3–17.

Gourley, J. J., B. Kaney, and R. A. Maddox, 2003: Evaluating the calibrations of radars: A software approach. Preprints, *31st Int. Conf. on Radar Meteorology*, Seattle, WA, Amer. Meteor. Soc., 459–462. [Available online at https://ams.confex.com/ams/32BC31R5C/techprogram/paper_64171.htm.]

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.

Hoban, N. P., J. G. Cunningham, and D. Zittel, 2014: Estimating Systematic WSR-88D differential reflectivity biases using Bragg Scattering. *30th Conf. on Environmental Information Processing Technologies*, Atlanta, GA, Amer. Meteor. Society, 2. [Available online at https://ams.confex.com/ams/94Annual/webprogram/Paper237404.html.]

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.

Kirstetter, P. E., J. J. Gourley, Y. Hong, J. Zhang, S. Moazamigoodarzi, C. Langston, and A. Arthur, 2015: Probabilistic precipitation rate estimates with ground-based radar networks. *Water Resour. Res.*, **51**, 1422–1442, doi:10.1002/2014WR015672.

Krajewski, W. F., G. Villarini, and J. A. Smith, 2010: Radar-rainfall uncertainties: Where are we after thirty years of effort? *Bull. Amer. Meteor. Soc.*, **91**, 87–94, doi:10.1175/2009BAMS2747.1.

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

Martinaitis, S. M., 2008: Effects of multi-sensor radar and rain gauge data on hydrologic modeling in relatively flat terrain. M. S. thesis, Dept. of Meteorology, Florida State University, 99 pp.

——, S. Cocks, Y. Qi, B. 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, 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.]

Park, H. S., A. V. Ryzhkov, D. S. Zrnić, and K.- E. Kim, 2009: The hydrometeor classification algorithm for the polarimetric WSR-88D: Description and application to an MCS. *Wea. Forecasting*, **24**, 730–748, doi:10.1175/2008WAF2222205.1.

Radar Operations Center, 2015: Guidance on adaptable parameters. WSR-88D Operator Handbook, Vol. 1, RPG, NOAA, U.S. Department of Commerce. [Available online at https://www.roc.noaa.gov/wsr88d/Program/OperationsManuals.aspx.]

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, doi:10.1175/BAMS-D-11-00052.1.

Ryzhkov, A. V., S. E. Giangrande, V. M. Melnikov, and T. J. Schuur, 2005: Calibration issues of dual-polarization radar measurements. *J. Atmos. Oceanic Technol.*, **22**, 1138–1155, doi:10.1175/JTECH1772.1.

——, M. Diederich, P. Zhang, and C. Simmer, 2014: Potential utilization of specific attenuation for rainfall estimation, mitigation of partial beam blockage, and radar networking. *J. Atmos. Oceanic Technol.*, **31**, 599–619, doi:10.1175/JTECH-D-13-00038.1.

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.

Smith, J. A., D. J. Seo, M. L. Baeck, and M. D. Hudlow, 1996: An intercomparison study of NEXRAD precipitation estimates. *Water Resour. Res.*, **32**, 2035–2046, doi:10.1029/96WR00270.

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.

Tang, L., J. Zhang, C. Langston, J. Krause, K. Howard, and V. Lakshmanan, 2014: A physically based precipitation/nonprecipitation radar echo classifier using polarimetric and environmental data in a real-time national system. *Wea. Forecasting*, **29**, 1106–1119, doi:10.1175/WAF-D-13-00072.1.

Wang, Y., P. Zhang, A. Ryzhkov, J. Zhang, and P. Chang, 2014: Utilization of specific attenuation for tropical rainfall estimation in complex terrain. *J. Hydrometeor.*, **15**, 2250–2266, doi:10.1175/JHM-D-14-0003.1.

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.

Zhang, J., and Y. Qi, 2010: A real-time algorithm for the correction of brightband effects in radar-derived QPE. *J. Hydrometeor.*, **11**, 1157–1171, doi:10.1175/2010JHM1201.1.

——, 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.

——, and Coauthors, 2012: Radar Quality Index (RQI)—A combined measure for 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, 2014: Initial operating capabilities of quantitative precipitation estimates in the Multi-Radar Multi-System. *28th Conf. of Hydrology*, Atlanta, GA, Amer. Meteor. Soc., 5.3. [Available online at https://ams.confex.com/ams/94Annual/webprogram/Paper240487.html.]

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