Using Ancillary Information from Radar-Based Observations and Rain Gauges to Identify Error and Bias

BRIAN R. NELSON,[a] OLIVIER P. PRAT,[b] AND RONALD D. LEEPER[b]

[a] *NOAA/NESDIS/National Centers for Environmental Information, Asheville, North Carolina*
[b] *North Carolina Institute for Climate Studies, North Carolina State University, Asheville, North Carolina*

(Manuscript received 31 July 2020, in final form 11 February 2021)

ABSTRACT: Ancillary information that exists within rain gauge and radar-based datasets provides opportunities to better identify error and bias between the two observing platforms as compared to error and bias statistics without ancillary information. These variables include precipitation type identification, air temperature, and radar quality. There are two NEXRAD-based datasets used for reference: the National Centers for Environmental Prediction (NCEP) Stage IV and the NOAA NEXRAD Reanalysis (NNR) gridded datasets. The NCEP Stage IV dataset is available at 4 km hourly and includes radar–gauge bias adjusted precipitation estimates. The NNR dataset is available at 1 km at 5-min and hourly time intervals and includes several different variables such as reflectivity, radar-only estimates, precipitation flag, radar quality indicator, and radar–gauge bias adjusted precipitation estimates. The NNR data product provides additional information to apply quality control such as identification of precipitation type, identification of storm type and Z–R relation. Other measures of quality control are a part of the NNR data product development. In addition, some of the variables are available at 5-min scale. We compare the radar-based estimates with the rain gauge observations from the U.S. Climate Reference Network (USCRN). The USCRN network is available at the 5-min scale and includes observations of air temperature, wind, and soil moisture, among others. We present statistical comparisons of rain gauge observations with radar-based estimates by segmenting information based on precipitation type, air temperature, and radar quality indicator.

KEYWORDS: Precipitation; Radars/Radar observations

1. Introduction

The accepted procedure to measure the quality of radar rainfall estimates has typically been to compare rain gauge data with radar-based rainfall estimates using the rain gauge as the de facto observation. Historical process has dictated that the rain gauge estimate is considered truth, and common practice has been to compare the gridded (Cartesian or polar) radar rainfall product (reflectivity, rainfall, etc.) with the point measurement at the corresponding temporal scale. Some examples of this process can be seen in Smith et al. (1996), Young et al. (1999), and Nelson et al. (2016). These methods provide a quantitative process to identify errors in the radar-based product. Caveats in these methods are provided in the studies, and potential problems mostly relating to point to pixel comparisons are outlined in studies such as Kitchen and Blackall (1992), Ciach and Krajewski (1999), and Habib et al. (2004). The caveats have to do with subgrid variability of the rainfall process at various temporal scales and the fact that comparing a point estimate to an area estimate introduces errors. Kirstetter et al. (2015) built a probabilistic model at fine scales to evaluate the error in the rainfall estimates from the Multi-Radar Multi-Sensor QPE system by precipitation type. They also used the National Weather Service (NWS) Hydrometeorological Automated Data System (HADS) in situ data for comparisons. Another study by Chen et al. (2013) compared the National Mosaic Quantitative Precipitation Estimation (Q2) product to the NCEP Stage IV project over the conterminous United States (CONUS). Chen et al. (2013) present results by season and location [i.e., River Forecast Center (RFC)] as well as striating the error based on the radar quality indicator by location. In this study we approach the method of describing errors using rain gauges as the reference, but we also use variables that are available as ancillary information to try and characterize the errors more completely. We attempt to build on the previous studies by presenting information using precipitation type, air temperature, and radar quality indicator in comparisons with the U.S. Climate Reference Network (USCRN). NOAA's NEXRAD Reanalysis (NNR) Climate Data Record has six variables at 1-km resolution; hourly radar–gauge estimates, hourly radar-only estimates, 5-min radar rain rates, 5-min precipitation flags, 5-min reflectivity mosaic, and 5-min radar quality indicator (Nelson 2017). In this study we also have available to us the USCRN dataset, which includes the following variables: 5-min precipitation estimate, soil moisture, air temperature, wind speed, and an hourly rain gauge estimate. The goal of this study is to use the ancillary information to separate cases, calculate errors for each case and characterize errors in a different manner than what has typically been done. This paper is organized as follows: section 2 provides an overview of the methodology and data products, section 3 shows results, and section 4 provides conclusions and recommendations.

2. Data and methodology

a. *NOAA's NEXRAD reanalysis climate data record*

NOAA's NEXRAD Reanalysis (NNR) Climate Data Record (CDR) generates six variables related to quantitative precipitation

Corresponding author: Brian R. Nelson, brian.nelson@noaa.gov

DOI: 10.1175/JHM-D-20-0193.1

TABLE 1. Precipitation type identifiers from NOAA's NEXRAD reanalysis.

Value	Precipitation type
−1	Missing
0	Zero
1	Stratiform/warm stratiform
2	Bright band
3	Snow
4	Overshoot
6	Convective
7	Hail
9	Warm rain
10	Cool stratiform
91	Tropical convective
96	Tropical stratiform

estimation over a defined period (2002–11). All variables are 1-km spatial resolution. The variables are MOS2D, a two-dimensional gridded reflectivity; PRATE, an uncorrected precipitation rate; ROQPE, a radar-only quantitative precipitation estimate; GCQPE, a gauge corrected quantitative precipitation estimate; PFLAG, a precipitation flag; and RQIND, a radar quality indicator. ROQPE and GCQPE are provided at hourly scale and MOS2D, PRATE, PFLAG, and RQIND are provided at 5-min scale. For in-depth details about the products, see Nelson (2017).

Of particular interest in NNR are the hourly precipitation estimates (gauge-corrected and radar-only) as well as the precipitation flag and 5-min radar rain rates. Hourly radar–gauge, radar-only estimates, and 5-min rain rates are generated by the Multi-Radar Multi-Sensor algorithm (Zhang et al. 2016).

The precipitation flag provides information on the type of precipitation that is estimated (Zhang et al. 2016) such as snow, hail, convective, stratiform, etc. as well as no rain cases. Table 1 provides the types of flags that are generated. In this work we use all of the defined variables with the exception of the MOS2D. A brief overview of each variable is given here.

1) ROQPE

Radar-only precipitation rates are obtained by applying Z–R relationships to the mosaicked hybrid scan reflectivity field pixel by pixel. Zhang et al. (2011) provide the overview of precipitation rate generation.

2) GCQPE

Bias correction of radar-only QPEs is based on an additive radar rainfall error model. The details can be found in Zhang et al. (2011, 2014). Rain gauge observations used in this blending procedure are mostly taken from HADS.

3) PRATE

The algorithm to generate precipitation rate is performed at the 5-min scale using the mosaicked two-dimensional reflectivity and the Z–R relation defined in ROQPE.

4) PFLAG

Multiple precipitation regimes often coexist within a single radar site. An automated precipitation classification based on the 3D radar reflectivity structure and atmospheric environmental data is used. The classification of precipitation regimes consists of a series of physically based heuristic rules (see Zhang et al. 2011; Grams et al. 2014; Xu et al. 2008; Qi et al. 2013). Each grid point is assigned a precipitation type based on

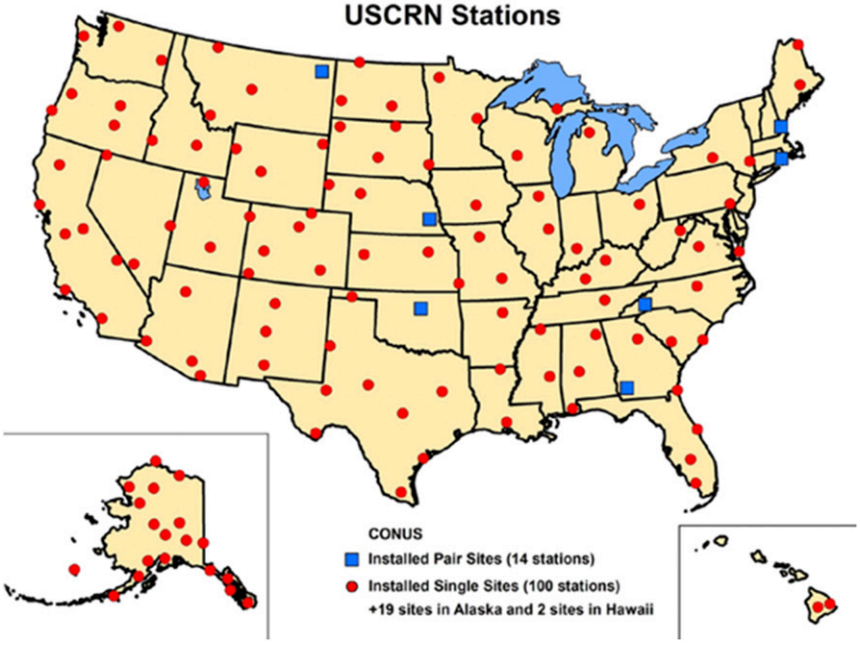

FIG. 1. Location of USCRN sites. CONUS sites are used for comparison of the radar-based products in this study.

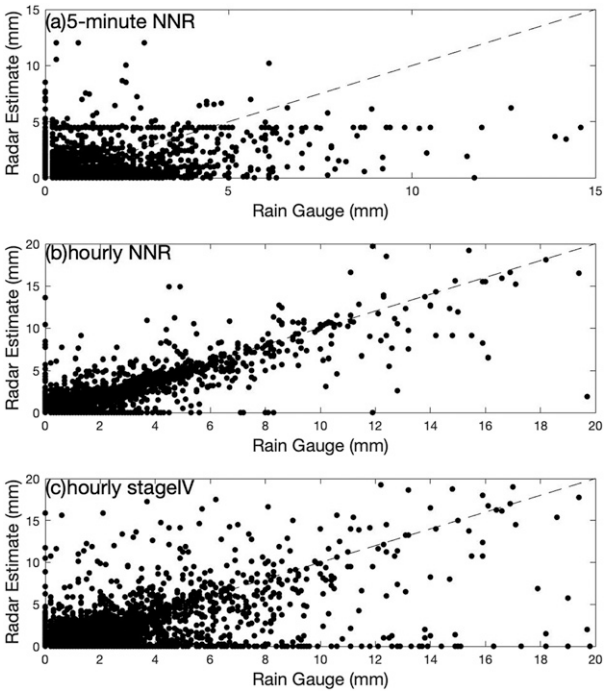

FIG. 2. (a) The USCRN 5-min observations vs the unadjusted NNR 5-min precipitation rate. (b) The USCRN hourly observation vs the gauge corrected NNR hourly accumulation. (c) The USCRN hourly observation vs the Stage IV hourly accumulation.

3D reflectivity structure and the environmental thermal and moisture fields.

5) RQIND

The radar quality indicator is a dimensionless variable that is a combined measure for beam blockage and vertical profile of reflectivity effects. Zhang et al. (2012a) describe the equations for determining the RQIND with a range from zero [full beam blockage and/or vertical profile of reflectivity (VPR) effects] to 1 (no beam blockage or VPR effects).

b. U.S. Climate Reference Network

The USCRN is a systematic and sustained network of climate monitoring stations with sites across the CONUS, Alaska, and Hawaii. The primary purpose of the USCRN network is to monitor air temperature, precipitation, and soil moisture/soil temperature. In addition to these parameters, each station measures ground surface (IR) temperature, solar radiation, wind speed, relative humidity, wetness from precipitation, and several values that monitor the operating condition of the equipment. Some of the secondary parameters contribute to improving the confidence in the observational measurements, and provide insight into the reliability and performance of the primary sensors (Diamond et al. 2013). Of the parameters measured by the USCRN network we use information on precipitation and ground surface temperature at 5-min and hourly intervals. The USCRN precipitation gauge is surrounded by a Small Double Fence Intercomparison Reference (SDFIR) shield and has a heating element in the gauge throat so this system measures snow better than standard Cooperative Observer Program (COOP) or Automated Surface Observing System (ASOS) gauges. Figure 1 shows the location of the USCRN stations used in this study. The USCRN is used for comparison in this study because the observations are not used in the gauge corrected radar observations. A note about the comparisons is that the USCRN 5-min precipitation requires an initial precipitation amount of 0.2 mm in any 5-min period, which can be problematic for comparisons with radar observations. For example, it may take multiple 5-min periods to reach a 0.2-mm increment recorded in a single 5-min period, while the stratiform rain is falling at a steady but lighter rate, making it seem like the radar is underestimating precipitation rate when it is not. Therefore, we restricted comparison to whenever the rain gauge estimate was at least 0.2 mm.

c. NCEP Stage IV radar-based estimates

The NCEP Stage IV product, herein referred to as Stage IV, is a near-real-time product that is generated at NCEP. It is based on the NEXRAD Precipitation Processing System (PPS) (Fulton 1998) and the NWS RFC precipitation processing. Originally the Stage IV product was intended for assimilation into atmospheric forecast models to improve quantitative precipitation forecasts (QPF) (Lin and Mitchell 2005). However, the product as it is currently generated and archived has become quite popular for various applications and we use it in this study as a comparison and sanity check albeit at a coarser spatial resolution. The overview of Stage IV is that Stage IV data are the mosaicked data from the 12 RFCs in the CONUS (Alaska is not included in Stage IV). And the RFC data are generated by the NWS PPS algorithm that uses mosaicked NEXRAD radar data and operational rain gauge data for bias adjustment. The rain gauge data used at each RFC varies depending on the RFC. However, some of the main rain gauge networks are the HADS, ASOS, and Remote Automatic Weather Stations (RAWS) and Automatic Local Evaluation in Real-Time (ALERT). An extensive assessment as well as a description of the processing system of the Stage IV dataset can be found in Nelson et al. (2016). While the Stage IV data are available since 2002, we only use the concurrent data to the rain gauge and NNR available data (2007–09).

d. Methodology

A comparison of radar rainfall estimates with rain gauges should provide a measure of their relative closeness and should take into consideration the subjectivity of comparing a gridded estimate with a point estimate (Kitchen and Blackall 1992). There are several statistical measures that when combined can tell the story of this closeness. However, including the information such as temperature, type of precipitation and an indication of radar quality, we can provide some additional information on the characteristics of the relative closeness of the estimates. While data are available for various temporal ranges we compared data from 2007 to 2009 since this was the period of overlapping data for all datasets.

1) FRACTIONAL STANDARD ERROR

The fractional standard error (FSE) provides a measure of the error in a relative sense. The FSE is a normalized

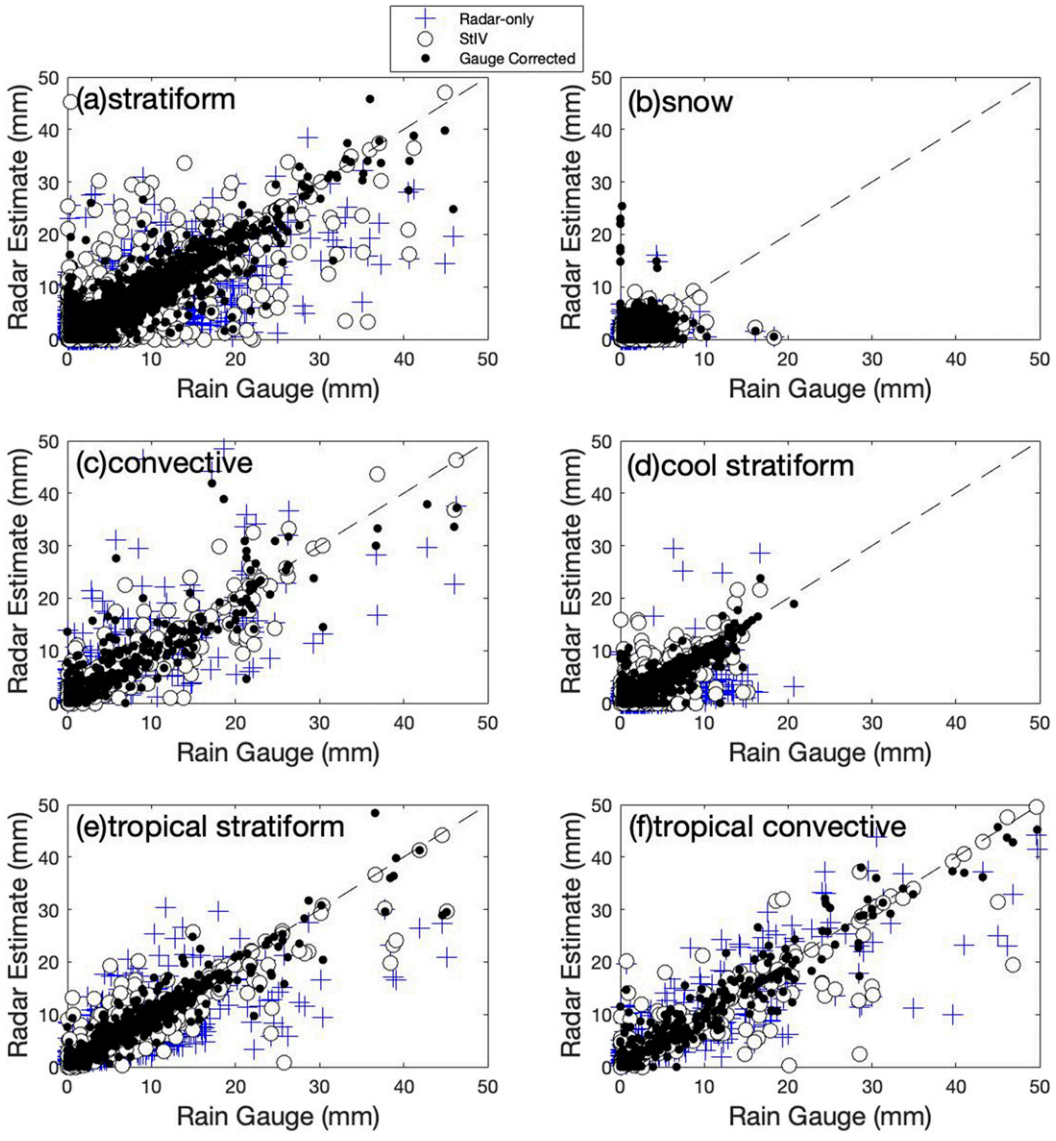

FIG. 3. Scatterplots of the hourly rain gauge estimate (x axis) vs the hourly radar based estimate (y axis) for the gauge-corrected radar, radar only, and Stage IV radar: (a) for precipitation flags identified as stratiform, (b) for precipitation flags identified as snow, (c) for precipitation flags identified as convective, (d) for precipitation flags identified as cool stratiform, (e) tropical stratiform, and (f) for precipitation flags identified as tropical convective.

root-mean-square error that is normalized by the average gauge value for the given condition. The normalization of the root-mean-square error allows for a comparison of the error across scales and conditions. The FSE provides a measure of the error for various conditions of rain rate defined from the rain gauge measurement G_n and the radar-based observation R_n:

$$\mathrm{FSE} = \frac{\sqrt{\frac{1}{N}\sum_{n=1}^{N}(G_n - R_n)^2}}{\frac{1}{N}\sum_{n=1}^{N}(G_n)}. \quad (1)$$

2) CORRELATION

Correlation is defined as a measure of interdependence of variable quantities. In most radar rainfall studies the sample Pearson correlation coefficient is used as the metric to illustrate this interdependence (Habib et al. 2001).

3) BIAS

Bias in terms of comparison of radar estimates with gauge estimates can be simply defined as the long-term ratio of the rain gauge measurement G_n as compared to the radar-based observation R_n. In this study we use a multiplicative bias defined as

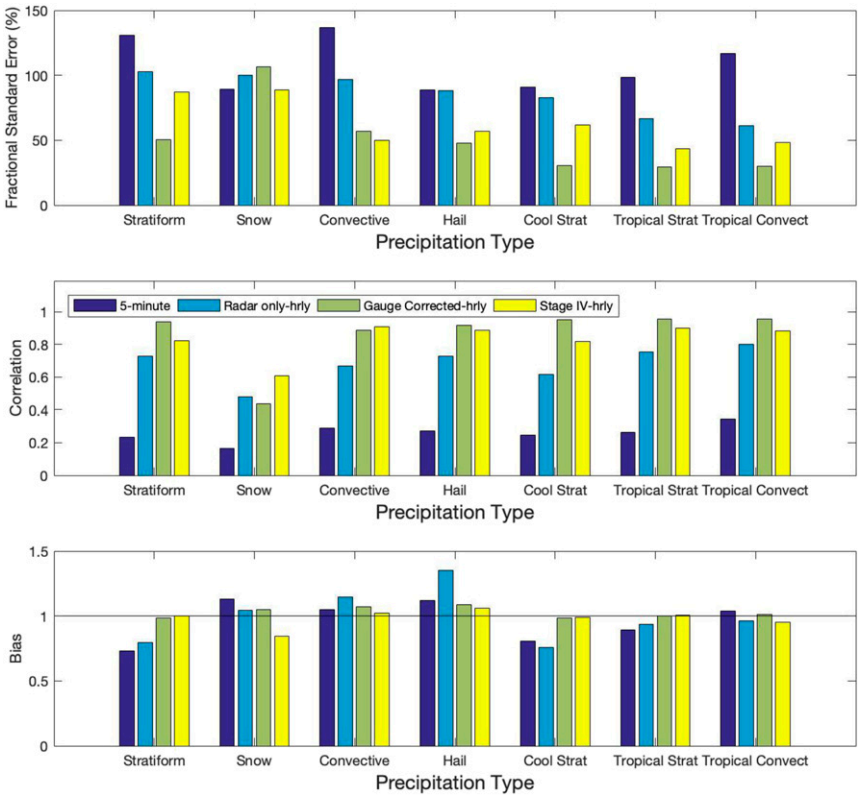

FIG. 4. (top) The fractional standard error, (middle) the correlation, and (bottom) the bias for each type of precipitation for the four different rain gauge–radar comparisons (5-min gauge–precipitation rate, hourly gauge–radar only, hourly gauge–gauge-corrected radar, and hourly gauge–Stage IV).

$$\text{bias} = \frac{\sum_{i=1}^{N} R_n}{\sum_{i=1}^{N} G_n}. \tag{2}$$

The rain gauge measurement G_n is assumed the reference, hence its summation is in the denominator (a bias value of 1.0 being unbiased).

3. Results

a. Radar gauge comparison

Radar rainfall observations from NOAA's NEXRAD Reanalysis are derived in the manner given by Zhang et al. (2016). The method lays out the process for deriving a precipitation estimate given the input information such as the modeled atmospheric temperature, the flagging process to identify precipitation type, and a Z–R relationship as defined in Zhang et al. (2016). The process of gauge correction via local gauge bias-correction is also defined in Zhang et al. (2016).

Initially we look at comparisons of the radar rainfall rain rates (PRATE) at the 5-min scale. In this study we convert the 5-min rain rate that is provided in millimeters per hour (m h^{-1}) to a 5-min accumulation in millimeters (mm) for comparison to the 5-min rain gauge estimate. Figure 2a shows this comparison at the 5-min scale. The 5-min radar observations are not corrected with in situ measurements and thus show large bias as compared to the rain gauge measurements. Figure 2b shows the comparisons of the hourly rain gauge measurements with the gauge corrected radar-based observations, and Fig. 2c shows the improvement in the gauge corrected radar-based observations as compared to the Stage IV radar-based estimates. What is noticeable is the improvement from the 5-min scale. Five-minute comparisons of rain gauge data with radar rainfall estimates are often problematic (Habib et al. 2001). In addition, at the hourly scale the improvement from the gauge corrected observations (GCQPE) as compared to the Stage IV is evident due the improved estimation algorithm such as better Z–R relations, identification and removal of anomalous propagation, and brightband identification.

b. Effect of precipitation type

The method for deriving precipitation observations provides for the identification of precipitation type given information on surface temperature, wet bulb temperature, expected hail, and the composite reflectivity (see Zhang et al. 2016). The resulting precipitation types are no precipitation, snow, hail, warm stratiform, cool stratiform, convective, tropical stratiform, and tropical convective. In this study we provide statistical analysis

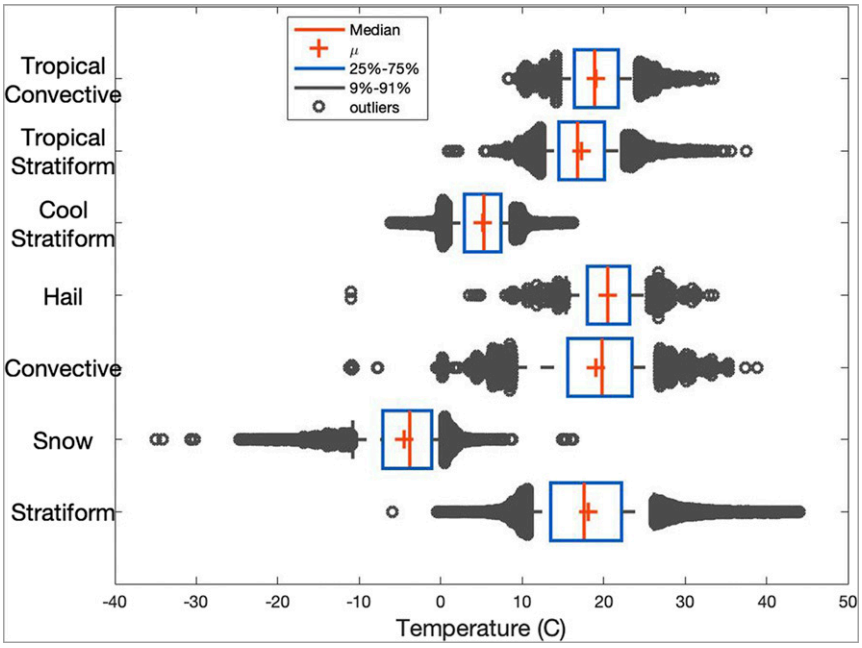

FIG. 5. Temperature extracted from the USCRN database vs the type of precipitation flag.

as described in section 2 for FSE, correlation, and bias for rain gauge versus radar-based observations (5-min, radar-only hourly, gauge corrected hourly, and NWS Stage IV hourly) for each type of precipitation as described above. Figure 3 shows the scatterplots for each of three radar-based observations (radar-only, gauge corrected, and Stage IV) categorized by precipitation flag: stratiform (Fig. 3a), snow (Fig. 3b), convective (Fig. 3c), cool stratiform (Fig. 3d), tropical stratiform (Fig. 3e), and tropical convective (Fig. 3f). In Fig. 3 the largest scatter is in the comparison of rain gauge and radar observations identified as snow and the least scatter is in the observations identified as tropical stratiform.

Figure 4a provides the comparisons for FSE for each type of precipitation. The largest FSE is in the convective and stratiform regime for the 5-min comparisons and then the least FSE is in the same precipitation regime for the NNR hourly gauge corrected precipitation. This provides some support for the gauge correction part of the NNR algorithm (Zhang et al. 2014). Overall, there is a low FSE for the gauge corrected precipitation comparisons except in the snow regime. In fact, in the snow regime the FSE increases for the gauge corrected comparisons (Cocks et al. 2016; Martinaitis et al. 2015; Zhang et al. 2012b). Errors in the snow regime can be due to timing of snow reports, stuck gauges, and the inherent variability in the reflectivity to snow conversion factor. And overall the FSE is lower in the gauge corrected precipitation estimates as compared to the Stage IV estimates with the exception of snow and a small increase in the convective regime. The smallest FSE are in the tropical stratiform and tropical convective regimes for the gauge corrected comparisons. These errors being smaller are likely due to the nature of the rainfall regime being more widespread spatially and temporally. Therefore, the radar(s) are better able to "see" the rainfall and have a better representativeness of the storm/system.

Figure 4b provides the comparisons for correlation for each type of precipitation. Of note is the low correlation in the 5-min comparisons as expected and the low correlations in all comparisons for the snow regime despite the high-quality frozen precipitation measurement by the USCRN stations. Also, the correlations are largest in the gauge corrected comparisons except in the snow regime and slightly lower in the convective regime. There are large improvements in the correlation from the radar-only estimates to the gauge corrected estimates (except snow), and the correlation in the gauge corrected comparisons are close to the Stage IV comparisons.

Figure 4c shows the bias results for each type of precipitation. A bias of 1.0 shows that on average there is relatively no over or underestimation. A bias over 1.0 suggests an overestimation of the radar-based observation and a bias under 1.0 suggests an underestimation of the radar-based observation. For both the gauge corrected comparisons and the Stage IV comparisons the bias is relatively low with the exception of the snow regime for Stage IV (under estimation) and convective for the gauge corrected (slight over estimation). The largest biases are in the 5-min observations with large under estimation in the stratiform labeled regimes and over estimation in the snow, convective and hail regimes. An exception is the large over estimation in the radar-only comparisons for hail. This is expected as a 5-min reflectivity identified as hail will then be converted to a large rainfall rate but then the gauge correction (Vasiloff et al. 2009; Zhang et al. 2014) procedure seems to do a good job of correction in the hail regime as the bias is greatly reduced.

c. *Effect of temperature*

The USCRN network provides air temperature at 1.5 m above the ground as a variable for each observation. We parsed

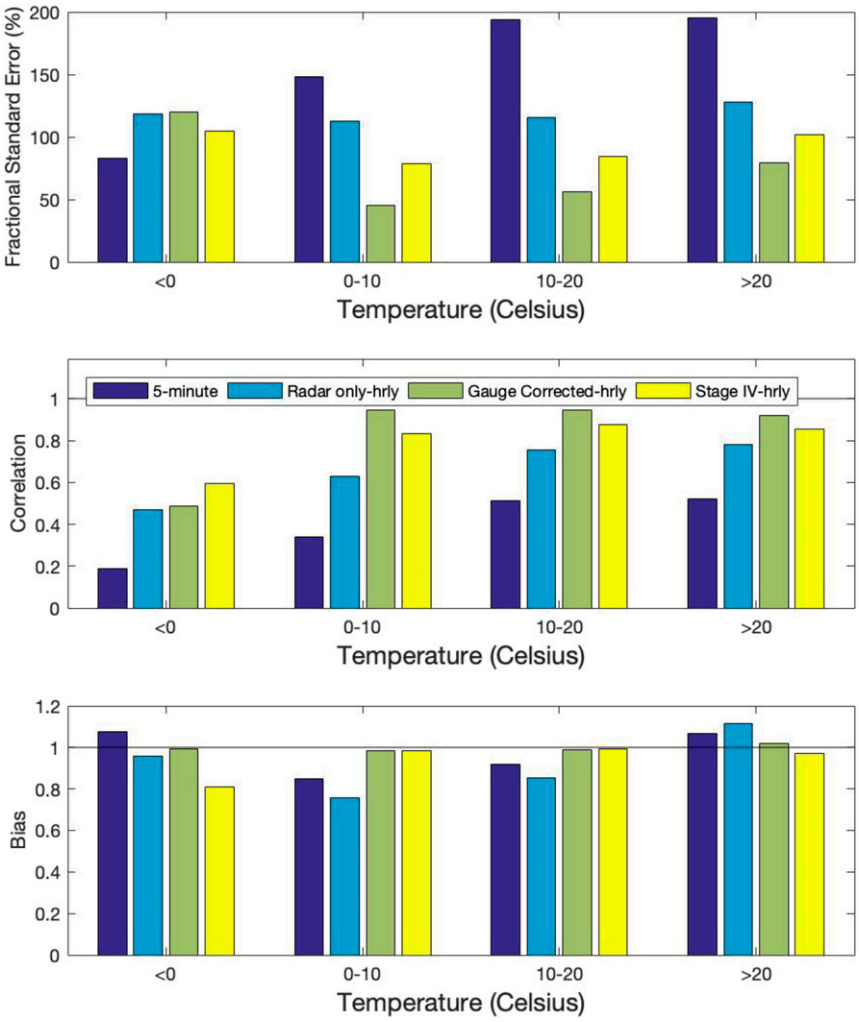

FIG. 6. (top) The fractional standard error, (middle) the correlation, and (bottom) the bias for a range of air temperature for the four different rain gauge–radar comparisons (5-min gauge–precipitation rate, hourly gauge–radar only, hourly gauge–gauge-corrected radar, and hourly gauge–Stage IV).

this temperature as a function of precipitation type. Figure 5 shows the ranges of temperature for each of the precipitation types. The range of temperatures identified in the USCRN observations are from −40°C to approximately 45°C. And as seen in the figure they generally agree with the precipitation typing scheme. However, the ranges of temperatures appear to be larger than expected in the stratiform precipitation, snow, and cool stratiform regimes. For example, the stratiform type range of temperatures is large from −10° to 40+°C. This could be due to an error in the precipitation flagging algorithm being influenced at times by hydrometeor conditions above the surface (i.e., at the radar scanning height) for these different regimes. For the most part the identification of precipitation type appears to match well with the related surface air temperature variable from the USCRN location. This led us to segment the statistics of FSE, correlation and bias based on the air temperature itself. Figure 6a provides the FSE based on a range of air temperature values. For the ranges of temperature greater than 0°C the 5-min comparisons are largest and the gauge corrected comparisons are smallest. But for freezing temperatures the 5-min comparisons have the smallest FSE, which compares well with the error analysis based on snow precipitation type.

Figure 6b provides the correlation based on the ranges of air temperature values. The smallest correlation for all comparisons is for freezing temperatures. Correlations increase considerably for the greater than zero ranges of air temperature and the gauge corrected comparisons have slightly larger correlation values versus the Stage IV comparisons. Likewise, in Fig. 6c the presentation of bias shows a similar story to correlations except that for temperatures greater than zero the 5-min comparisons show slightly better bias as compared to the radar-only comparisons. Again, the gauge corrected comparisons show low or almost no bias as well as the Stage IV except

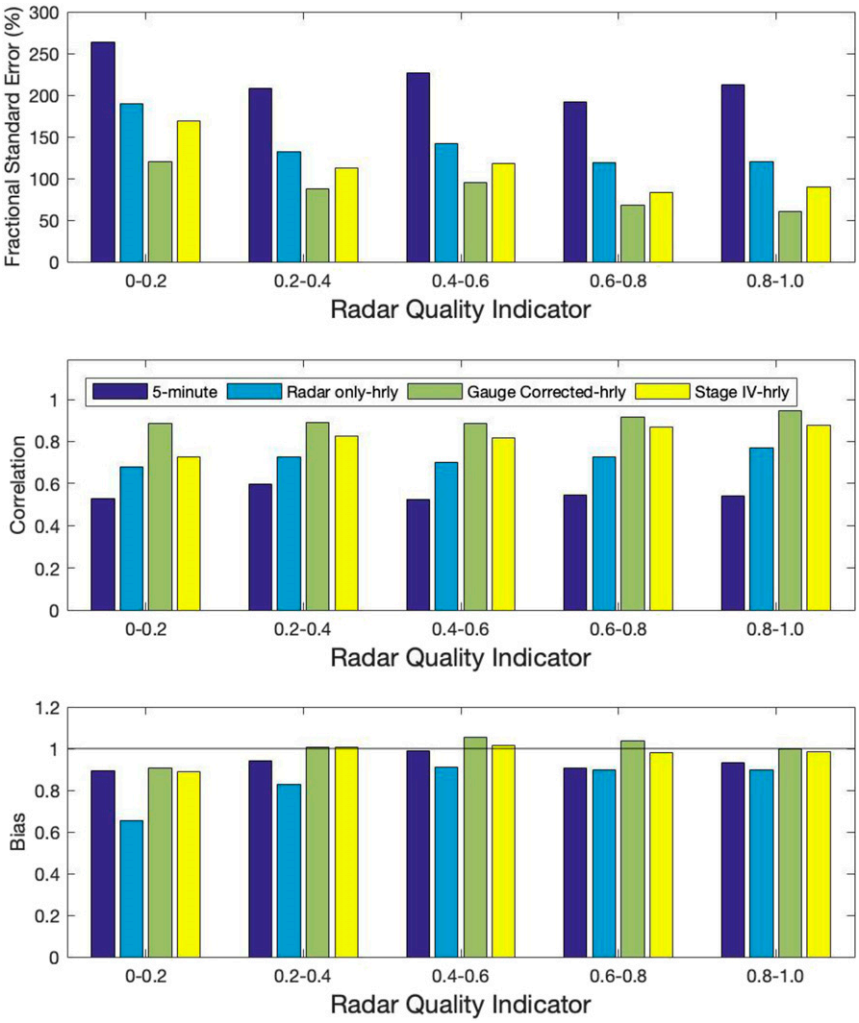

FIG. 7. (top) The fractional standard error, (middle) the correlation, and (bottom) the bias for radar quality indicator for the four different rain gauge–radar comparisons (5-min gauge–precipitation rate, hourly gauge–radar only, hourly gauge–gauge-corrected radar, and hourly gauge–Stage IV).

that for freezing temperatures the Stage IV shows a significant underestimation.

d. Effect of radar quality

Radar quality index (RQI) is defined in Zhang et al. (2012a) as a combined measure for beam blockage and VPR effects. VPR effects have to do with the variation of the reflectivity as it is returned from the atmosphere. There is variation in the vertical direction as well as with distance from that radar (Vignal et al. 1999). Identification of this variation aids in correcting the reflectivity that is used for precipitation estimation (Kirstetter et al. 2010). The RQI ranges from 0 (more blockage and VPR effects) to 1 (least blockage and VPR effects). In our study we segmented rain gauge measurements and radar-based observations by RQI bins (0–0.2, 0.2–0.4, 0.4–0.6, 0.6–0.8, and 0.8–1.0) and assessed the FSE, correlation, and bias (Fig. 7). For all statistical measures the best performance is at the 0.8–1.0 RQI bin. This is to be expected since the highest value of RQI relates to the least blocked areas with the least VPR effects. For FSE for each radar-based comparison the trend is for reduced FSE for increasing RQI and the gauge corrected comparisons show the smallest FSE with Stage IV having the next smallest FSE. The story is similar for correlation with increasing correlation for increasing RQI for all comparisons of the radar-based observations and the gauge corrected comparisons having the largest correlations. Finally, the comparisons become more unbiased with increasing RQI. However, for the gauge corrected comparisons there appears to be a slight over estimation at the 0.4–0.6 and 0.6–0.8 RQI bins. For RQI less than 0.6 the 5-min comparisons have less bias than the radar-only comparisons. Overall the statistics of the radar-based comparisons show improvement with increased radar quality. This is likely due to the advanced quality algorithms in the NNR that consider the effects of radar

artifacts (Lakshmanan et al. 2007a,b) as compared to the Stage IV observations.

4. Conclusions

In this study we have compared rain gauge estimates from the USCRN network for the period 2007–09 versus the NNR for both the 5-min and hourly time scales. We used information in the NNR dataset including the radar-only observation, a precipitation flag, and the radar quality index, and information from the USCRN such as surface air temperature to better discern the factors governing the relationship of radar estimates to gauge observations. In addition, we used the NCEP Stage IV product that is provided at a coarser spatial resolution as a separate comparison dataset and sanity check. A statistical analysis of the gauge and radar comparisons provides several important conclusions.

1) Five-minute comparisons of rain gauge estimates with radar observations are still problematic due to factors including subgrid variability, biased observations, and quality of observations. In this discussion bias is considered over or under estimation as referenced to an in situ observation. Quality of observation is related to the quality control measures applied to the observation or estimate.
2) The precipitation type can be an effective proxy for identification of the likely magnitude and sign of error and bias. Large errors are evident in the radar-only precipitation estimates for cool season type precipitation like stratiform and snow. These errors are reduced for radar-only precipitation estimates in the tropical type precipitation regimes. The gauge-corrected precipitation estimates (including the NCEP Stage IV) have reduced errors as compared to the radar-only precipitation estimates for all precipitation types except for snow, which can indicate the need for an improved snow identification algorithm and/or this can indicate the difficulty of using gauges for correction due to the nature of measuring snow with rain gauges (e.g., timing, melting, quantity). Correlations are largest for the gauge corrected precipitation estimates (including NCEP Stage IV) as compared to the radar-only estimates as expected due to the blending of the rain gauge observations. This is the case across all precipitation regimes except snow. Biases for the gauge corrected precipitation estimates are consistently close to 1.0 (unbiased), which are markedly improved from the radar-only estimates.
3) Temperature is an indicator for both precipitation type and error and bias. In the temperature ranges of less than 0°C large errors are evident in the radar-based observations, and in the type of precipitation indicated by radar. This problem has always been difficult since precipitation in the freezing regime provides a highly variable reflectivity estimate based on the returned electromagnetic power from the hydrometeor back to the radar.
4) Radar quality indicator is less of an indicator for error and bias except for the values in the 0.8–1.0 bin value, which show the best agreement of the radar-based observations with the rain gauge estimates. Correlation and FSE both improve for each increased bin of RQI. This is as expected since an increasing RQI suggests a better precipitation estimate due to better radar coverage either vertically or horizontally.
5) The bias correction procedure for blending radar and rain gauges in the NNR performs relatively well as compared to the NCEP Stage IV analysis.
6) The NNR is able to identify several types of precipitation and assign different Z–R relations as compared to Z–R relations (only four) assigned to the Stage IV observations. This ability to identify different precipitation types is a factor in reduced error and bias and improved correlation. However there is still room for improvement in cool season precipitation and snow/frozen precipitation regimes.

In conclusion, radar precipitation estimates have been shown to be sensitive to precipitation type, surface air temperature and the quality of the radar reflectivity estimates. Some uncertainties are no doubt due to the quality of gauge precipitation measurements, especially those at 5-min intervals. This paper provides information on the error, correlation and bias in the NNR radar-only and gauge-corrected precipitation estimates as well as the operational NCEP Stage IV precipitation estimates. These statistics can be valuable for studies of precipitation given supporting information on the meteorological conditions of the events. In addition, improvements in precipitation estimation algorithms could be focused to certain aspects of the rainfall/frozen type precipitation regimes.

REFERENCES

Chen, S., and Coauthors, 2013: Evaluation and uncertainty estimation of NOAA/NSSL next-generation national mosaic quantitative precipitation estimation product (Q2) over the continental United States. *J. Hydrometeor.*, **14**, 1308–1322, https://doi.org/10.1175/JHM-D-12-0150.1.

Ciach, G. J., and W. F. Krajewski, 1999: Radar-rain gauge comparisons under observational uncertainties. *J. Appl. Meteor.*, **38**, 1519–1525, https://doi.org/10.1175/1520-0450(1999)038<1519:RRGCUO>2.0.CO;2.

Cocks, S. B., S. M. 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.

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

Fulton, R. A., 1998: WSR-88D polar-to-HRAP mapping. NOAA Tech. Memo., 33 pp.

Grams, H. M., J. Zhang, and K. L. Elmore, 2014: Automated identification of enhanced rainfall rates using the near-storm environment for radar precipitation estimates. *J. Hydrometeor.*, **15**, 1238–1254, https://doi.org/10.1175/JHM-D-13-042.1.

Habib, E., W. F. Krajewski, and G. J. Ciach, 2001: Estimation of rainfall interstation correlation. *J. Hydrometeor.*, **2**, 621–629, https://doi.org/10.1175/1525-7541(2001)002<0621:EORIC>2.0.CO;2.

——, G. J. Ciach, and W. F. Krajewski, 2004: A method for filtering out raingauge representativeness errors from the verification

distributions of radar and raingauge rainfall. *Adv. Water Resour.*, **27**, 967–980, https://doi.org/10.1016/j.advwatres.2004.08.003.

Kirstetter, P.-E., G. Delrieu, B. Boudevillian, and C. Obled, 2010: Toward an error model for radar quantitative precipitation estimation in the Cévennes–Vivarais region, France. *J. Hydrol.*, **394**, 28–41, https://doi.org/10.1016/j.jhydrol.2010.01.009.

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

Kitchen, M., and R. M. Blackall, 1992: Representativeness errors in comparison between radar and gauge measurements of rainfall. *J. Hydrol.*, **134**, 13–33, https://doi.org/10.1016/0022-1694(92)90026-R.

Lakshmanan, V. T., A. Fritz, T. Smith, K. Hondl, and G. J. Stumpf, 2007a: An automated technique to quality control radar reflectivity data. *J. Appl. Meteor. Climatol.*, **46**, 288–305, https://doi.org/10.1175/JAM2460.1.

——, T. Smith, G. J. Stumpf, and K. Hondl, 2007b: The warning decision support system–integrated information. *Wea. Forecasting*, **22**, 596–612, https://doi.org/10.1175/WAF1009.1.

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, https://ams.confex.com/ams/Annual2005/techprogram/paper_83847.htm.

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.

Nelson, B. R., 2017: C-ATBD: Precipitation – NEXRAD QPE CDR. Doc. CDRP- ATBD-0917, NCDC Climate Data Record Program, 29 pp., https://www1.ncdc.noaa.gov/pub/data/sds/cdr/CDRs/Precipitation-NEXRAD_QPE/AlgorithmDescription%20_01B-22.pdf.

——, O. P. Prat, D.-J. Seo, and E. Habib, 2016: Assessment and implications of NCEP Stage IV quantitative precipitation estimates for product intercomparisons. *Wea. Forecasting*, **31**, 371–394, https://doi.org/10.1175/WAF-D-14-00112.1.

Qi, Y., J. Zhang, and P. Zhang, 2013: A real-time automated convective and stratiform precipitation segregation algorithm in native radar coordinates. *Quart. J. Roy. Meteor. Soc.*, **139**, 2233–2240, https://doi.org/10.1002/qj.2095.

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**, 2035–2045, https://doi.org/10.1029/96WR00270.

Vasiloff, S. V., K. W. Howard, and J. Zhang, 2009: Difficulties with correcting radar rainfall estimates based on rain gauge data: A case study of severe weather in Montana on 16–17 June 2007. *Wea. Forecasting*, **24**, 1334–1344, https://doi.org/10.1175/2009WAF2222154.1.

Vignal, B., H. Andrieu, and J. D. Creutin, 1999: Identification of vertical profiles of reflectivity from volume scan radar data. *J. Appl. Meteor. Climatol.*, **38**, 1214–1228, https://doi.org/10.1175/1520-0450(1999)038<1214:IOVPOR>2.0.CO;2.

Xu, X., K. Howard, and J. Zhang, 2008: An automated radar technique for the identification of tropical precipitation. *J. Hydrometeor.*, **9**, 885–902, https://doi.org/10.1175/2007JHM954.1.

Young, C. B., B. R. Nelson, A. A. Bradley, J. A. Smith, C. D. Peters-Lidard, A. Kruger, and M. L. Baeck, 1999: An evaluation of NEXRAD precipitation estimates in complex terrain. *J. Geophys. Res.*, **104**, 19 691–19 703, https://doi.org/10.1029/1999JD900123.

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

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

——, ——, D. Kingsmill, and K. Howard, 2012b: Radar-based quantitative precipitation estimation for the cool season in complex terrain: Case studies from the NOAA Hydrometeorology Testbed. *J. Hydrometeor.*, **13**, 1836–1854, https://doi.org/10.1175/JHM-D-11-0145.1.

——, ——, C. Langston, B. Kaney, and K. Howard, 2014: A real-time algorithm for merging radar QPEs with rain gauge observations and orographic precipitation climatology. *J. Hydrometeor.*, **15**, 1794–1809, https://doi.org/10.1175/JHM-D-13-0163.1.

——, 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.