Evaluation and Uncertainty of MRMS v12 Dual-polarized Radar Quantitative Precipitation Estimation Product

HENRY M. SANTER*

National Weather Center Research Experiences for Undergraduates Program
Norman, Oklahoma
University of Maryland, College Park
College Park, Maryland

HEATHER M. GRAMS

NOAA/OAR/National Severe Storms Laboratory
Norman, Oklahoma

ABSTRACT

Two radar-based quantitative precipitation estimation (QPE) products from the Multi-Radar/Multi-Sensor (MRMS) system are evaluated against hourly gauge-based rainfall accumulations. The two products are Q3RAD, the QPE deployed with the initial operating version of MRMS, and Q3EVAP, a new product that uses rain relationships based on specific attenuation A and applies an evaporation correction. The evaluation takes place over the entire conterminous United States (CONUS) over an 18-month period starting in August 2018. Regional analysis of both products suggests that Q3EVAP has a more linear relationship with rainfall than Q3RAD, and in fact is usually closer to the true hourly rainfall than Q3RAD. Q3EVAP mitigates a wet bias that Q3RAD tends to exhibit, but in turn is more prone to underestimation, particularly in areas of worse radar coverage. Both products tend toward overestimation under very warm conditions, but Q3EVAP limits the extent of this with the evaporation correction. For very high rain rates, both products underestimate significantly, but Q3EVAP consistently provides a similar or better estimate compared to Q3RAD. This is attributed primarily to higher linearity between specific attenuation and rainfall than with other radar variables. The results of this study should be useful to users wanting to understand Q3EVAP's limitations and to developers in pursuit of probability-based QPE.

1. Introduction

Reliable, high-resolution quantitative precipitation estimation (QPE) provides crucial information for real-time flash flood guidance, forecast model verification, hydrological modelling, agriculture, and infrastructure and resource management. Creating QPEs over the conterminous United States (CONUS) requires both a massive amount of data and the ability to process it into a meaningful product in a timely manner.

To address the need for data, a wide array of sensor systems for measuring rainfall exist, each with its own advantages and disadvantages. Rain gauges measure rainfall directly, and are often considered to be well representative of the true rainfall rates in their immediate surroundings, but are sparsely distributed. Remote sensing systems achieve a much better coverage, but measure by less direct means that can be subject to greater errors. Weather radar in par-

ticular is able to achieve great coverage and resolution in space and time, and the WSR-88D radar network has been a large contributor to NWS QPE generation since the late 1990s (Fulton et al. 1998). Between 2011 and 2013, the WSR-88D network was upgraded to dual-polarization capability, which enabled a hydrometeor classification (HC) algorithm (Park et al. 2009) and provided additional information that could be used to improve QPE.

Developments in computational methods and internet speed and reliability have made it increasingly possible to move radar data to a centralized location for processing and distribution (Kelleher et al. 2007). The Multi-Radar/Multi-Sensor (MRMS) system built at the National Severe Storms Lab (NSSL) is a realization of these developments. MRMS integrates data from the WSR-88D radar network, approximately 30 Canadian radars, satellite data, NWP model information, and rain gauge and lightning data to create a suite of different products. In doing so, it is able to leverage the strengths of the different inputs to create data sets that are high-resolution in space (1 km) and time (2 min), and are optimized to leverage the

^{*}Corresponding author address: Henry Santer, University of Maryland at College Park, 120 David L. Boren Blvd E-mail: henrysanter3@gmail.com

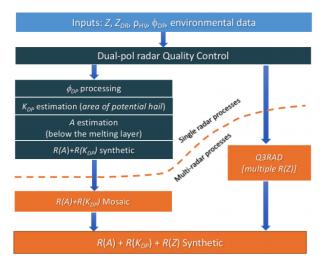


FIG. 1: Flowchart summarizing the creation of Q3RAD and the synthetic QPE that the evaporation correction is applied to to create Q3EVAP. Figure from Zhang et al. (2020).

highest quality data sources at any given location. (Zhang et al. 2016).

The deployment of MRMS allowed for a significant step toward accurate, high-resolution radar QPE in the form of a synthetic, HC-based product called Q3RAD (Zhang et al. 2016). Q3RAD chooses from a selection of rainfall relationships using Z (reflectivity) based on the type of precipitation occurring. However, R(Z) relationships are sensitive to calibration biases, and subject to errors from partial radar beam blockage (PBB). This has motivated investigation in recent years into rainfall relationships based on specific differential phase (K_{DP}) and specific attenuation (A) (Ryzhkov et al. 2014), which are immune to those measurement errors. A new synthetic radar product called Q3DP has been in development that applies R(A) relationships in areas of pure rain, $R(K_{DP})$ relationships in areas with potential hail and ice, and R(Z) relationships (i.e. reverts to Q3RAD) in areas above the melting layer (ML) with mixed-phase precipitation (Zhang et al. 2017). An evaporation correction has been added to Q3DP to reduce false precipitation echoes from virga. The resulting MRMS product is called Q3EVAP, and is slated to be ready for use operationally by the NWS by October 2020. Figure 1 provides an overview of how Q3RAD and Q3DP are created.

The transition to Q3EVAP represents a notable improvement in accuracy over current QPE products, but is still subject to errors. The focus of this study is to complement the work of Zhang et al. (2020) by quantifying the error structure of Q3EVAP with respect to rainfall intensity, geographic location, and the surrounding environmental conditions. The methodology used for this study

is inspired by the work of Chen et al. (2013), and adapts their analysis to compare Q3EVAP and Q3RAD against hourly rain gauge accumulations. The next section will describe the data sets and statistical methods used. The results of the analysis will be in section 3. Section 4 will have a summary and discussion of the results, and closing remarks.

2. Data and Methods

This study uses hourly rain gauge data, paired with Q3RAD and Q3EVAP accumulations at each gauge location. The data was gathered between August 2018 and February 2020, and each sample is also accompanied by location data, temperature, humidity information, and melting layer height. Figure 2 shows the distribution and concentrations of observations across the CONUS.

a. Radar Products

Q3EVAP calculates the rainfall rate R from a combination of specific attenuation A, specific differential phase K_{DP} , and reflectivity Z. The choice of rainfall relationship used is based on the likelihood that the radar is sampling frozen or mixed-phase precipitation, which is assessed by reflectivity and the melting layer (ML) height. In areas of rain without possible ice or hail (Z < 50 dBZ, below ML), an R(A) relationship is used. Past studies have shown that R(A) methods are less susceptible to radar calibration errors (Ryzhkov et al. 2014; Cocks et al. 2019) and partial blockage of the radar beam (Cocks et al. 2019). Additionally, R tends to have a higher linearity with R than with other standard radar variables (Ryzhkov et al. 2014). In areas where hail is possible (R > 50 dBZ, below ML),

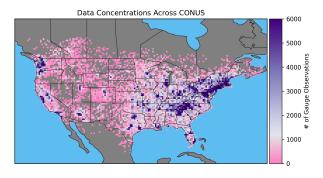


FIG. 2: Distribution of gauge observations across the CONUS. Each pixel is approximately 21 km long on the side.

an $R(K_{DP})$ relationship is used. $R(K_{DP})$ relationships share many of the strengths of R(A) relationships, but calculating K_{DP} inherently requires smoothing that results in a lower resolution than R(A) relationships can provide

(Ryzhkov et al. 2014). In and above the melting layer, Q3EVAP reverts to Q3RAD. To prevent discontinuities in Q3EVAP, the areas of different rainfall relationships are combined with a weighted averaging scheme (Qi and Zhang 2017). An evaporation correction is then applied to the resulting product as in Martinaitis et al. (2018).

Q3RAD was the radar-only QPE product initially introduced with the MRMS system. Q3RAD makes extensive use of the MRMS PCP_FLAG product, which classifies surface precipitation into one of seven categories, depending on the reflectivity, surrounding temperature, wet-bulb temperature, and ML height. Different precipitation types are then treated with different R(Z) relationships. For a detailed description of Q3RAD and the initial operating capabilities of MRMS QPE, see Zhang et al. (2016).

The MRMS Radar Quality Index (RQI) product

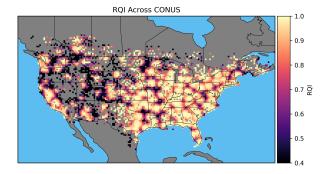


FIG. 3: Radar Quality Index (RQI) across the CONUS over the course of this study.

acts as an aggregate measure for uncertainty based on calibration errors, PBB, and variability in the vertical profile of reflectivity. RQI does not fully capture radar QPE uncertainty, but still correlates with some radar QPE errors (Chen et al. 2013). For each set of gauge and radar measurements in our data, we also have the corresponding RQI at the time of observation. Error sensitivity to RQI should serve as a good metric for sensitivity of QPE products to distance from the nearest radar, and potentially for sensitivity to PBB issues. Figure 3 shows average RQI values across the CONUS over the course of this study.

b. Gauge Network Data

The gauge data used for this study is sourced from the National Centers for Environmental Prediction Meteorological Assimilation Data Ingest System (MADIS) database. MADIS is a database of observations from NOAA and non-NOAA sources, processed and packaged for convenient use by the meteorology community. Before it is used by MRMS, MADIS gauge data is run through an

automated quality control scheme that takes surrounding temperature and QPE readings into account. Observations that read drastically over or under the surrounding QPE information are discarded, as are observations from gauges that are deemed partially or fully frozen (Zhang et al. 2016).

c. Comparison Methods

The purpose of this study is to understand the structure of the uncertainty in Q3EVAP measurements, and compare it to Q3RAD. We binned the data by gauge-based rainfall intensity, by location, and by various environmental parameters, and then modelled how the uncertainty in Q3EVAP and Q3RAD measurements varies across each set of bins. The primary metrics for uncertainty used were mean accumulation difference (MAD), root-mean-squared error (RMSE), and Pearson correlation coefficient (CC):

$$\begin{aligned} \text{MAD} &= \frac{\sum_{i=1}^{N} (Q_{i} - G_{i})}{N} \\ \text{RMSE} &= \sqrt{\frac{\sum_{i=1}^{N} (Q_{i} - G_{i})^{2}}{N}} \\ \text{CC} &= \frac{\sum_{i=1}^{N} (Q_{i} - \bar{Q})(G_{i} - \bar{G})}{\sqrt{\sum_{i=1}^{N} (Q_{i} - \bar{Q})^{2} \sum_{i=1}^{N} (G_{i} - \bar{G})^{2}}} \end{aligned}$$

where Q_i and G_i are the rainfall estimates from the radar QPE and gauge accumulation, respectively, for the *i*th observation, and N is the total number of observations.

3. Results

a. Regional Analysis

Spatial maps of MAD, RMSE, and CC were produced for each product at about 21 km resolution from the study data, again using the gauge-accumulated rainfall as the reference "truth" value. Figures 4a-b show MAD values across the CONUS for Q3EVAP and Q3RAD. In general, low RQI is a good indicator for underestimation Both products underestimate significantly in the West, particularly along the Cascade and Sierra Nevada mountain ranges, which is attributable to poor radar coverage near the surface due to the surrounding terrain. This phenomenon also presents over the Rocky Mountains, where RQI is poorest. Q3EVAP underestimates moderately over much more of California and the Pacific Northwest (PNW) than Q3RAD. There is an overestimation trend in the California Central Valley and the flatter regions of Washington, where the terrain is more even, and Q3EVAP reduces this bias. Q3RAD exhibits a significant overestimation bias over the Great Plains and parts of the midwest, due to a tendency to

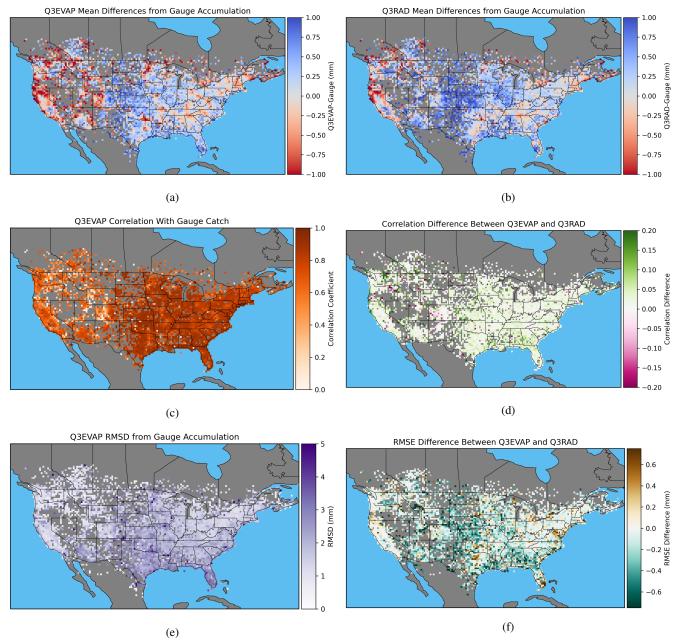


FIG. 4: MAD values for Q3EVAP (a) and Q3RAD (b), followed by Q3EVAP values and differences from Q3RAD values for CC (c-d) and RMSE (e-f).

FIG. 5

flag precipitation as tropical in these areas. Q3EVAP mitigates this bias. Underestimation bias and low RQI again correlate in the Appalachian Mountains and some remote areas in the Eastern United States. In these areas, the underestimation bias is more widespread and slightly stronger for Q3EVAP than for Q3RAD. Overall, Q3EVAP tends to measure lower QPEs across the country than Q3RAD, possibly due in part to the new evaporation

correction.

Figures 4c-f show CC values across the CONUS for Q3EVAP, the difference in CC from Q3RAD to Q3EVAP, and similar plots for RMSE values. Over most of the CONUS, Q3EVAP exhibits a CC increase of approximately 0.05. Exceptions to this seem closely related to RQI: almost all of the places where Q3EVAP has lower CC than Q3RAD are areas of RQI below 0.5. In the Great

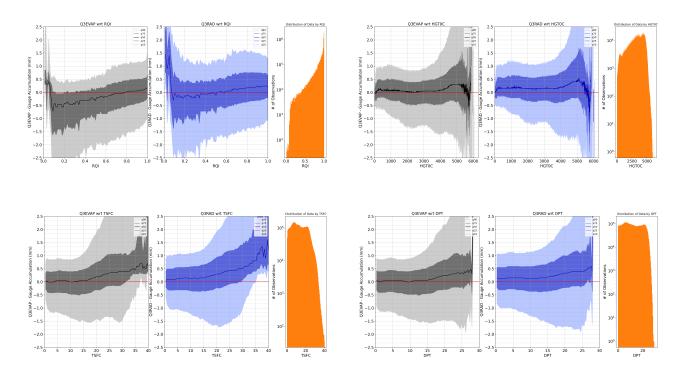


FIG. 6: Residual quantile plots for Q3EVAP/Q3RAD with respect to RQI, ML height (HGT0C), TSFC, and DPT. Orange plots show the number of observations that went into each bin for computing the quantiles.

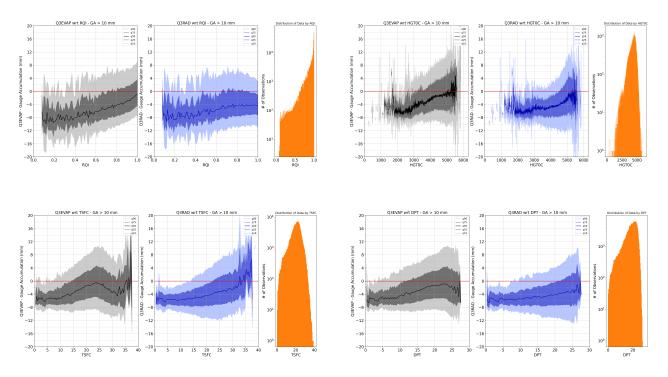


FIG. 7: As in figure 6, but only for rainfall amounts above 10 mm.

Plains and in the South, CC is often above 0.9, inhibited only by poor near-surface coverage in the Appalachian

Northeast and West, where RMSE for Q3RAD was often already under 2 mm, Q3EVAP's performance was generally similar. RMSE for Q3EVAP was higher across a number of metropolitan areas - more research is required to understand the causes of this.

b. Error Modelling of Q3EVAP/Q3RAD

Error quantiles were computed for each product as functions of RQI and various environmental conditions (Figure 6). Both QPEs tend to underestimate rainfall for low RQI, and overestimate it slightly for high RQI. The underestimation biases indicate both products are sensitive to beam height (Chen et al. 2013). Q3EVAP's underestimation bias is more pronounced than Q3RAD's; it is possible that this is the result of the evaporation correction overcorrecting based on information ingested at high beam heights. For all ranges of dew point (DPT) values in the study, O3RAD exhibits a systematic overestimation bias of 0.15 mm, which worsens starting at around 15° C. Q3EVAP eliminates this wet bias until 15° C, and then lessens it in magnitude. The latter

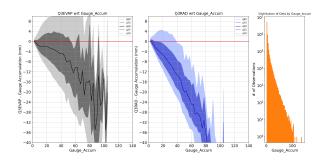


FIG. 8: Residual quantile plots for Q3EVAP and Q3RAD with respect to gauge-accumulated rainfall.

improvement is attributed to the higher linearity between A and rainfall rate. The distributions of residuals are similar when viewed as a function of surface temperature (TSFC), but Q3RAD overestimates more significantly when TSFC exceeds 30° C, and Q3EVAP mitigates that overestimation. With respect to the height of the melting layer bottom (HGT0C), both products show a slight overestimation bias that increases when HGT0C > 4000 m. This bias could be the result of falsely applying tropical rain relationships in nontropical environments, or could be indicative of sub-radar beam evaporation in environments with warm vertical profiles, where high melting layer heights are common. Still, Q3EVAP again exhibits a generally lower bias than Q3RAD.

The study data was skewed heavily toward lower rainfall rates - approximately half of all of the gauge observations recorded fell below 2 mm. To understand how the QPE products performed under more significant rainfall intensities, error models were also generated for hourly accumulations exceeding 10 mm of rainfall (Figure 7). For higher rain rates, Q3EVAP consistently underestimates more than Q3RAD does at low RQI, but at high RQI Q3EVAP errors center much closer to 0 than Q3RAD errors. This can be attributed to Q3EVAP reverting to Q3RAD under poor radar conditions (and reading consistently lower than Q3RAD because of the evaporation correction), and performing better than Q3RAD under favorable conditions when it is possible to use R(A) relationships. This phenomenon is also present looking at HGT0C; lower ML heights suggest a higher likelihood of intersecting the ML and reverting to Q3RAD, and only at higher (> 3000 m) ML heights does Q3EVAP demonstrate improvement over Q3RAD. As a function of DPT, Q3EVAP again shows error of similar or smaller magnitude, with the most significant decrease in magnitude occurring between 15 and 25 ° C. Residuals with respect to TSFC also improve for Q3EVAP in the mid ranges before dropping off again. Q3RAD still exhibits a tendency toward overestimation at the highest temperature range.

c. Rainfall Intensity Analysis

Figure 8 shows error models with respect to rain gauge accumulation for both products. For the very lowest range of accumulation values (under 2.5 mm), both products overestimate slightly: the mean residual for Q3EVAP in this range is 0.23 mm, and the mean residual for O3RAD is 0.38 mm. For larger gauge accumulation values, both products underestimate more significantly, on the order of tens of millimeters. The residuals for O3EVAP are frequently smaller in magnitude than those of Q3RAD, but are also more widely distributed. Additionally, the rate at which underestimation grows with respect to gauge accumulation is greater for Q3RAD than for Q3EVAP. It is important to note, however, that the volume of data contributing to the residuals for the higher accumulation ranges is orders of magnitude smaller than for the very smallest

To understand more precisely how the uncertainties for both products change with respect to gauge accumulation, density estimates were created for the distributions of QPE measurements under Q3EVAP and Q3RAD for gauge

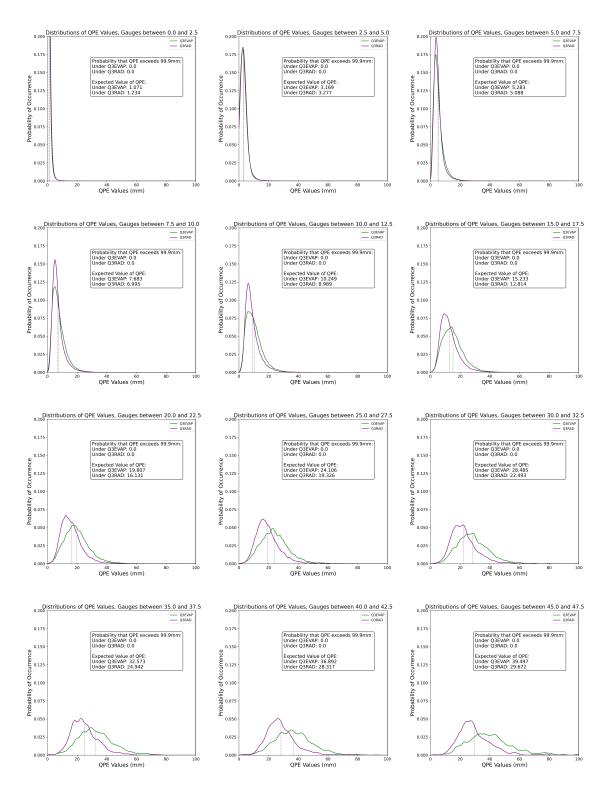


FIG. 9: Selected probability density estimates of Q3RAD/Q3EVAP measurements when gauge accumulations fall within 2.5 mm intervals, up to 50 mm.

accumulations in 2.5 mm intervals, up to 50 mm (past which there was not enough data to fit to meaningful looking distributions) (Figure 9). The distributions support the conclusion that both products measure similarly at very low rain rates, but at accumulation ranges as early as 5 - 7.5 mm the distributions for Q3EVAP and Q3RAD differ: Q3EVAP follows the gauge accumulations more closely in the expected value sense, but has more variance than Q3RAD. Both products also exhibit skewness toward lower measurements for lower accumulation values. Q3EVAP becomes closer to symmetric by about 25 mm, but Q3RAD never fully eliminates the skewness.

4. Summary and Conclusions

The Multi-Radar/Multi Sensor (MRMS) system generates quantitative precipitation estimation (QPE) products at high resolution in space (1 km) and time (2 min) over the CONUS. These products feed into flash flood forecasting, model verification, resource management, and other necessary services. Because of this, it is important to understand sources of uncertainty in QPE measurements, and to understand how uncertainty changes in moving from one QPE product to the next. The radar QPE product in use with the initial operating version of MRMS is called Q3RAD. The newest MRMS radar QPE product is called Q3EVAP, and is a synthetic $R(A) + R(K_{DP}) + R(Z)$ product that applies an evaporation correction to reduce false precipitation echoes from sub-radar beam evaporation. In this study, we evaluated Q3EVAP and Q3RAD against hourly gauge accumulations over an 18-month period over the CONUS. The evaluations were performed with respect to the location of each sample, the amount of rainfall, the surrounding environmental conditions at the time each sample was taken, and the MRMS Radar Quality Index (RQI) product, which quantifies the uncertainty of a radar measurement based on distance from the radar and partial radar beam blockage. findings from the evaluation are summarized below:

- Q3EVAP measured lower rain rates than Q3RAD across most of the CONUS. In particular, it mitigated a strong wet bias over the Great Plains, and led to more widespread dry bias over the Eastern and Western United States.
- Dry bias in both products tended to be related to relatively low RQI. We attributed this to the evaporation correction lowering rain rates further than necessary when precipitation was missed because of higher radar beam heights.
- Q3EVAP had an increased correlation coefficient with gauge-accumulated rainfall compared to Q3RAD, excepting only some regions in the western mountain ranges with lower RQI.

- Q3EVAP had a smaller RMSE than Q3RAD in the Central and Southern United States, and was similar in other regions. RMSE for Q3EVAP was higher over some major metropolitan areas.
- Q3RAD exhibited a systematic overestimation bias under most conditions that Q3EVAP largely eliminated. Both products tended toward overestimation under very warm conditions.
- When Q3EVAP was able to use R(A) relationships, it performed significantly better than Q3RAD in estimating higher rainfall rates.
- Both products still undermeasured high rainfall rates.
 Q3EVAP was closer to the true rainfall rate than
 Q3RAD, but was less precise.

The information provided by this study should be of use to product developers and users for understanding Q3EVAP's limitations and identifying pathways for future development. Further investigation into the sources and features of error in Q3EVAP should focus on performance differences east and west of the continental divide because of the significantly different environmental and radar quality conditions in each region. A similar analysis will be performed comparing Q3RAD against Q3DP to separate the influence of the evaporation correction from our understanding of how R(A) and $R(K_{DP})$ relationships compare with O3RAD. Future work should leverage the information this study provides in pursuit of a probabilistic QPE system. Ideally, a Q3EVAP measurement should lead to a probability density estimate of different true rainfall amounts, based on the surrounding conditions.

Acknowledgments. The authors are grateful for funding for this work provided by National Science Foundation Grant No. AGS-1560419. We thank Dr. Micheal Simpson for his work in putting together the data set, and Dr. Jian Zhang for her advice and guidance over the course of the study. Additionally, special thanks to Dr. Daphne LaDue and Ms. Alexandra Marmo for their support and leadership during the 2020 National Weather Center Research Experience for Undergraduates program.

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 (4), 1308–1322, doi:10.1175/JHM-D-12-0150.1, URL https://journals.ametsoc.org/jhm/article/14/4/1308/107600/Evaluation-and-Uncertainty-Estimation-of-NOAA-NSSL.

Cocks, S. B., and Coauthors, 2019: A Prototype Quantitative Precipitation Estimation Algorithm for Operational S-Band Polarimetric Radar Utilizing Specific Attenuation and Specific Differential Phase. Part II: Performance Verification and Case Study Analysis. J.

- Hydrometeor., 20 (5), 999–1014, doi:10.1175/JHM-D-18-0070.1, URL https://journals.ametsoc.org/jhm/article/20/5/999/344205/ A-Prototype-Quantitative-Precipitation-Estimation, publisher: American Meteorological Society.
- Fulton, R. A., J. P. Breidenbach, D.-J. Seo, D. A. Miller, and T. O'Bannon, 1998: The WSR-88D Rainfall Algorithm. Wea. Forecasting, 13, 19.
- Kelleher, K. E., and Coauthors, 2007: Project CRAFT: A Real-Time Delivery System for NEXRAD Level II Data Via the Internet. Bulletin of the American Meteorological Society, 88 (7), 1045–1058, doi:10.1175/BAMS-88-7-1045, URL https://journals.ametsoc.org/bams/article/88/7/1045/59138/Project-CRAFT-A-Real-Time-Delivery-System-for, publisher: American Meteorological Society.
- Martinaitis, S. M., H. M. Grams, C. Langston, J. Zhang, and K. Howard, 2018: A Real-Time Evaporation Correction Scheme for Radar-Derived Mosaicked Precipitation Estimations. *J. Hydrometeor.*, 19 (1), 87–111, doi:10.1175/JHM-D-17-0093.1, URL https://journals.ametsoc.org/jhm/article/19/1/87/69545/A-Real-Time-Evaporation-Correction-Scheme-for, publisher: American Meteorological Society.
- 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. Weather and Forecasting, 24 (3), 730–748, doi:10.1175/2008WAF2222205.1,
 URL https://journals.ametsoc.org/waf/article/24/3/730/39114/
 The-Hydrometeor-Classification-Algorithm-for-the, publisher:
 American Meteorological Society.
- Qi, Y., and J. Zhang, 2017: A Physically Based Two-Dimensional Seamless Reflectivity Mosaic for Radar QPE in the MRMS System. *J. Hydrometeor.*, 18 (5), 1327–1340, doi:10.1175/JHM-D-16-0197.
 1, URL https://journals.ametsoc.org/jhm/article/18/5/1327/69730/A-Physically-Based-Two-Dimensional-Seamless, publisher: American Meteorological Society.
- Ryzhkov, A., M. Diederich, P. Zhang, and C. Simmer, 2014: Potential Utilization of Specific Attenuation for Rainfall Estimation, Mitigation of Partial Beam Blockage, and Radar Networking. *Journal of Atmospheric and Oceanic Technology*, 31 (3), 599–619, doi:10.1175/JTECH-D-13-00038.1, URL https://journals.ametsoc.org/jtech/article/31/3/599/39281/ Potential-Utilization-of-Specific-Attenuation-for, publisher: American Meteorological Society.
- Zhang, J., L. Tang, S. Cocks, P. Zhang, A. Ryzhkov, K. Howard, C. Langston, and B. Kaney, 2020: A Dualpolarization Radar Synthetic QPE for Operations. *J. Hydrometeor.*, doi:10.1175/JHM-D-19-0194.1, URL https: //journals.ametsoc.org/jhm/article/doi/10.1175/JHM-D-19-0194.1/ 345338/A-Dual-polarization-Radar-Synthetic-QPE-for.
- Zhang, J., and Coauthors, 2016: Multi-Radar Multi-Sensor (MRMS) Quantitative Precipitation Estimation: Initial Operating Capabilities. Bulletin of the American Meteorological Society, 97 (4), 621–638, doi:10.1175/BAMS-D-14-00174.
 1, URL https://journals.ametsoc.org/bams/article/97/4/621/216222/Multi-Radar-Multi-Sensor-MRMS-Quantitative, publisher: American Meteorological Society.
- Zhang, J., and Coauthors, 2017: MRMS dual-polarization radar synthetic QPE. 38th Conference on Radar Meteorology, 8, URL https://ams.confex.com/ams/38RADAR/meetingapp.cgi/Paper/320875.