Using Citizen Science Reports to Evaluate Estimates of Surface Precipitation Type

BY SHENG CHEN, JONATHAN J. GOURLEY, YANG HONG, QING CAO, NICHOLAS CARR, PIERRE-EMMANUEL KIRSTETTER, JIAN ZHANG, AND ZAC FLAMIG

The Multi-Radar Multi-Sensor (MRMS) system uses data from the Next Generation Radar network (NEXRAD) combined with model analyses from the Rapid Refresh (RAP) system to provide precipitation rate and type products on a grid with 1-km horizontal resolution every 2 min (Zhang et al. 2011). The model analysis fields were found to be useful for estimating precipitation type *at the surface*, which differs from the various Hydrometeor Classification Algorithms (HCA) designed to identify hydrometeors at the height of radar sampling (e.g., Park et al. 2008). These sampled hydrometeors undergo changes in terms of their sizes, shapes, orientation, and phase as they fall and reach the surface. During the cool-season months, these changes often occur below the height of radar sampling. As a result, the MRMS algorithm uses model surface analyses to aid in the decision tree logic. To date, there has not yet been a systematic evaluation of the MRMS surface precipitation type products beyond case studies. An opportunity exists to employ newly collected citizen-scientist (also referred to as "crowdsourced") reports made possible through the meteorological Phenomena Indication Near the Ground (mPING) project to accomplish algorithm evaluations (Elmore et al. 2014).

Human weather observers have the ability to use multiple clues in order to make a decision about a given precipitation type. They typically begin with eyesight in order to assess the fall speed of the hydrometeor, its color, size, and shape. If uncertainty remains, then they may use other observations such as touch to determine its phase (liquid, frozen, or beginning to melt). Lastly, they can introduce ancillary observations such as local temperature or even perform additional experiments such as examining the particle with a magnifying glass to finalize their decision. Instruments do not have these adaptive capabilities and measure the variables according to their design. Discriminating precipitation types has its challenges and often requires multiple, expert observations. However, most people can readily distinguish more distinct precipitation types, such as rain from snow.

In this study, we provide the first systematic evaluation of rain and snow precipitation types estimated from the MRMS algorithm. This is of particular interest to satellite algorithm developers now that the recently launched Global Precipitation Mission (GPM; Hou et al. 2014) is measuring precipitation from space up to latitudes of 65°, and the MRMS product suite (including surface precipitation type) is the primary database for space-based algorithmic evaluations, development, and improvements (Kirstetter et al. 2012). Moreover, MRMS has recently been transitioned to operations at the National Centers for Environmental Prediction in the National Weather Service (NWS). We analyze the spatial distribution,

AFFILIATIONS: CHEN AND HONG—Advanced Radar Research Center, National Weather Center, and School of Civil Engineering and Environmental Science, University of Oklahoma, Norman, Oklahoma; GOURLEY—NOAA/National Severe Storms Laboratory, and School of Meteorology, University of Oklahoma, Norman, Oklahoma; CAO—Research and Innovation, Enterprise Electronics Corporation, Enterprise, Alabama; CARR— Advanced Radar Research Center, National Weather Center, and School of Meteorology, University of Oklahoma, Norman, Oklahoma; KIRSTETTER—Advanced Radar Research Center, National Weather Center, and NOAA/National Severe Storms Laboratory, Norman, Oklahoma; ZHANG—NOAA/National Severe Storms Laboratory, Norman, Oklahoma; FLAMIG—Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, Advanced Radar Research Center, National Weather Center, and NOAA/National Severe Storms Laboratory, Norman, Oklahoma

CORRESPONDING AUTHOR: Jonathan J. Gourley, National Weather Center, 120 David L. Boren Blvd., Norman, OK 73072

E-mail: jj.gourley@noaa.gov

DOI:10.1175/BAMS-D-13-00247.1

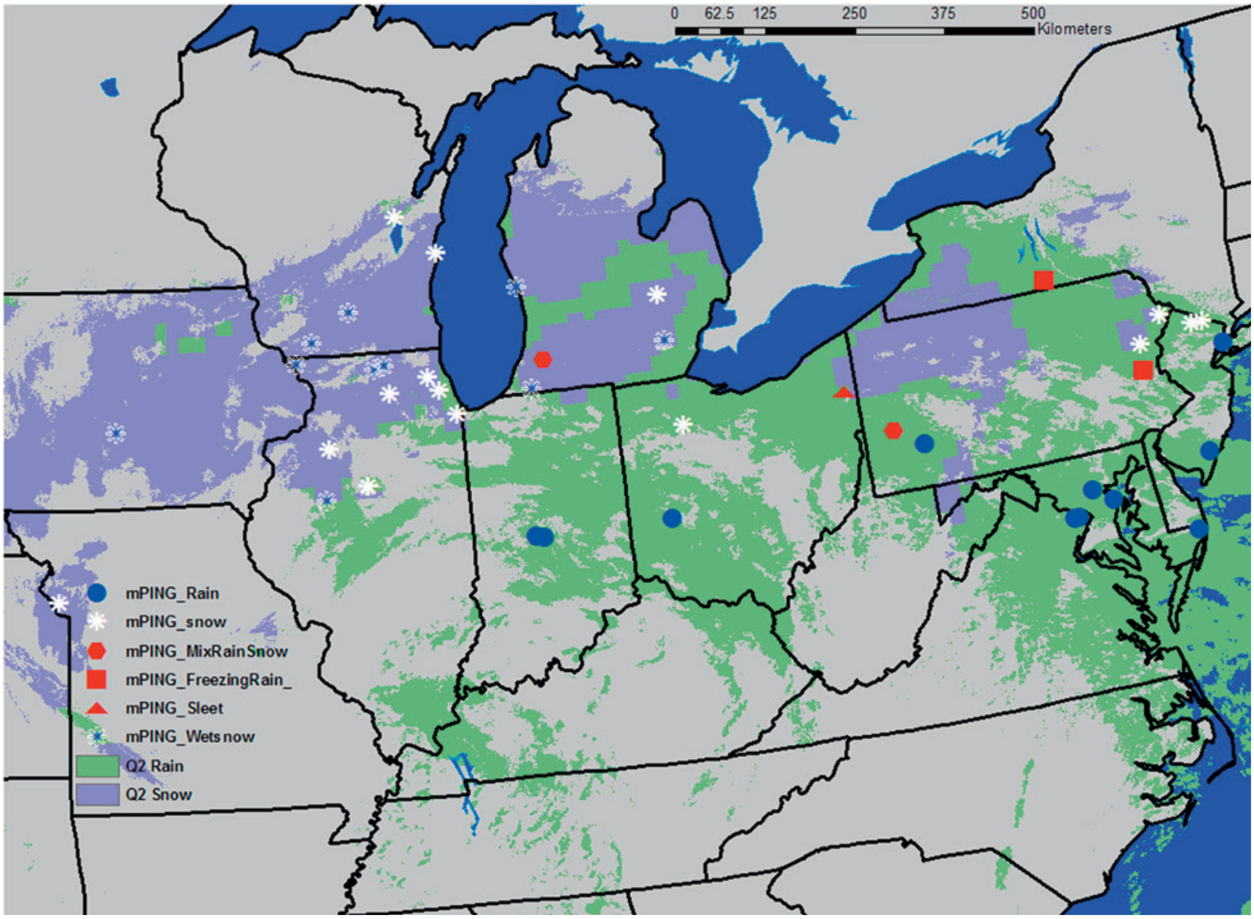

Fig. 1. Citizen-scientist reports of observed weather using the meteorological Phenomenon Indication Near the Ground (mPING) smartphone app overlaid on surface precipitation type product from the Multi-Radar Multi-Sensor (MRMS) system on 2340 UTC 26 Feb 2013. Refer to the legend for the different precipitation type categories.

density, and temporal variation of mPING reports and then compare them to MRMS precipitation type products during the cool season from 19 December 2012 through 31 April 2013.

MPING CITIZEN SCIENTIST REPORTS. In this study, we make use of the precipitation type reports from the mPING project. In short, members of the public can download an app free of charge on their global positioning system (GPS)-enabled smartphones and report the weather they are observing at their location. The reports include the time and geolocation of standard weather types including rain, snow, and several mixed-precipitation categories; hail sizes; severity of wind damage; and flash flooding severity. The reports automatically populate a database and can be viewed in real time on a website (www.nssl.noaa.gov/projects/ping/).

A total of 18,987 reports of snow and 23,356 reports of rain from 19 December 2012 through 30 April 2013 were recorded and are utilized in this study. We have chosen to only use those mPING reports of pure rain and pure snow in order to reduce observer uncertainty and bias in the myriad mixed types that can be reported. The interested reader should refer to Elmore et al. (2014) for an evaluation of mPING transitional precipitation types. The rain and snow categories are assumed to be associated with the least amount of uncertainty or at least can be considered unbiased. Moreover, the MRMS algorithm provides a basic rain–snow segregation product—thus the additional mPING precipitation types are superfluous in this study. Beyond removing the mixed-phase precipitation reports, no postprocessing procedures have been applied to assess the quality of the citizen-scientist reports.

MRMS PRECIPITATION TYPE PRODUCTS. The MRMS precipitation type algorithm uses decision tree logic to segregate rain from snow at the surface. The first decision removes echoes that are deemed to be too light to be associated with surface precipitation. If the RAP-analyzed surface temperature is greater than or equal to 5°C and the associated reflectivity value at the lowest, unblocked elevation angle is less than 5 dBZ, then the echoes are not considered for precipitation rate or type estimation. If the surface temperature is less than 5°C, then the threshold is dropped down to 0 dBZ, in recognition that snow has a lower dielectric constant than liquid water. Following removal of weak, nonprecipitating echoes, two thresholds are used to segregate rain and snow. The MRMS algorithm uses dry- and wet-bulb temperatures from the latest, hourly RAP analysis product. If the surface wet-bulb temperature is less than 0°C and the surface dry-bulb temperature is less than 2°C, then the surface precipitation type is set to frozen; otherwise it is set to rain. The use of two thresholds accounts for situations in which there are surface temperatures just above freezing, but there is wet snow reaching the surface.

INTERCOMPARISON OF MPING AND MRMS SURFACE PRECIPITATION TYPE REPORTS. All mPING rain and snow reports are matched to the nearest grid cell of the MRMS surface precipitation type product with 1-km horizontal resolution, produced every 2 min. Figure 1 shows an example of several mPING surface precipitation type reports overlaid on the MRMS precipitation type product at 2340 UTC on 26 February 2013. The red symbols correspond to sleet, mixed rain/snow, and freezing rain. It is noteworthy that all of these symbols for transitional precipitation types are located within 50 km of the MRMS rain–snow delineation line. This is a case where the horizontal temperature gradients are weak, which causes the threshold-based MRMS algorithm to "toggle" back and forth between rain and snow decisions versus depicting a distinct rain–snow line. The blue-filled circles indicating mPING rain reports are all located in the MRMS rain regions, but are all well displaced to the south and east of the rain–snow boundary line. Similarly, most of the mPING wet snow and snow reports are collocated with MRMS surface snow precipitation types. Mismatches are noted in northern Ohio and in northern New Jersey. In both regions with misclassifications, the MRMS rain–snow line is within a horizontal distance of 50 km. The temperatures were very near the thresholds in these regions, which indicates large uncertainty in the estimated precipitation types. Some suggestions on future probabilistic approaches to surface precipitation typing are provided later.

Maps of mPING rain and snow reports from 19 December 2012 to 31 April 2013 are shown in Figs. 2a and 2b, respectively. The rain reports in Fig. 2a represent a superposition of cool-season rainfall, population density, and perhaps cognizance of the mPING app itself. For instance, we see "hot spots" in Oklahoma City, Oklahoma; Tulsa, Oklahoma; Chicago, Illinois; Seattle, Washington; and Washington, D.C. The snow reports are much less common in the southern United States as expected, and we see relative maxima in Denver, Colorado; Kansas City, Missouri; Minneapolis, Minnesota; Chicago, Illinois; Cleveland,

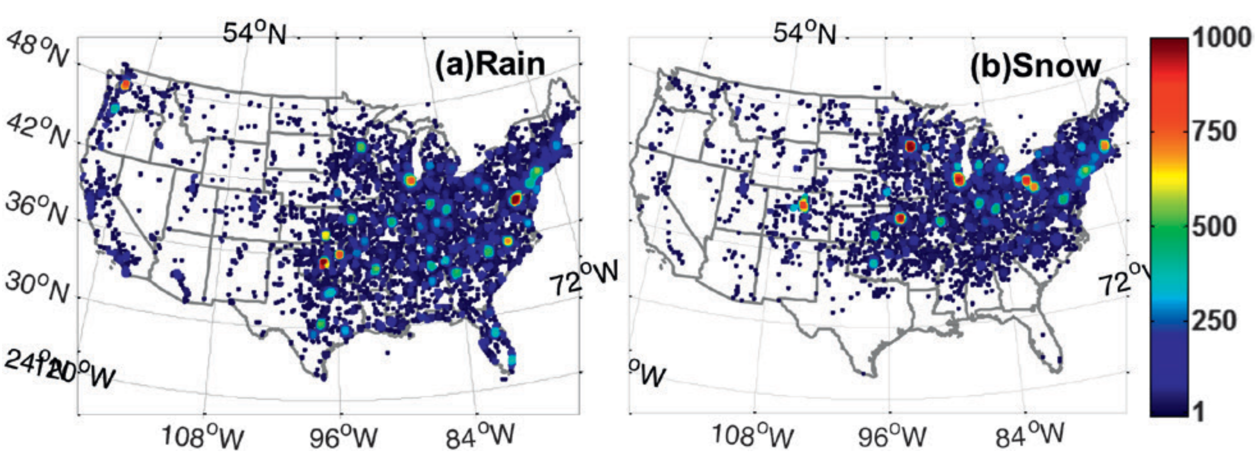

Fig. 2. Total count of mPING reports of (a) rain and (b) snow from 19 Dec 2012 to 31 Apr 2013. "Hot spots" referred to in text are indicated by red dots.

Ohio; and Boston, Massachusetts. There are far fewer reports of rain and snow in the sparsely populated, vast Intermountain West region between the Rocky Mountains and the Sierra Nevada range. Figure 3 shows a time series of daily rain and snow reports for the study period. There are no discernible trends indicating increasing reporting using the app. The reporting frequency seems to be more related to the episodic nature of synoptic-scale precipitation systems yielding mixed-phase events during the cool season. In April, the snow reports become much less common, while the rain reports continue as before.

Table 1 provides the population densities for several of these cities as well as the sample size of mPING reports and the height of the lowest, unblocked radar beam above the ground. Overall, we see excellent low-level radar coverage over the cities, which was one of the principal factors in siting the NEXRAD radars. So we would not expect there to be MRMS precipitation type biases over cities due to varying radar

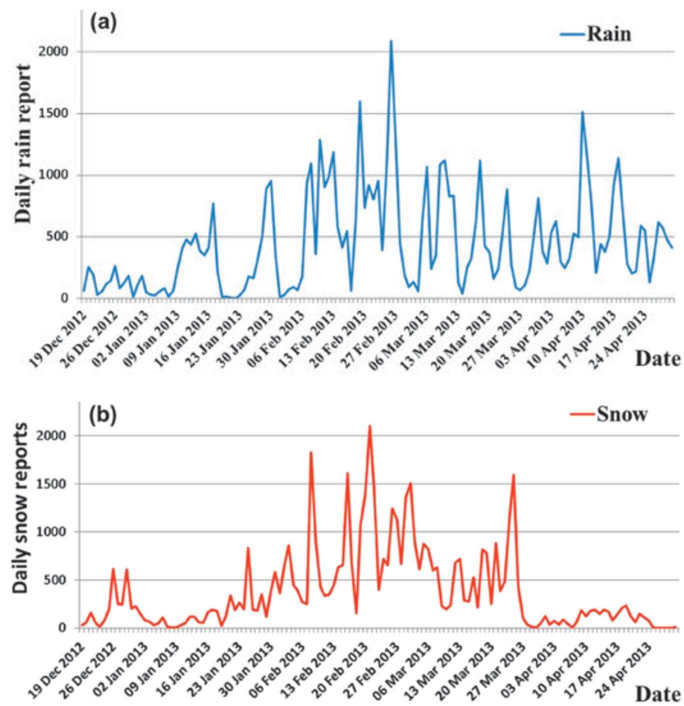

Fig. 3. Time series of mPING daily reports of (a) rain and (b) snow from 19 Dec 2012 to 31 Apr 2013.

Table 1. Probability of detection (POD, %) and sample sizes for rain and snow products for selected cities with high-density reports over the United States. The radar beam heights are above ground level (AGL). Cities are Boston (BOS); New York City (NYC); Washington, D.C. (DCA); Chicago (CHI); Denver (DEN); Oklahoma City (OKC); Minneapolis-St. Paul (MSP); and Seattle (SEA). The population densities are based on www.factmonster.com/ipka/A0763098.html.

City	Type	POD (%)	MRMS-Rain samples	MRMS-Snow samples	Total reports	Radar beam height (km)	Population
BOS	PING- Snow	73.2	132	361	848	0.37	636,479
	PING-Rain	94.4	335	20			
NYC	PING- Snow	72.2	117	304	827	0.62	8,336,697
	PING-Rain	93.4	379	27			
DCA	PING- Snow	72.4	111	291	921	0.37	632,323
	PING-Rain	96.3	500	19			
CHI	PING- Snow	72.1	105	272	954	0.19	2,714,856
	PING-Rain	93.6	540	37			
DEN	PING- Snow	73.7	84	339	1047	0.01	634,265
	PING-Rain	98.2	581	43			
OKC	PING- Snow	81.6	63	280	1109	0.01	599,199
	PING-Rain	93.6	717	49			
MSP	PING- Snow	80.1	84	339	1047	0.12	392,880
	PING-Rain	93.1	581	43			
SEA	PING- Snow	69.0	27	60	197	0.34	634,535
	PING-Rain	98.2	108	2			

coverage at low levels, but this effect may become more evident in more rural areas at long distance from radar. The largest number of total seasonal reports (1,109) occurred in the greater Oklahoma City metropolitan area. This is no surprise given the origination of the app in Norman, just outside of Oklahoma City. Comparison of the MRMS surface precipitation types to mPING reports is limited to the flatter terrain in the populated eastern two-thirds of the United States as well as cities west of the Sierra Nevada range in California, Oregon, and Washington.

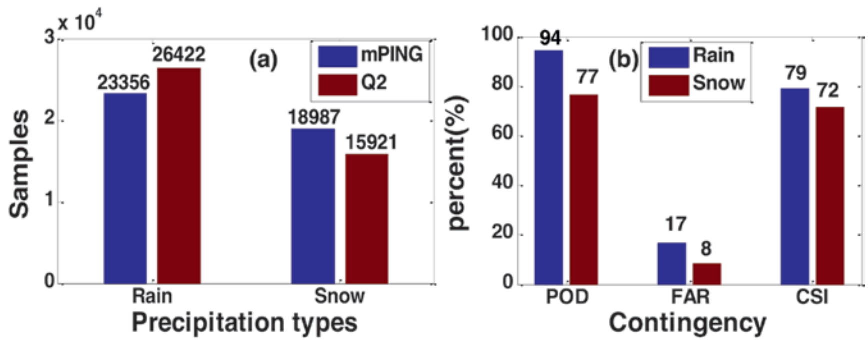

Fig. 4. (a) Sample size of mPING reports of rain and snow (blue bars) and matched surface precipitation types from the MRMS algorithm (red bars). (b) Contingency table statistics (expressed in percent) for the MRMS rain and snow products across the conterminous United States using the mPING reports as the reference. POD is probability of detection, FAR is false alarm rate, and CSI is critical success index.

Figure 4 shows the sample sizes for mPING reports of rain and snow for the > 4-month period of

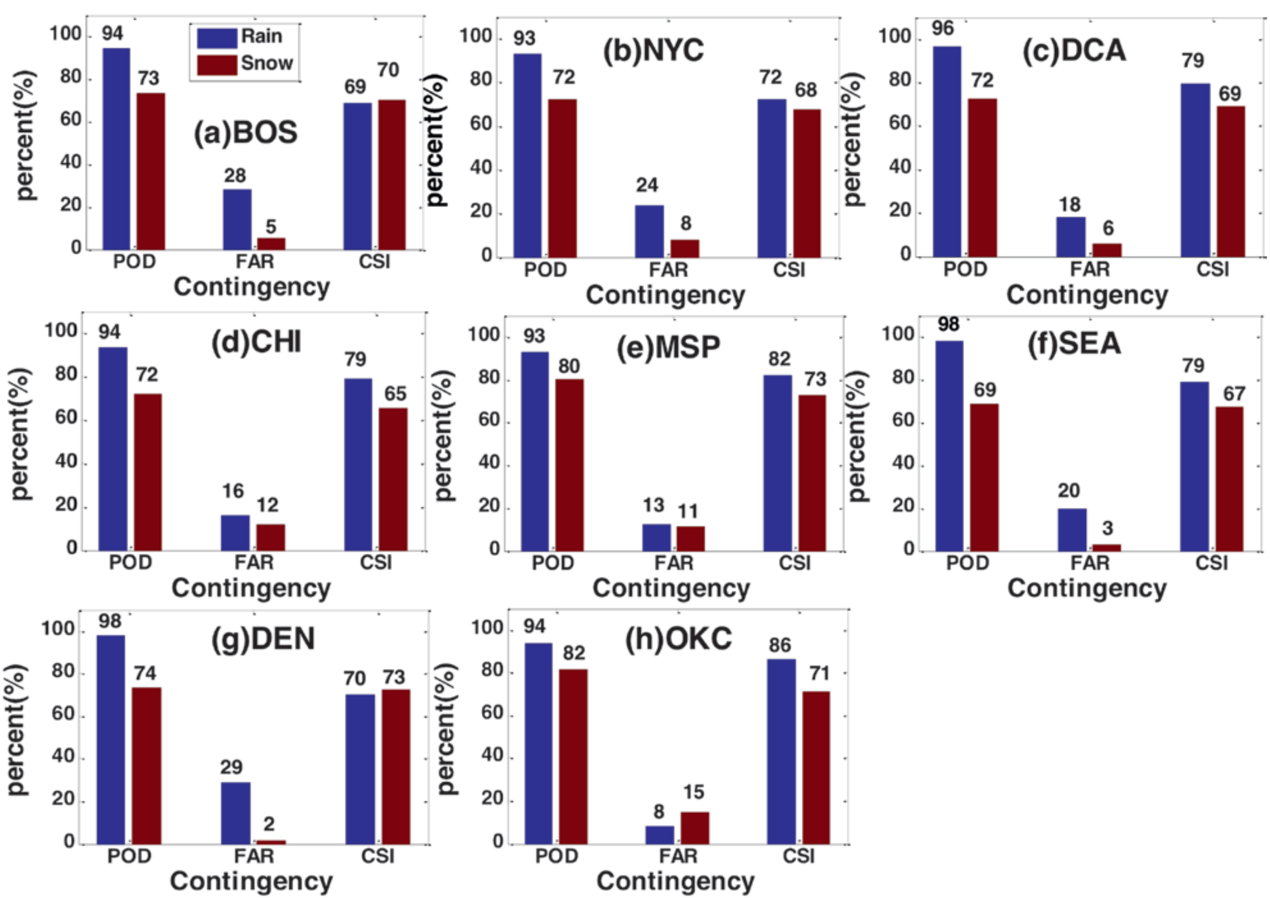

Fig. 5. Contingency table statistics (expressed in percent) for the MRMS rain and snow products across the conterminous United States using the mPING reports as the reference over high-density reporting regions in (a) Boston; (b) New York City; (c) Washington, D.C.; (d) Chicago; (e) Minneapolis-St. Paul; (f) Seattle; (g) Denver; and (h) Oklahoma City. POD is probability of detection, FAR is false alarm rate, and CSI is critical success index.

study. We can see from Fig. 4a that comparison of the MRMS precipitation type algorithm for pixels collocated with the mPING reports reveals a propensity for MRMS to yield too much rain relative to mPING reports. The same phenomenon for a single case was illustrated in Fig. 1. This indicates that the temperature thresholds used in the MRMS algorithm could be slightly modified in order to improve algorithm results by removing a bias toward estimating too much rain. Another potential explanation for the bias is that the MRMS precipitation types are based on the latest hourly RAP temperature analysis. It is possible that in the vicinity of a rapidly advancing cold front the MRMS precipitation type will be "delayed," causing the surface precipitation type to be incorrectly assigned to rain when it is actually snow. In any case, when we compute simple contingency table statistics using mPING reports as referenced in Fig. 4b, the MRMS algorithm has a critical success index (CSI) of 79% in rain and 72% in snow. The lower CSI in estimating snow is due to a degradation in the probability of detecting snow (POD=77% versus 94% for rain). This weakness might be attributed to the relatively shallow nature of winter snowstorms, making them less visible to radar with distance, and also due to the insensitivity of the radar signal in snow at S band due to weaker dielectric properties of snow.

Figure 5 is similar to Fig. 4b, but is specific to all the cities listed in Table 1 with large sample sizes. First, we note a high POD (> 93%) for rain in all cities. The snow POD is lower and ranges from 69% to 82%. The highest false alarm rates (FAR) in estimating rain relative to snow occur in Boston; New York City; Washington, D.C.; and Denver. These are cities that experience the most heavy snow events. This seemingly consistent characteristic is reversed, however, in Oklahoma City, which has a higher FAR with snow (15%) than with rain (8%). The FARs and CSIs with rain and snow are similar to each other in Oklahoma City, Minneapolis, and Chicago. These cities are in the central Plains and are affected by continental air masses, which might explain their similar statistical performances.

Overall, the relatively high CSI scores indicate that the MRMS precipitation type product is reliable and can serve as a useful reference to evaluate the performance of spaceborne active and passive sensors in detecting surface precipitation types. This is of particular importance to the current GPM mission. Users of the MRMS datasets may choose cases in which either the rain–snow line is clearly delineated (indicating sharp temperature gradients), or they can trust pixels that are well displaced from the rain–snow line itself. There is much greater algorithmic uncertainty in close proximity to regions that have temperatures quite close to the decision-tree thresholds.

SUMMARY AND CONCLUSIONS. To date, there are very few spatially consistent datasets with quantified errors in snow that can be used to evaluate the space-based algorithms. The primary objective of this study is to provide an initial evaluation of the MRMS precipitation type algorithm in segregating rain from snow. This benchmarking is of importance to all users of MRMS precipitation type products, which is timely given the launch of GPM and transition of MRMS to operations in the NWS. The evaluation is uniquely performed using unique reports from citizen scientists who use the mPING app freely available on GPS-enabled smartphones.

The MRMS precipitation type algorithm had overall good skill in detecting rain and snow at the surface but with lower detectability of snow. Multisensor approaches that incorporate observations from remote-sensing systems such as active and passive sensors on airborne and spaceborne platforms should be considered to improve the detectability of snow. There was a slight propensity for MRMS to yield rain in regions where it should have produced snow. This indicates a slight modification needed for the wet- and dry-bulb surface temperature thresholds (presently 0° and 2°C, respectively) to segregate rain and snow. Deterministic precipitation types from MRMS will have the greatest uncertainty in regions with temperatures that are close to the thresholds. Errors in surface precipitation type will also be prevalent for rain–snow lines that are progressing rapidly in relation to the hourly updates afforded by the RAP analysis. Current users of the MRMS surface precipitation type datasets should use caution in these regions of high algorithmic uncertainty. Future efforts will incorporate the model temperature fields and their spatiotemporal gradients in order to compute probabilistic surface precipitation-type fields.

The good performance (CSI ranging from 65% to 73% in snow and 69% to 86% in rain) and consistency in results from city to city give an indication that the mPING rain and snow reports provide useful information about the quality of the MRMS precipitation type algorithm. This study utilized mPING reports of rain and snow alone, with no consideration of the more ambiguous classes of mixed precipitation. Note that this study was performed using the raw reports without applying any quality control measures, such as validating reports with nearby observations, or from independent, automated station reports. Additional insights into algorithm behaviors are anticipated following quality control procedures that need to be developed for crowdsourced data. Some limitations of the citizen scientist reports include lack of significant sample sizes in sparsely populated regions. This means they may not be suitable for evaluating MRMS rain–snow segregation products in complex terrain. Further, the app requires a GPS-enabled smartphone and is presently only available in English. Future versions of the app will require foreign-language support in order to apply to regions outside the United States. Future work should investigate the other surface precipitation types beyond just rain and snow.

ACKNOWLEDGMENTS. This work was partially supported by HyDROS Lab at the University of Oklahoma. Partial funding was provided by the NOAA/Office of Oceanic and Atmospheric Research under NOAA-University of Oklahoma Cooperative Agreement NA17RJ1227.

FOR FURTHER READING

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.

Hou, A. Y., and Coauthors, 2014: The Global Precipitation Measurement Mission. *Bull. Amer. Meteor. Soc.,* **95,** 701–722, doi:10.1175/BAMS-D-13-00164.1.

Kirstetter, P.-E., and Coauthors, 2012: Toward a framework for systematic error modeling of spaceborne precipitation radar with NOAA/NSSL ground radar-based National Mosaic QPE. *J. Hydrometeor.,* **13,** 1285–1300, doi:10.1175/JHM-D-11-0139.1.

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

Zhang, J. K. Howard, 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.