ASCE

Evaluation of Multiradar Multisensor and Stage IV Quantitative Precipitation Estimates during Hurricane Harvey

Shang Gao, S.M.ASCE[1]; Jiaqi Zhang, S.M.ASCE[2]; Dongfeng Li, S.M.ASCE[3]; Han Jiang, S.M.ASCE[4]; and Zheng N. Fang, M.ASCE[5]

Abstract: Radar-based quantitative precipitation estimate (QPE) serves as input for flood forecasting, and its importance gets magnified during catastrophic storms, e.g., Hurricane Harvey in 2017. The record-breaking rainfall from Hurricane Harvey covered vast spatial extents and lasted for a 5-day period, providing a unique chance for evaluating radar errors, especially their spatiotemporal dependence. Using the rainfall data of Hurricane Harvey, the authors utilize a new method for sampling ground-based rainfall measurements over radar pixels (i.e., spatial reference rainfall) based on subpixel rainfall variability. The new method aims to enlarge the sample size and allow for compressively evaluating the QPE. Two hourly QPE products, the Next Generation Weather Radar (NEXRAD) Stage IV and Multiradar Multisensor (MRMS), are chosen for the evaluation due to their roles in major flood forecasting activities; and a dense rain gauge network covering the whole of Harris County, Texas, provides the spatial rainfall reference in this analysis. Comparative analyses are conducted based on Hurricane Harvey and other two flood-inducing storms occurring in 2015 and 2016 over Harris County. The results imply that the Stage IV and MRMS overestimate and underestimate, respectively, the total rainfall by a small factor, while both QPEs tend to overestimate very light precipitation. In addition, the study suggests that the spatial correlation of radar error from both QPEs be described as powered exponential functions of interpixel distance. This study also includes hydrologic simulations for an urban watershed, demonstrating the importance of both the accuracy and spatial resolution of QPE in representing the mean areal precipitation (MAP) over catchments. The insight gained from this study provides guidance for further improving the QPE performance, and the new sampling approach for spatial reference rainfall can be applied to comprehensively evaluate long-term radar rainfall data. **DOI: 10.1061/(ASCE)NH.1527-6996.0000435.**

Author keywords: Stage IV radar rainfall; Multiradar multisensor (MRMS) radar rainfall; Radar quantitative precipitation estimate (QPE) evaluation; Spatial reference rainfall; Hurricane Harvey.

Introduction

Hurricane Harvey made landfall as a category 4 hurricane along the middle Texas coast on August 25th, 2017. The storm then stalled with its center near the Houston–Galveston area for 4 days, generating historic amounts of rainfall (more than 60 in.) over Southeastern Texas (Blake and Zelinsky 2018). The extreme precipitation of Hurricane Harvey caused catastrophic flooding, inundating more than 300,000 structures and over 500,000 cars, and necessitated 120,000 rescues (Blake and Zelinsky 2018; FEMA 2017). During such chaos, the emergency management entities were overwhelmed and in urgent need of assistance. Accurate and timely flood forecasting could help emergency responders to prioritize the limited resources at hand and make the most effective decisions in crises. As the driving input for hydrological simulations, good-quality precipitation measurements play a crucial role in flood prediction. For instance, the rain gauge network has been conventionally used for flood warning. But due to the lack of spatial density of the rain gauge network, its role in various forecasting activities has been gradually replaced by quantitative precipitation estimate (QPE) from weather radar systems.

In the early to middle 1990s, the National Weather Service (NWS) installed the Next Generation Weather Radar (NEXRAD) system that currently comprises 160 WSR-88D radars across the United States (NCEI 2018). Over almost two decades, the NEXRAD precipitation products have undergone a series of improvements and are currently in the fourth stage. NEXRAD precipitation products (Stages I, II, III, and IV) from the River Forecast Centers (RFCs) have been applied to hydrometeorology (Smith et al. 2001, 2002; Zhang and Smith 2003), hydrologic analyses (Vieux and Bedient 1998; Bedient et al. 2000; Fang et al. 2008, 2011; Juan et al. 2015; Torres et al. 2015; Bass et al. 2016; Gao and Fang 2018), and remote sensing validation (Krajewski and Smith 2002; Habib and Krajewski 2002). The Stage IV product is a national mosaic of regional multisensor (combination of quality-controlled WSR-88D, satellite, and rain gauge data) precipitation estimates that are produced hourly at the NWS RFCs for operational hydrologic forecasting at the

[1]Graduate Student, Dept. of Civil Engineering, Univ. of Texas at Arlington, Arlington, TX 76019. Email: shang.gao@mavs.uta.edu

[2]Graduate Student, Dept. of Civil Engineering, Univ. of Texas at Arlington, Arlington, TX 76019. ORCID: https://orcid.org/0000-0003-1071-6742. Email: jiaqi.zhang@mavs.uta.edu

[3]Graduate Student, Dept. of Civil Engineering, Univ. of Texas at Arlington, Arlington, TX 76019. Email: dongfeng.li@mavs.uta.edu

[4]Graduate Student, Dept. of Civil Engineering, Univ. of Texas at Arlington, Arlington, TX 76019. ORCID: https://orcid.org/0000-0002-1905-9841. Email: han.jiang2@mavs.uta.edu

[5]Assistant Professor, Dept. of Civil Engineering, Univ. of Texas at Arlington, Arlington, TX 76019. (corresponding author). ORCID: https://orcid.org/0000-0001-9871-8405. Email: nickfang@uta.edu

Note. This manuscript was submitted on September 25, 2018; approved on August 25, 2020; published online on December 4, 2020. Discussion period open until May 4, 2021; separate discussions must be submitted for individual papers. This paper is part of the ***Natural Hazards Review***, © ASCE, ISSN 1527-6988.

Hydrologic Rainfall Analysis Project (HRAP) grid of approximately 4×4 km^2 spatial resolution (Lin and Mitchell 2005; Habib et al. 2009).

As another emerging radar-based weather sensing/monitoring system, the Multiradar Multisensor (MRMS) system has become operational at the National Centers for Environmental Protection (NCEP) since September 2014 (Zhang et al. 2014, 2016). By integrating over 180 radars, 7,000 hourly rain gauges, and numerical weather prediction outputs across the continental United States, MRMS has four types of QPE products including (1) radar-only QPE, (2) gauge-only QPE, (3) local gauge bias-corrected radar QPE (Q3gc), and (4) gauge and precipitation climatology merged QPE (Zhang et al. 2016). The hourly rain gauge data used for bias-correction in MRMS are quality-controlled from the HADS (2017a). The new national water model (NWM) that has been operational since August 2016 (NOAA 2016) utilizes hourly precipitation forcing in real time from the MRMS system to improve street-level water information services over the continental United States.

Precipitation measurements should be taken at a sufficiently high spatial and temporal resolution to represent the dynamic characteristics of storms. However, due to the indirect nature of radar measurement, radar-based rainfall products are subject to uncertainties that are inevitably propagated through further hydrologic simulations (Krajewski and Smith 2002). Therefore, previous studies were conducted to evaluate the radar rainfall products, especially the NEXRAD radar products. Habib et al. (2008) utilized a dense rain gauge network in Mississippi to evaluate various aspects of the errors in the Stage III products and the associated implications on streamflow simulation and concluded that bias adjustment can improve runoff prediction significantly. Wang et al. (2008) compared the Stage IV product with rain gauge measurements for a watershed in Texas and found that Stage IV products were better at rain detection than a rain gauge network for a studied watershed. Habib et al. (2009) used a dense rain gauge network in Louisiana to validate the Stage IV product and demonstrated its improved performance mainly due to the continuous algorithm update. These research efforts were motivated by a common focal point: how can point measurements of rainfall (rain gauge data) be used to represent the surface rainfall over an HRAP pixel (4×4 km^2) over a small temporal scale (say an hour), or simply how can hourly surface reference rainfall be accurately provided? Filtering criteria have been applied in these previous studies to select qualified radar pixels for evaluation, which might have resulted in the exclusion of useful data samples. In this study, the authors intend to simultaneously acquire the accuracy and the size of spatial reference rainfall data samples based on a new method and to demonstrate this approach with Hurricane Harvey (2017).

Hurricane Harvey was regarded as one of the most severe tropical cyclones in United States history according to spatial coverage and peak rainfall amount. The highest total rainfall recorded by a rain gauge during Hurricane Harvey was 154 cm (60.58 in.) in Nederland, northeast of Houston, which was nearly 9 in. higher than the previous record of 132 cm (52 in.) from Hurricane Hiki, in August 1950 (Blake and Zelinsky 2018). Fig. 1 shows a comparison of the total precipitation over the Harris County generated by Hurricane Harvey and other two flood-inducing storms in Houston (2015 Memorial Day storm on May 25th, 2015, and 2016 Tax Day storm on April 17th, 2016) based on the MRMS gauge-corrected rainfall (Q3gc product). The comparison highlights the exceptional total rainfall amount of Hurricane Harvey due to its intensity and, more importantly, its 5-day duration. In addition, the areal extent of heavy rainfall from Hurricane Harvey was truly overwhelming, with almost the entire Harris County receiving over 70 cm (about 2.3 ft) of rainfall. From the perspective of QPE evaluation, any rainfall events with large areal coverages and long durations tend to generate a large number of data samples. Therefore, Hurricane Harvey provides an opportunity for the authors to investigate spatial and temporal structures of radar errors that would otherwise be hindered using small-scale short-duration storms.

Despite its importance, accurate QPE does not necessarily guarantee the optimal representation of precipitation as an input to hydrologic models. This is because the mean areal precipitation (MAP) is the actual input to catchments or grid boxes that make up the watershed or model domain in hydrologic models. Given that the true MAP cannot be measured easily, comparisons between the simulated flow/stage from hydrologic models and the observed values can indirectly reflect the uncertainty associated with MAP. Previous research efforts have been invested in examining the uncertainty within streamflow simulation induced by radar error (e.g., Yilmaz et al. 2005; Habib et al. 2008; Gourley and Vieux 2005). These studies were conducted upon the premise that the MAP is the primary, if not the only, contributor to the uncertainty in model output. In other words, uncertainties from other sources (e.g., model structure, model parameters, and state variables) need to be minimized. For instance, if a hydrologic model is well calibrated for a highly urbanized watershed with saturated soil, the uncertainty in simulated streamflow can then be linked to that in MAP instead of an infiltration process. During Hurricane Harvey, an urban watershed, Brays Bayou in Harris County, was considered a sufficient study area, as heavy rainfall rendered the soil fully saturated. Therefore in this study, the Brays Bayou watershed during Hurricane Harvey was selected to investigate the implication of MAP estimation on hydrologic simulation.

In this study, an innovative sampling approach is introduced to increase the number of references for evaluating radar rainfall products. More references in turn allow us to examine the conditional performance of radar data as well as the spatial and temporal structures of rainfall error. Given the importance of Stage IV and MRMS hourly QPEs in operational flood forecasting, the authors are motivated to comprehensively evaluate their performance during Hurricane Harvey to achieve better preparedness and decision making for future floods. Hurricane Harvey, as an unprecedented tropical rainfall event, presented a unique research opportunity for the authors to demonstrate the new sampling method in terms of increasing sample size and a better understanding of the spatiotemporal structure of radar error via the following objectives:

1. To construct the spatial reference rainfall datasets over the scale of the radar pixels from a dense rain gauge network during the 2015 Memorial Day storm, 2016 Tax Day storm, and 2017 Hurricane Harvey in Harris County, Texas;
2. To examine the radar errors in terms of bias, conditional dependence on rainfall intensities, and spatiotemporal structure; and
3. To investigate the implication of radar errors for the accuracy of hydrologic simulation and prediction by analyzing the runoff behaviors simulated by hydrologic models of Brays Bayou in Harris County.

Study Area and Data

Harris County is the third-most populous county in the United States and includes the largest city in Texas—Houston. Fig. 2 shows the area of this study, which is Harris County with 4,602 km^2 located in the State of Texas. Fig. 3 shows the Brays Bayou watershed, where the hydrologic simulation is examined in this study. As one of the flood-prone urban watersheds, Brays

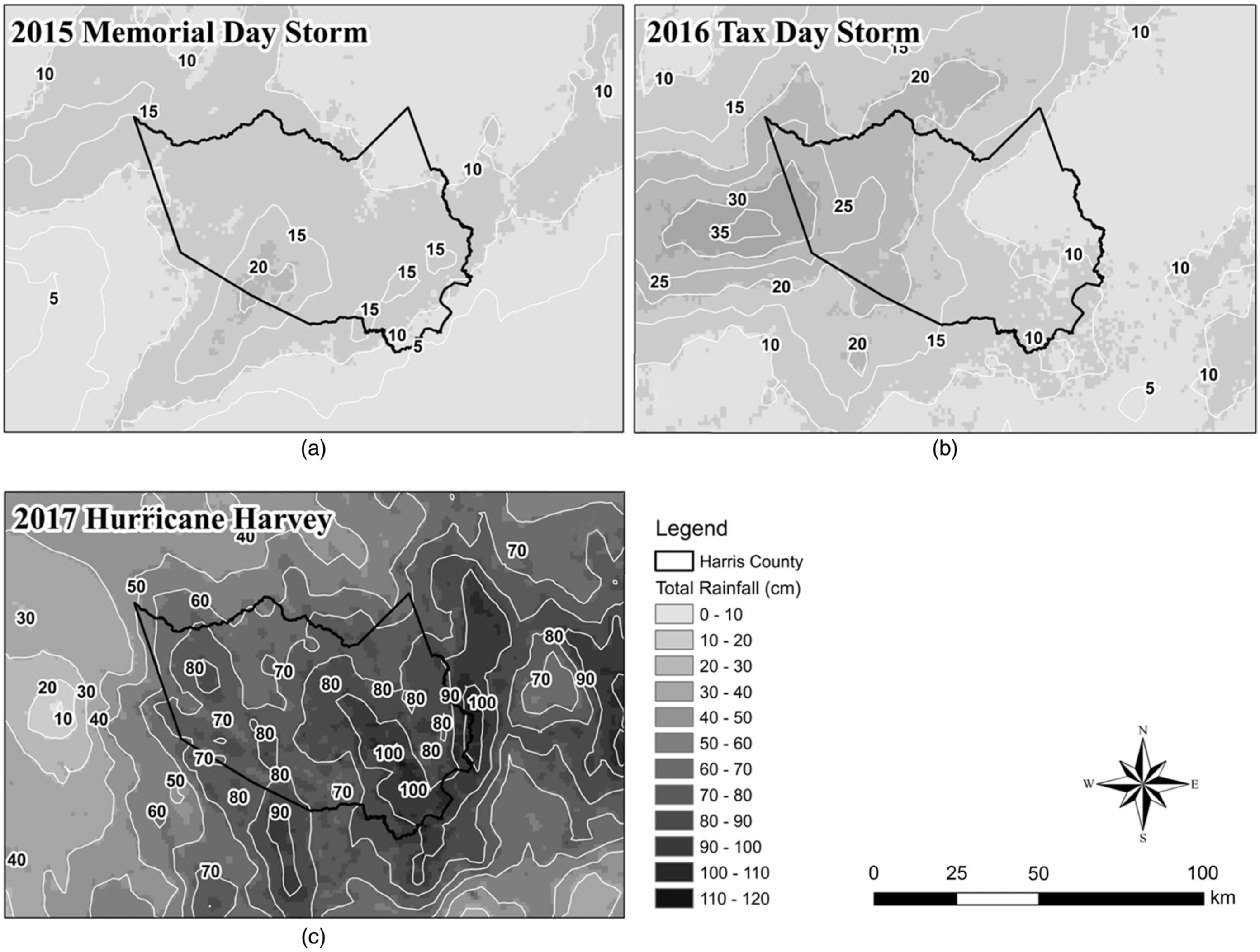

Fig. 1. Cumulative rainfall based on the MRMS data of (a) 2015 Memorial Day storm on May 25th, 2015; (b) 2016 Tax Day storm on April 17th, 2016; and (c) Hurricane Harvey on August 25th, 2017, respectively.

Bayou is 95% developed with a population of more than 722,000 people, making it one of the most urbanized watersheds in Harris County, Texas (HCFCD 2017b). The high tendency of flooding in this watershed is due to flat slopes, impermeable land surface and clay soils, and the explosive rainfall (Bedient et al. 2003, 2007; Fang et al. 2008, 2011, 2014; Bass et al. 2016; Gao and Fang 2018). There are four junctions with reliable streamflow observation during Hurricane Harvey from United States Geological Survey (USGS) gauges (**Junction 1**/USGS8074760@Belle Park Dr., **Junction 2**/USGS8074810@S. Gessner Rd., **Junction 3**/USGS8075000@Main St., and **Junction 4**/USGS8075110@MLK Blvd.).

The rain gauge observations were collected from the website of the HCFWS (2017c). This flood warning system, which measures rainfall amounts and monitors water levels in bayous and major streams on a real-time basis, is operated and maintained by the Harris County Flood Control District (HCFCD). The system relies on 154 gauge stations strategically placed throughout Harris County bayous and their tributaries. For QPE evaluation, the surface reference rainfall needs to be independent of QPE products (Habib et al. 2009). In other words, the rain gauges serving as truth should not overlap those used by NWS in developing or adjusting the bias of QPE estimates. Therefore, the HADS rain gauges (white dots in Fig. 2) are not utilized for evaluating the radar QPE, as they are already incorporated in processing the Stage IV and MRMS hourly gauge-corrected products.

The radar rainfall data used in this study are the Stage IV product and the MRMS Q3gc product, which are later referred to as Stage IV and MRMS, respectively. The Stage IV data are provided by the West Gulf River Forecast Center (WGRFC), whose service boundary fully encompasses the study area, Harris County. The MRMS are radar-only estimates locally adjusted by hourly HADS gauge data using an inverse distance weighting (IDW) scheme (Zhang et al. 2016; Cocks et al. 2017). As aforementioned, the Stage IV and MRMS have the same temporal resolution of one hour, but different spatial resolutions of 4×4 km^2 and 1×1 km^2, respectively. From the available period of both Stage IV and MRMS, three severe storm events, 2015 Memorial Day storm, 2016 Tax Day storm, and Hurricane Harvey were selected for analysis, with their starting and ending times summarized in Table 1.

Approach and Methods

Surface Reference Rainfall

MRMS has a high spatial resolution of 1×1 km^2 and thus are directly compared with rain gauge measurements in this study.

Nat. Hazards Rev., 2021, 22(1): 04020057

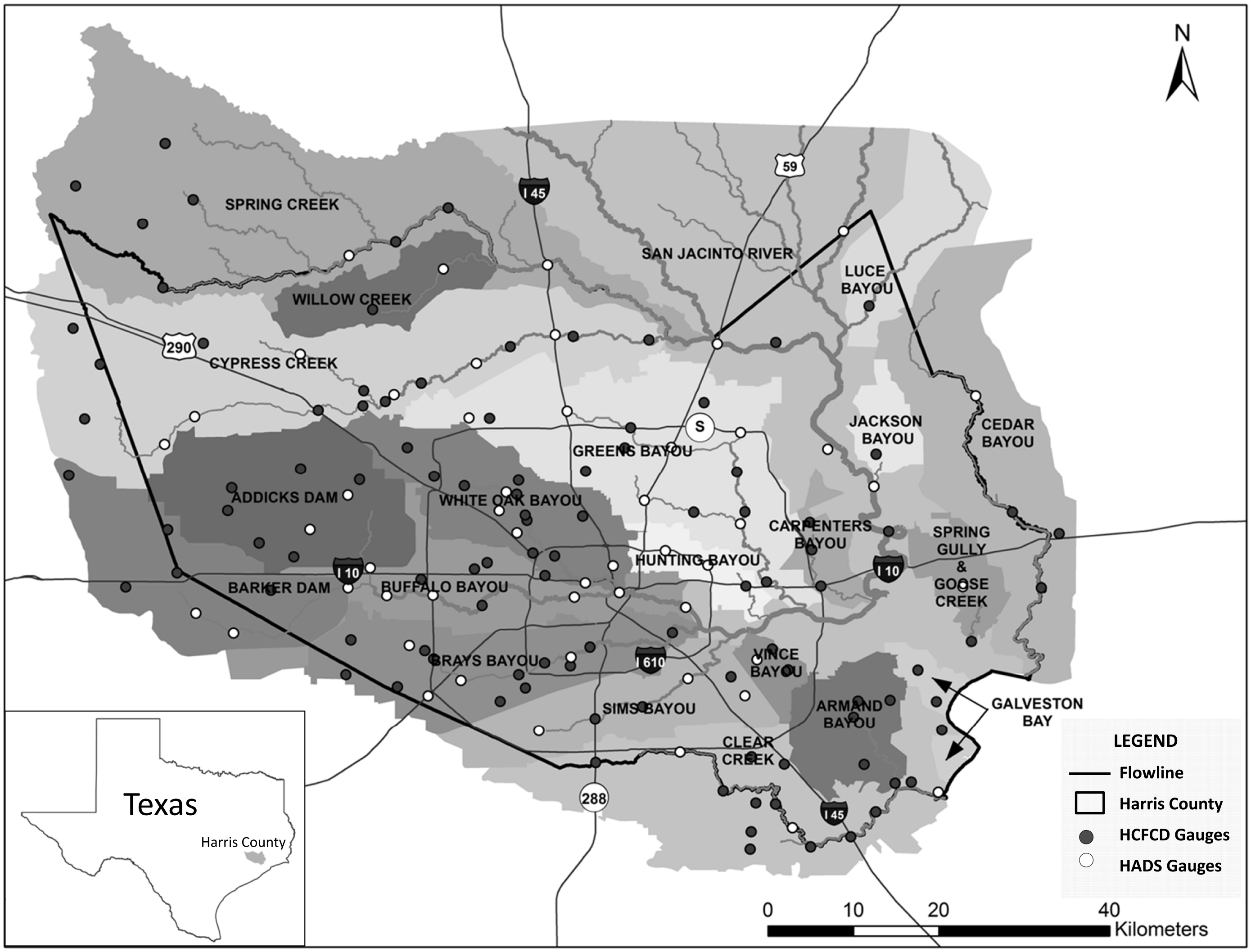

Fig. 2. Study area and the rain gauge network.

However, for Stage IV, it is recognized that rain gauges at hourly or smaller scales may be limited by their near-point sampling and may not provide an acceptable approximation of surface rainfall over the 4×4 km^2 pixel (Habib et al. 2009). Therefore, the critical issue in evaluating hourly Stage IV QPE has been how to get a measurement representing the areal average rainfall over an HRAP pixel (Kitchen and Blackall 1992). In previous studies, a high-density rain gauge network (4–10 gauges within each HRAP pixel) has been used (Ciach and Krajewski 1999; Habib and Krajewski 2002; Ciach et al. 2003; Habib et al. 2004), which is however limited in spatial coverage and costly to implement and maintain. Wang et al. (2008) utilized a method to select Stage III hourly radar estimates (also with 4×4 km^2 resolution) only during uniform rainfall for evaluation. Although Wang et al. (2008) managed to utilize even sparse rain gauge network (e.g., one gauge per HRAP pixel), their definition for uniform rainfall is not based on the spatial variability of rainfall within the target radar pixel (subpixel variability) but that among the target radar pixel and its eight neighboring pixels (interpixel variability).

In order to make improvements, the authors investigate the subpixel spatial variability of hourly rainfall within an HRAP pixel and further determine if the gauge(s) within the HRAP pixel can sufficiently represent its areal average rainfall intensity. This new method features evaluating one type of radar QPE of coarser spatial resolution (i.e., Stage IV) using another kind of radar QPE of finer spatial resolution (i.e., MRMS). In essence, Stage IV pixels are compared to the corresponding rain gauge value(s) only when rainfall is spatially uniform within this Stage IV pixel. This is conditioned on the comparison between the MRMS QPE values at the gauge(s) and those encompassed by the HRAP pixel boundary. Fig. 4 illustrates this sampling scheme where the black square in bold is an HRAP pixel of interest; the grey squares are the MRMS pixels encompassed by the HRAP pixel, and the three shaded grey squares represent MRMS pixels with rain gauges inside. At a given hour, if the averaged MRMS rainfall of the grey squares (including the shaded ones) is sufficiently close to (90%–110% of) the averaged MRMS rainfall of the shaded grey squares, it will then be determined that the averaged rain gauge measurements can represent the MAP of the HRAP pixel. Here, the authors employ a 10% threshold to determine if the mean of the rain gauge measurements could approximate the MAP of the HRAP pixel. It is worth noticing that this method does not necessarily assume that MRMS QPE in the HRAP pixel is true or unbiased but that the MRMS QPE can adequately preserve the spatial variability of rainfall within the 4×4 km scale. By applying the selection scheme to all the HRAP pixels containing at least one rain gauge and for all the hours during

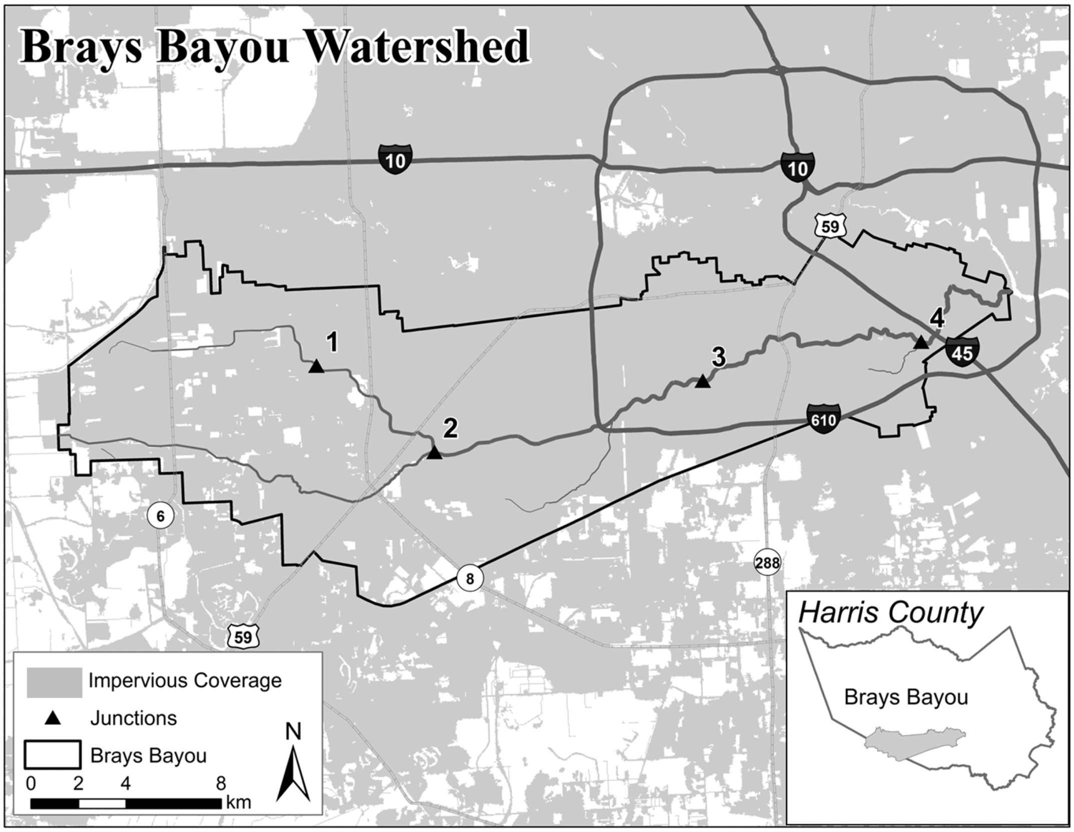

Fig. 3. Brays Bayou watershed, Houston, Texas.

Table 1. Starting and ending times for the three selected historical storms

Storm	Start time (CDT)	End time (CDT)	Duration (h)
2015 Memorial Day Storm	May 25, 2015 19:00	May 27, 2015 7:00	36
2016 Tax Day Storm	April 17, 2016 13:00	April 19, 2016 6:00	41
2017 Hurricane Harvey	August 25, 2017 00:00	August 30, 2017 00:00	120

storm events, the authors can then make the best use of data to comprehensively evaluate the QPEs via statistical metrics introduced in the following section. In comparison, the new method can generate a spatial reference rainfall sample three times the size of that from the traditional method in which only the HRAP pixels with two or more gauges are selected.

Statistical Metrics

Error

Error is defined as the deviation of radar rainfall estimates from the rain gauge observations, as shown by Eq. (1)

$$\varepsilon = \frac{R_e}{R_o} \tag{1}$$

where R_e and R_o = hourly rainfall intensities from radar QPEs and rain gauges, respectively. Due to Eq. (1), an error that is greater (smaller) than 1 means overestimation (underestimation).

Overall Bias

In order to investigate the systematic performance of the QPEs, overall bias (OB) is used to measure the averaged deviation of the rainfall estimates (R_e) from observations (R_o), as represented by Eq. (2)

$$OB = \frac{E[R_e]}{E[R_o]} \tag{2}$$

In this study, OB is calculated with respect to individual storms, which means expectation ($E[]$) is calculated by temporally averaging the estimates or observations over the duration of a storm event. To examine the spatial variability of OB, the authors calculate OBs for each radar pixel separately. In addition, for radar pixels that encompass multiple rain gauges, the R_o is the arithmetic mean of the values from corresponding gauges.

Conditional Bias

It has been found by previous researchers that radar rainfall bias can depend on the magnitude of estimated rainfall intensity

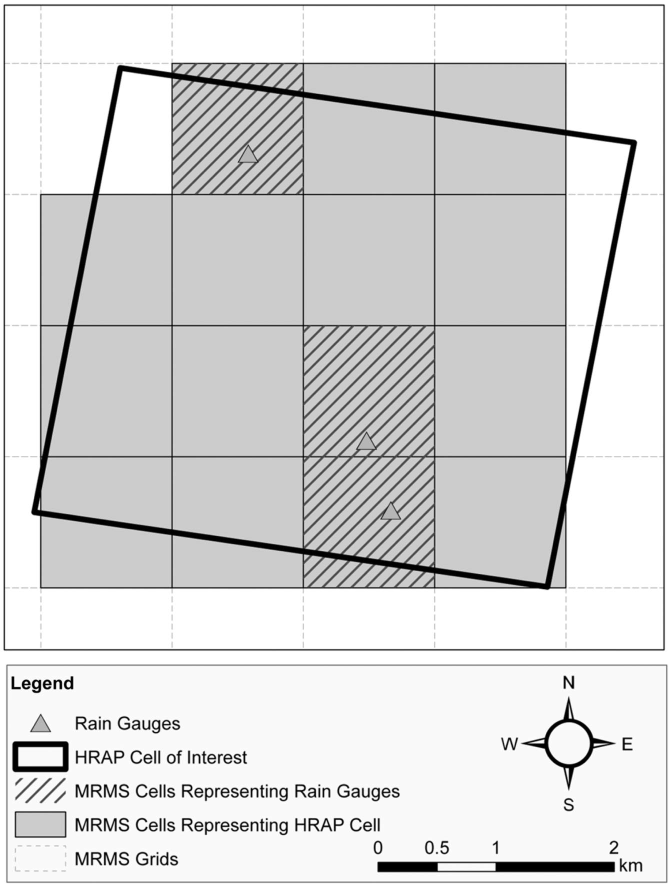

Fig. 4. Selection scheme for hourly spatial reference rainfall.

(Ciach et al. 2000). Using moving average windows, the authors calculate the conditional bias (CB) using Eq. (3)

$$CB = \frac{E[R_e|a < R_e \leq b]}{E[R_o|a < R_o \leq b]} \tag{3}$$

where a and b = lower and upper limits of the moving average window.

Spatial and Temporal Autocorrelations of Error

In addition to bias analysis, spatial and temporal dependence of error need to be assessed to infer the adequacy of bias adjustment in improving radar QPEs. The authors estimate the spatial autocorrelations of ε from marginal samples of ε at each time step, while the temporal autocorrelations are based on marginal samples of ε at all selected pixels. Eq. (4) is the Moran's I (Moran 1950), as a measure of the spatial autocorrelation

$$I(d) = \frac{\frac{1}{W}\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(\varepsilon_i - \bar{\varepsilon})(\varepsilon_j - \bar{\varepsilon})}{\frac{1}{n}\sum_{i=1}^{n}(\varepsilon_i - \bar{\varepsilon})^2} \tag{4}$$

where $I(d)$ = Moran's I as a function of distance d; ε_i and ε_j = errors at location i and j; w_{ij} = weight of 0 or 1, 1 meaning that ε_i and ε_j are within a given distance class and 0 being all the other cases; and W = sum of all w_{ij}; and n = sample size. The temporal autocorrelation (Box and Jenkins 1976) is defined as Eq. (5):

$$r(\tau) = \frac{\sum_{i=1}^{m-\tau}(\varepsilon_i - \bar{\varepsilon})(\varepsilon_{i+\tau} - \bar{\varepsilon})}{\sum_{i=1}^{m}(\varepsilon_i - \bar{\varepsilon})^2} \tag{5}$$

where $r(\tau)$ = autocorrelation of lag τ; ε_i and $\varepsilon_{i+\tau}$ = errors at ith hour and $(i+\tau)$ th hour; and m = sample size. It should be noted that only the marginal samples with sufficient size ($>= 200$) are included in the analysis. This is the very reason why Hurricane Harvey provided an opportunity for investigating the spatiotemporal structure of radar error, as the storm covered vast areas and lasted for a total of 5 days, generating large samples of radar error information. In previous studies (Kessler and Neas 1994; Habib et al. 2001b), sample size has been a limitation for estimating correlations at large spatial and temporal lags. However, the authors take advantage of Hurricane Harvey and present a more complete spatiotemporal structure of radar error during this event.

Hydrologic Simulation

A hydrologic (Hydrologic Engineering Center Hydrologic Modeling System, HEC-HMS) model is used to simulate the hydrologic response from Brays Bayou during Hurricane Harvey. The HEC-HMS model is part of the products from the Tropical Strom Allison Recovery Project (TSARP), which was initiated by the devastating impact from Tropical Storm Allison (2001). Calibration effort was invested to improve the hydrologic simulation of this model in several studies (Fang et al. 2011; Bass et al. 2016; Gao and Fang 2018). Fang et al. (2011) conducted an analysis using storm events with accumulated rainfall ranging from 3.3 cm (1.3 in.) to 19.5 cm (7.8 in.) and found that the model had predicted floods with an average of 3.6% difference in peak flows and an R^2 value of 0.90 for the overall performance from 2002 to 2010. Bass et al. (2016) updated soil/land use information in the model to best represent the actual land use conditions. Gao and Fang (2018) validated the model performances at all four USGS gauge locations during the 2015 Memorial Day storm and the 2016 Tax Day storm using the MRMS Q3gc product. With the well-calibrated HEC-HMS model, the authors compare simulated hydrographs from three rainfall input data, i.e., Stage IV, MRMS, and rain gauge. The two radar QPE products (Stage IV and MRMS) are processed into time series of MAP for each subbasin using the Hydrologic Engineering Center Meteorological Visualization Utility (HEC-MetVUE) program, while the rain gauge records are allocated into each subbasin using the Thiessen Polygon method (Brassel and Reif 1979). Using the runoff volume received by the most downstream USGS gauge and cumulative basin-averaged rainfall (rain gauge), the runoff coefficient is estimated to be 91%. This is largely due to the fact that Brays Bayou is 95% developed and 72% impervious and most importantly because of the enormous rain volume from Hurricane Harvey. Given the minor role of infiltration, the initial soil moisture is assumed to be fully saturated in the Green and Ampt method. Four statistics are used to quantitatively evaluate model performance in runoff and streamflow, as shown in Eqs. (6)–(9)

$$\text{Runoff volume error: } V_e = \frac{\sum_{i=1}^{n} Q_i^{sim}}{\sum_{i=1}^{n} Q_i^{obs}} - 1 \tag{6}$$

$$\text{Peak flow error: } P_e = \frac{Q_{\max}^{sim}}{Q_{\max}^{obs}} - 1 \tag{7}$$

$$\text{Root mean square error (RMSE): } RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Q_i^{obs} - Q_i^{sim})^2} \tag{8}$$

Nash-Sutcliffe model efficiency coefficient (NSE): NSE

$$= 1 - \frac{\sum_{i=1}^{n} (Q_i^{sim} - Q_i^{obs})^2}{\sum_{i=1}^{n} (Q_i^{obs} - \overline{Q_i^{obs}})^2} \quad (9)$$

where n = number of hours in the hydrographs; Q = runoff discharge, with the subscript '*max*' denoting the peak value; and the superscripts '*sim*' and '*obs*' = simulation and observation, respectively. The operator $\overline{Q_i^{obs}}$ is the arithmetic averaging of the discharge observation.

Results and Discussion

As an overview of all data samples involved in this analysis, Fig. 5 shows the scatter plots of Stage IV and MRMS hourly QPE against rain gauge data for the three investigated storm events (2015 Memorial Day storm, 2016 Tax Day storm, and 2017 Hurricane Harvey), along with the coefficient of determination (R^2), root square mean error (RMSE), and sample size (upper right table). The difference in sample size results from the Stage IV selection scheme that filters out about 50% of the data samples. The results show that both QPEs (Stage IV and MRMS) reached high R^2 values (over 0.8), showing good overall performance during the storms. Given that there is no clear indication of biases for either QPE solely based on the scatter plot, the authors conduct the following analysis in OB.

Overall Bias

OB is quantified based on Eq. (2) for each radar pixel and each storm to examine any spatial variability and inter-storm variability. Figs. 6 and 7 show the maps for OB calculated at selected radar pixels of Stage IV and MRMS for the 2015 Memorial Day storm (**6A** and **7A**), 2016 Tax Day storm (**6B** and **7B**), and 2017 Hurricane Harvey (**6C** and **7C**). The figures show that there is no distinct spatial pattern of OB for either QPE or any storm event. The mean OB values from individual storms and all storms combined indicate overestimation by Stage IV and underestimation by MRMS, except for the case of Stage IV during the 2016 Tax Day storm. In the case of Stage IV, approximately 43%, 66%, and 30% of the data samples have OB values smaller than 1 (signified by the white dots) for the 2015 Memorial Day storm, 2016 Tax Day storm, and 2017 Hurricane Harvey, respectively. In comparison, the majority of the OB values for MRMS are smaller than 1, with 95%, 90%, and 90% of the data samples represented by white dots for the three storms.

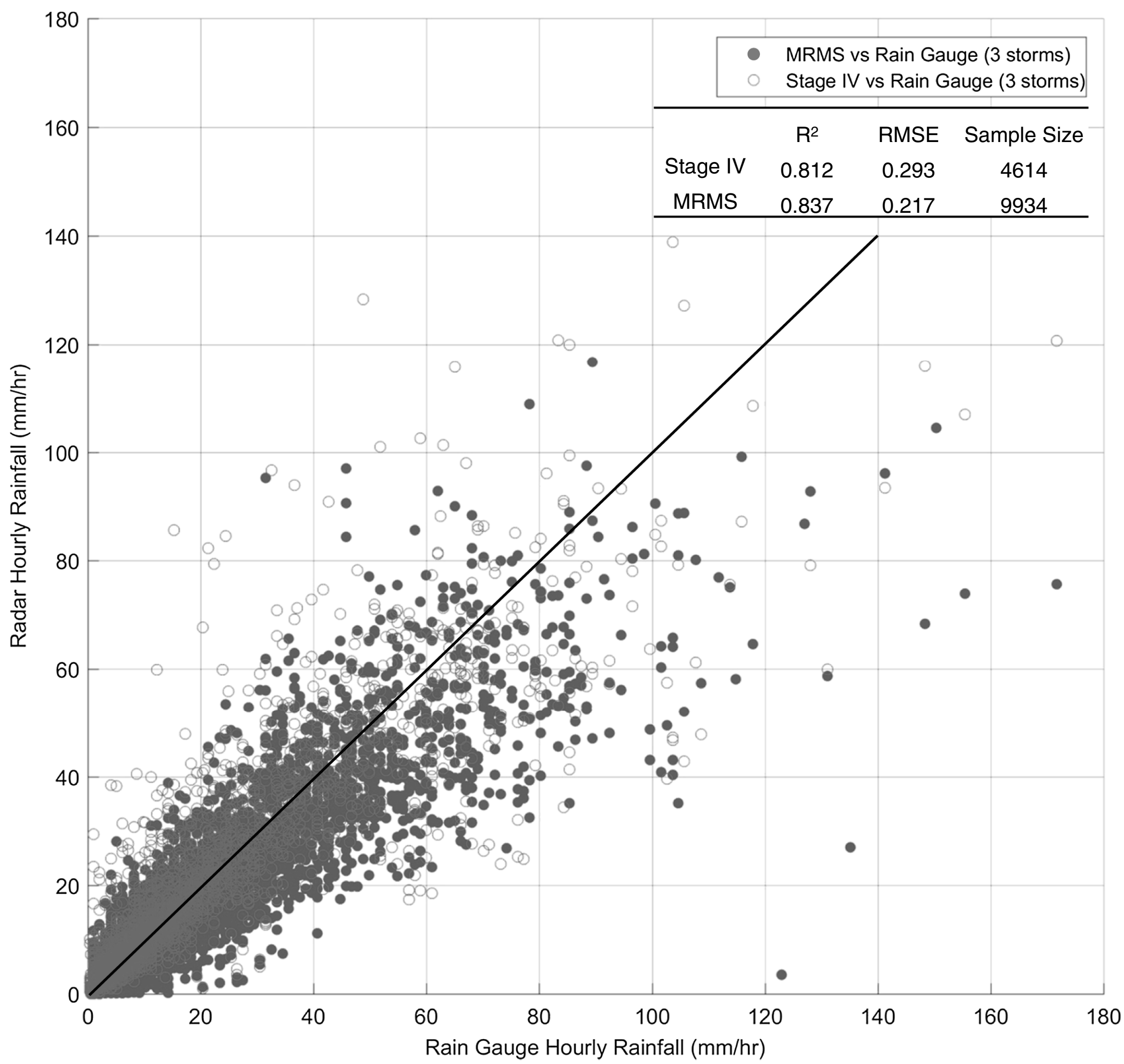

Fig. 5. Scatter plots of Stage IV and MRMS hourly QPE from the three storm events combined.

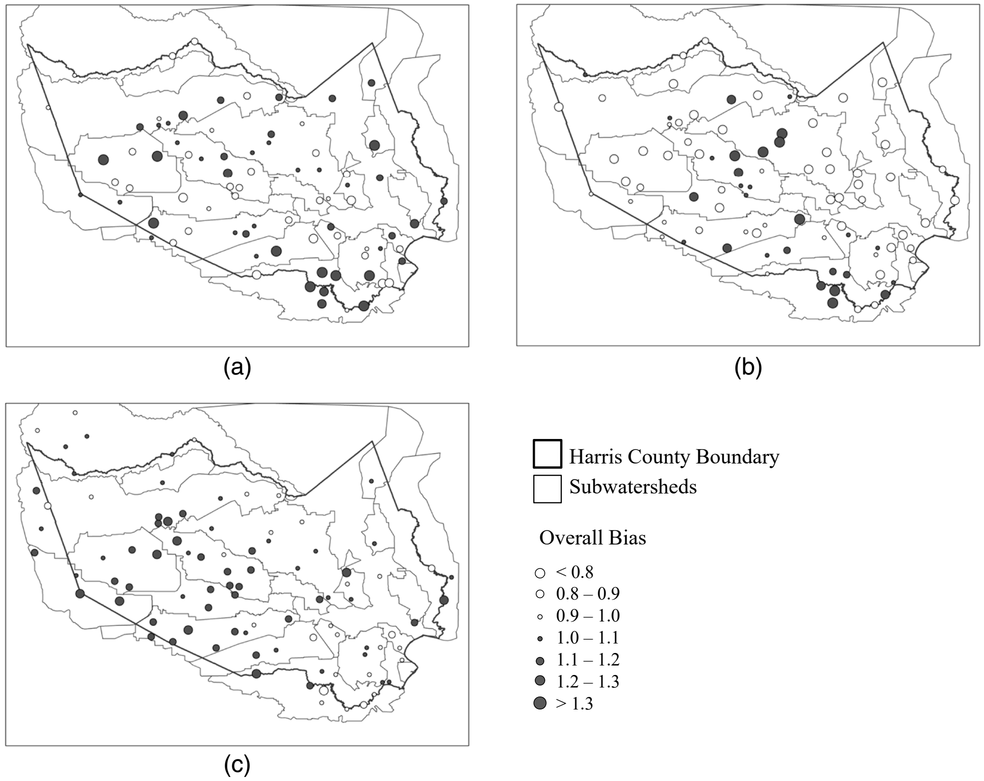

Fig. 6. OB of Stage IV for (a) 2015 Memorial Day Storm; (b) 2016 Tax Day Storm; and (c) 2017 Hurricane Harvey, respectively.

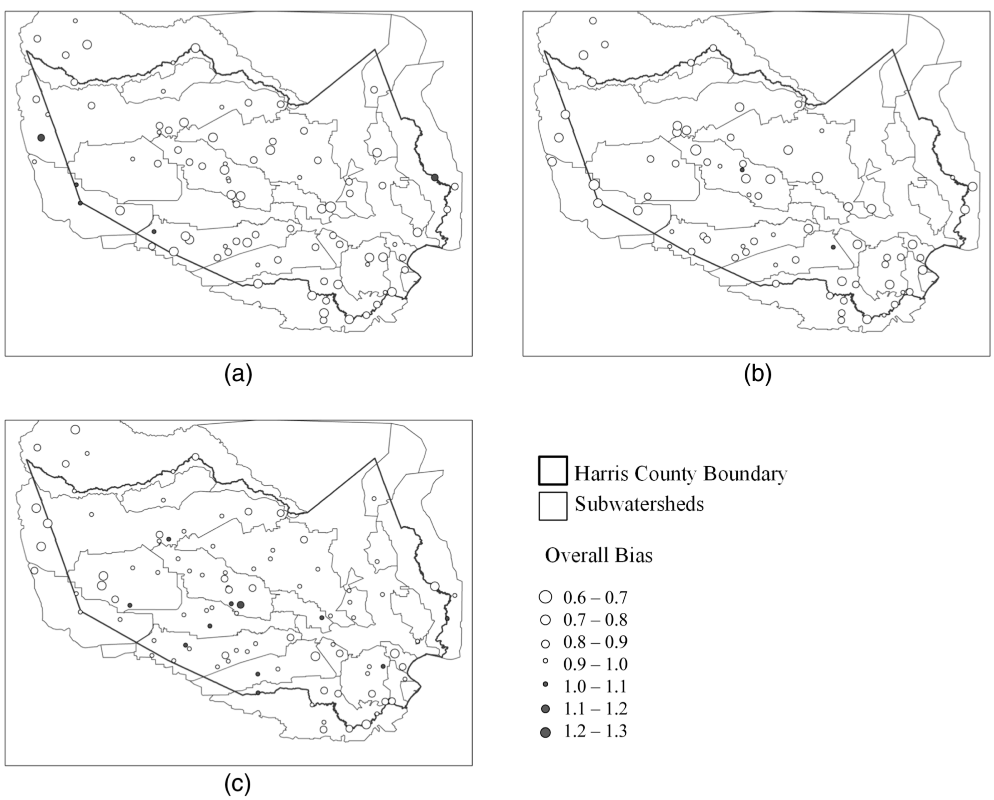

Fig. 7. OB of MRMS for (a) 2015 Memorial Day Storm; (b) 2016 Tax Day Storm; and (c) 2017 Hurricane Harvey, respectively.

The statistics, including average, min, max, and total sample size of the OB values for each storm separately and for all storms combined are also summarized in Table 2. It is worth noticing that the distinctly large sample size from Hurricane Harvey (2017) for both QPEs is due to the vast spatial coverage and long duration of the storm. In summary, MRMS shows the tendency of underestimating precipitation by a factor of 12% (OB = 0.88) for three storms combined, while Stage IV shows better performance in

Table 2. OB summary table

Rainfall product	Storm event	Average OB	Minimum OB	Maximum OB	Total sample size
MRMS	2015 Memorial Day	0.85	0.63	1.20	952
	2016 Tax Day	0.78	0.43	1.07	1,377
	2017 Hurricane Harvey	0.92	0.67	1.14	7,607
	All	0.88	0.43	1.20	9,936
Stage IV	2015 Memorial Day	1.12	0.62	3.70	417
	2016 Tax Day	0.93	0.49	2.08	520
	2017 Hurricane Harvey	1.06	0.77	1.28	3,677
	All	1.02	0.49	3.70	4,614

capturing the mean (temporally averaged) rainfall amount with an overestimation of 2% (OB = 1.02).

Conditional Behavior of Radar Error

While overall bias represents the average behavior of radar error, the conditional behavior of radar error depending on the rainfall intensity has been demonstrated by previous researchers and thus needs to be investigated (Ciach et al. 2000, 2007; Habib et al. 2008). Therefore, in this study, the authors also examine the radar errors conditioned on hourly rainfall intensity from rain gauges. Since this analysis is event-based and lacks a large volume of data, it is important to consider the distribution of utilized data samples before investigating the conditional behavior of radar error. Therefore, the authors present a 2D histogram of the combined data samples from all three storms plotted on log-scale radar error and log-scale hourly gauge rainfall intensity for Stage IV and MRMS in Figs. 8(a and b) respectively. The log-scale axes are used in plotting because both radar error and rainfall intensity are log-normally distributed. The results show that the two data samples (Stage IV and MRMS) share a similar distribution pattern along the X-axis, meaning Stage IV data samples, though smaller in size, do capture the same distribution of rainfall intensities as the MRMS data samples. Figs. 8(a and b) show that the sample population is divided by a gap located just above 1.5 mm/h. This is because readings from tipping bucket rain gauges are the number of tipping multiplied by the unit volume of each tip (in this case 0.04 in. or 1.016 mm). As for Stage IV [Fig. 8(a)], these data samples mostly show overestimation with radar errors over one, while in the case of MRMS [Fig. 8(b)], they seem to evenly distribute around one.

Keeping the data sample distribution in perspective, the authors further calculate the CB for both QPEs using the same measures to divide the sample population (equal interval in log-scale) as in the 2D histograms. Figs. 9(a and b) show the CB against the hourly gauge rainfall intensity both plotted in log-scale for Stage IV and MRMS, respectively. The results show that Stage IV overestimates very light rainfall (<1.5 mm/h), and the overestimation decreases with increasing rainfall intensity; Stage IV exhibits steady and good performance with a slight overestimation when rainfall ranges from 3.5 to 25 mm/h. In comparison, MRMS shows a small overestimation for rainfall lighter than 1.5 mm/h and a steady but slight underestimation for rainfall from 3.5 to 25 mm/h. Despite the limited spatial and temporal extent of data in this analysis, the findings echo with those reported by previous researchers. For instance, Nelson et al. (2016) conducted a comprehensive assessment of Stage IV for the Continental United States (CONUS) over the period 2002–2012 and found that larger overestimation exists for light rainfall for all RFCs and all seasons. The overestimation of light rainfall by the MRMS was founded by Cocks et al. (2017), who evaluated the MRMS Q3gc (gauge corrected) products east of the Rockies during the 2014 warm seasons. For both Stage IV and MRMS, this common issue of overestimating light rainfall is probably due to precipitation evaporating before reaching the gauge and gauge wetting losses, as speculated by previous researchers (Catizone et al. 2014; Cocks et al. 2017).

Spatial and Temporal Structure of Radar Error

Although the conditional behavior of radar error is recognized in the prior section, it is unclear whether the radar errors have distinct spatial and temporal dependence. As shown in Figs. 6 and 7, the maps of OB show no explicit spatial pattern in terms of the averaged deviation from the radar estimates to the gauge measurements. However, it is still possible that the radar errors can exhibit clustering patterns in space; therefore, it is necessary to decipher the spatial dependence of the radar error using spatial autocorrelations. Figs. 10(a and b) respectively show the spatial autocorrelation coefficients (Moran's I) of Stage IV and MRMS radar errors (ε) calculated based on Eq. (4) at intervals of radar pixel size (4 km for Stage IV and 1 km for MRMS) for the three storms. The sample size used for calculating each coefficient is also presented as the inverted bars in Figs. 10(a and b). The results show that Moran's I values are based on marginal samples collected at all time steps and only those with sample sizes larger than 200 are displayed; the remarkably larger sample size from Hurricane Harvey, as shown in both Figs. 10(a and b), is mainly caused by its 5-day duration. Due to the clear difference in sample size, Hurricane Harvey ought to yield the most representative result among the three storms. For Hurricane Harvey (2017), the Moran's I values for both QPEs are relatively high at short spatial lag, with Stage IV yielding 0.68 at 4 km and MRMS yielding 0.7 at 2 km. In addition, the Moran's I values, in all six cases (two QPEs and three storms), are fitted to powered exponential functions, meaning that the correlation of the radar error ε decays exponentially with increasing spatial spacing. This finding widely echoes with previous studies in radar error modeling where the spatial correlation function is parameterized by fitting a two-parameter power exponential function (Mandapaka et al. 2010; Dai et al. 2014; Ko et al. 2018).

In a similar fashion as Figs. 10(a and b), the temporal autocorrelations of radar error are presented in Figs. 11(a and b), except that the only results from Stage IV during Hurricane Harvey and MRMS during the 2016 Tax Day storm and Hurricane Harvey are displayed because of the sample size limit (>200). Due to its vast spatial coverage, Hurricane Harvey generates the most marginal samples collected at the selected radar pixels, as indicated by the inverted bars. The results show that no significant autocorrelation exists at any temporal lag for the storms or for either QPE, meaning that no persistence is observed in the temporal variation of radar error. Previous studies seemed to diverge regarding the temporal structure of radar errors: some found insignificant temporal correlations as this study does (e.g., Habib et al. 2008); while other studies on radar error modeling utilized autoregressive lag-one model

Nat. Hazards Rev., 2021, 22(1): 04020057

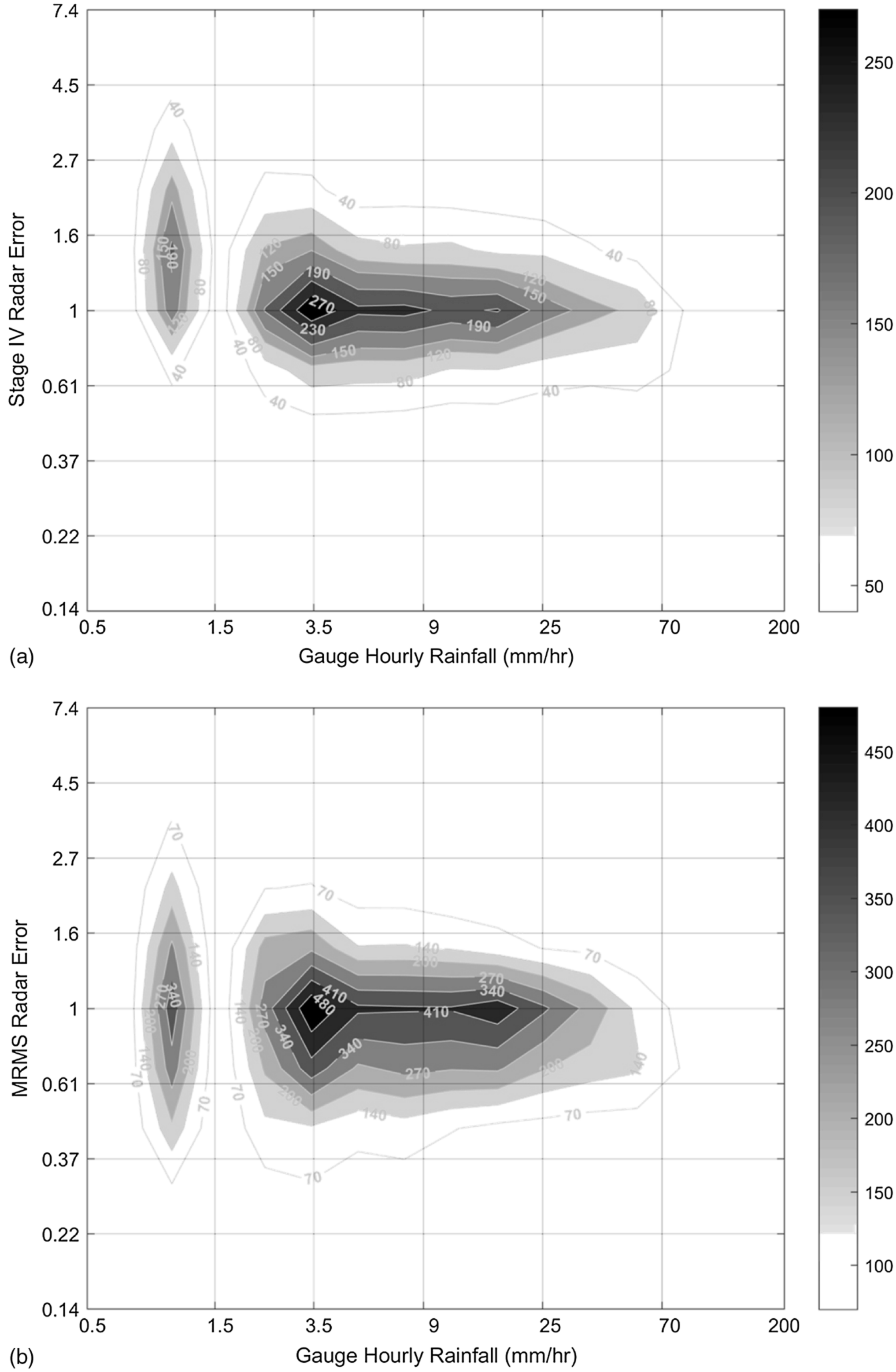

Fig. 8. 2D histogram of radar error from the three storm events combined for (a) Stage IV; and (b) MRMS.

assuming positive correlation at small time step (e.g., Ko et al. 2018). In spite of the differences among the previous studies, most of them lacked sufficient observations to firmly support any conclusion or assumption on the temporal structure of radar errors. More research is thereby needed to understand the spatiotemporal correlations of radar errors for various types of rainfall and radar data (Peleg et al. 2013). Herein, the authors intend to emphasize that sample size is vital for estimating the correlation coefficients and thus should be maximized (Kessler and Neas 1994; Habib et al. 2001a). Therefore, to augment and verify the findings from this study, the authors will apply the new sampling approach to long-term radar and rain gauge data in a future study.

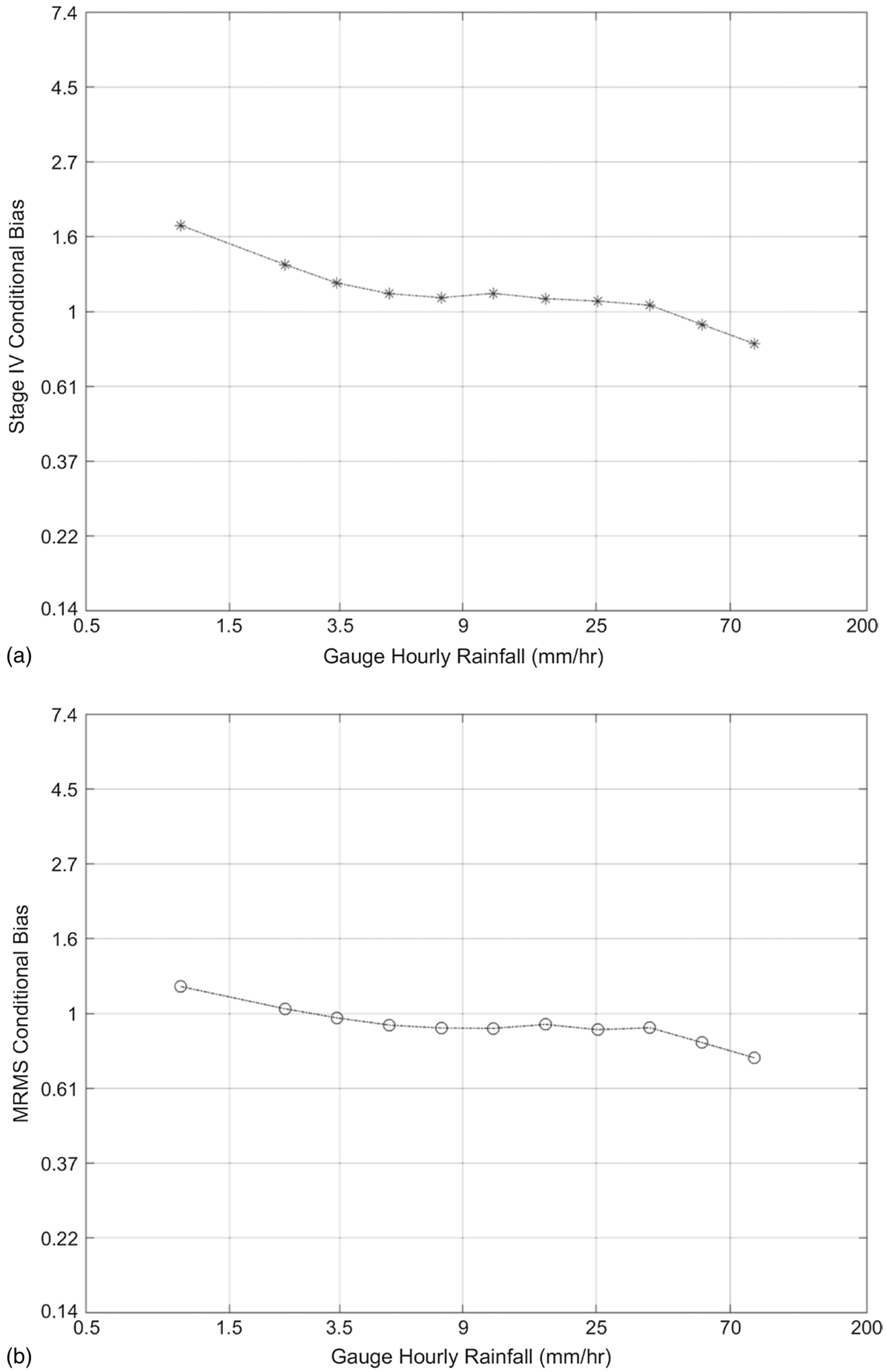

Fig. 9. CB from the three storm events combined for (a) Stage IV; and (b) MRMS.

Hydrologic Simulation

Because of the model calibration effort and the minor role of infiltration, the runoff simulation error is assumed to be mainly caused by the MAP estimation error instead of other modeling and parameter uncertainties. The calibrated HEC-HMS model was used to simulate runoff during Hurricane Harvey with three types of rainfall inputs, i.e., Stage IV, MRMS, and rain gauge. Comparisons are made between the simulated streamflow and the observed at four junctions along Bray Bayou (Junction 1/USGS8074760@Belle Park Dr.; Junction 2/USGS8074810@S. Gessner Rd.; Junction 3/USGS8075000@Main St.; and Junction 4/USGS8075110@MLK Blvd.) (see Fig. 3). Differences between simulated and observed hydrographs are summarized statistically (Table 3) using Eqs. (6)–(9) and presented visually in Fig. 12. Based on RMSE, NSE, and hydrograph shape, simulations driven

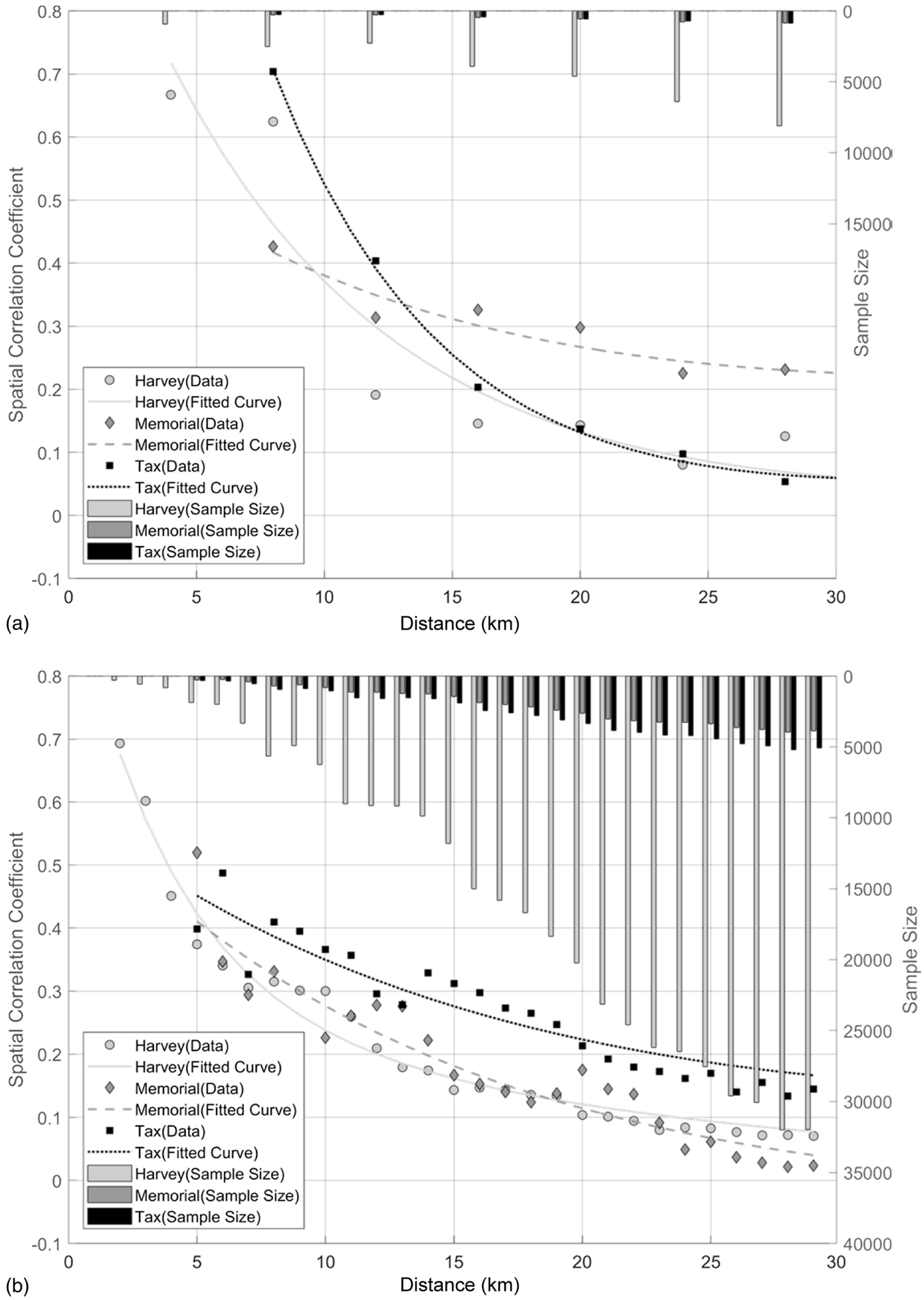

Fig. 10. Spatial autocorrelation coefficients of radar errors from the three storm events separately for (a) Stage IV; and (b) MRMS.

by MRMS and rain gauge equally generate better overall matches with the observations than those from Stage IV at all junctions. The P_e and V_e values indicate that peak flow and runoff volume are overestimated in the Stage IV hydrographs at most junctions, while the MRMS and rain gauge data generate better P_e and V_e with small overestimations or underestimations. The systematic overestimation by Stage IV-driven stimulation is expected as it aligns with the rainfall amount comparison.

To further dissect MAP estimation, it is decomposed into rainfall error and spatial resolution. Table 4 summarizes these two factors in the three rainfall inputs: (1) $\overline{OB}$ being the averaged OB ($\overline{OB}$) of the rainfall measurements enclosed by the Bray Bayou

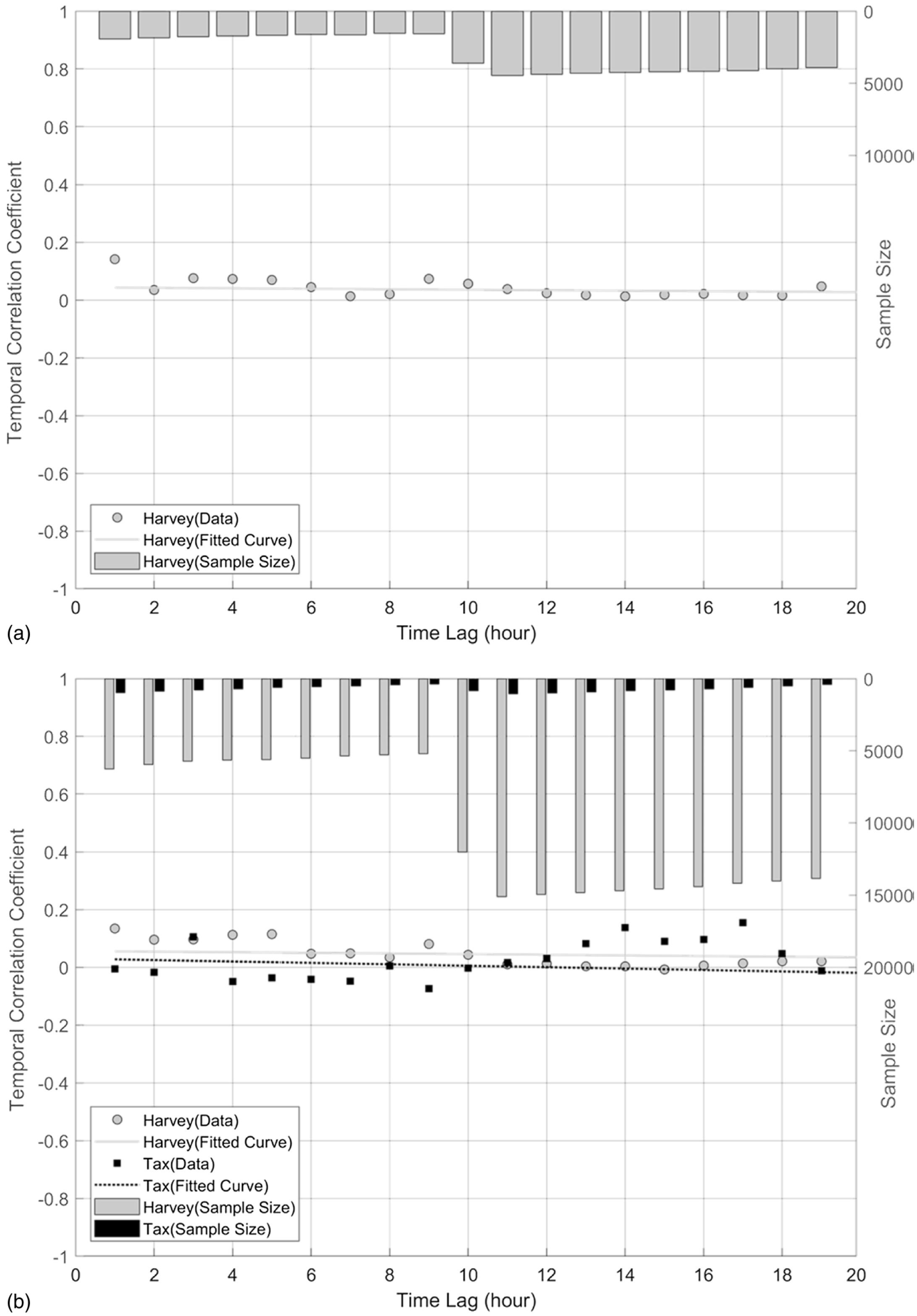

Fig. 11. Temporal autocorrelation coefficients of radar errors from the three storm events separately for (a) Stage IV; and (b) MRMS.

boundary; and (2) spatial resolution being the area of radar pixel in the cases of QPEs and the average area of Theissen polygons in the case of rain gauge. According to $\overline{OB}$ values, Stage IV and MRMS overestimate ($\overline{OB} = 1.11$) and underestimate ($\overline{OB} = 0.94$) the rainfall in Brays Bayou, respectively, which corresponds to their V_e values of 28% and −5% at Junction 4 (near the watershed outlet). Furthermore, it is worthwhile noticing the implication of spatial resolution on the accuracy of MAP estimation. Of all three rainfall inputs, MRMS has a superior spatial resolution, while Stage IV and rain gauge are coarser. When combining $\overline{OB}$ and spatial resolution, one can find that Stage IV produces lesser MAP estimation, as it has the largest $\overline{OB}$ and coarsest spatial resolution. The difference in spatial resolution also explains why the hydrologic simulations driven by rain gauge and MRMS perform

Table 3. Summary of hydrograph comparison

Statistic measures	Junction No.	Stage IV	MRMS	Rain gauge
V_e (%)	1	16	−18	−10
	2	38	0	9
	3	23	−8	0
	4	28	−5	3
P_e (%)	1	30	−5	3
	2	74	40	53
	3	2	−8	−7
	4	21	10	11
$RMSE$ (m^3/s)	1	33	19	17
	2	141	61	82
	3	164	71	91
	4	220	90	82
NSE	1	0.68	0.89	0.91
	2	0.15	0.84	0.71
	3	0.78	0.96	0.93
	4	0.63	0.94	0.95

Table 4. Averaged OB and spatial resolution of rainfall inputs

Rainfall input	$\overline{OB}$	Spatial resolution (km^2)
Stage IV	1.11	~16
MRMS	0.94	~1
Rain gauge	1	13

similarly despite that the rain gauge measurement is unbiased ($\overline{OB} = 1$).

Using one watershed (Brays Bayou) in Harris County, the authors cannot simply determine the better QPE product (Stage IV or MRMS) for flow simulations during Hurricane Harvey but would like to emphasize the significance of spatial resolution of QPE in MAP estimation. When rainfall estimates are unbiased, the uncertainty in MAP estimation can be analytically decomposed into two components: (1) the fractional coverage of rainfall over catchments, and (2) the spatial variability of rainfall itself, or inner variability (Entekhabi and Eagleson 1989; Barancourt et al. 1992;

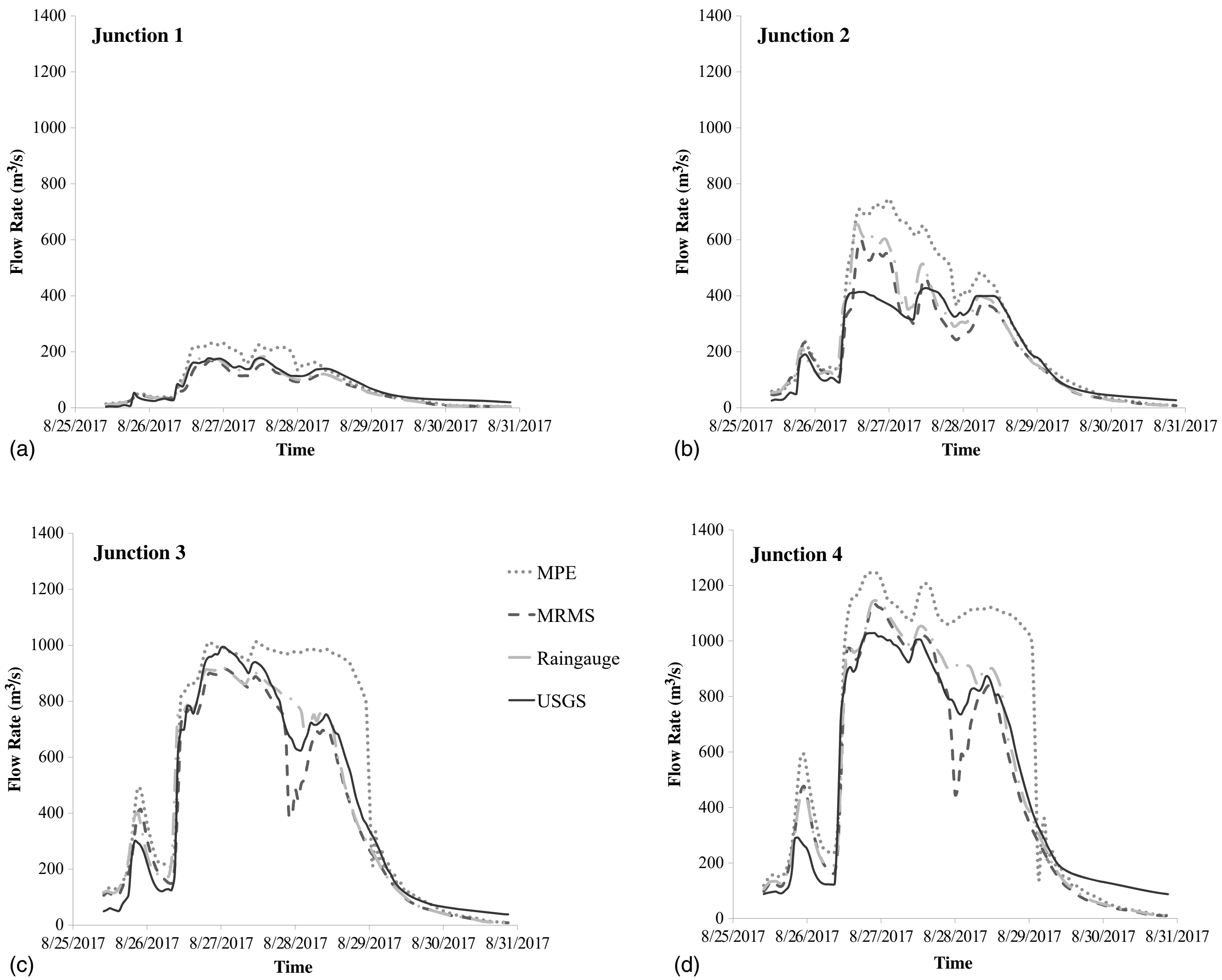

Fig. 12. Simulated and observed hydrographs for Hurricane Harvey in Brays Bayou at (a) Junction 1/USGS8074760@Belle Park Dr.; (b) Junction 2/USGS8074810@S. Gessner Rd.; (c) Junction 3/USGS8075000@Main St.; and (d) Junction 4/USGS8075110@MLK Blvd.

Seo and Smith 1996; Zhang and Seo 2017). Uncertainty in estimating the first component depends on the spatial resolution of QPE in a way that QPE with a higher resolution better represents the fractional coverage of rainfall over catchments.

Conclusions and Future Work

The authors investigate the performances of two hourly radar QPEs, the NEXRAD Stage IV and the MRMS Q3gc products, because of their important roles as precipitation input in major operational river forecasting activities. A new sampling approach for spatial reference rainfall is introduced in this study, which features resolving spatial variability of one QPE at the subpixel level by using another QPE with finer spatial resolution. Due to the vast spatial coverage and long duration, Hurricane Harvey (2017) provides a unique opportunity to demonstrate this new methodology. In comparison to the other two flood-inducing storm events (2015 Memorial Day storm and 2016 Tax Day storm) occurring in Harris County, Texas, Hurricane Harvey shows not only its exceptional rainfall magnitude but also the importance of sample size in studying the spatiotemporal characteristics of radar error. Thanks to the expixelent spatial scale and density of the HCFCD rain gauge network, the authors manage to effectively collect sufficient spatial reference rainfall samples with the new approach and then evaluate the radar errors in terms of bias, conditional dependence on rainfall intensities, and spatiotemporal structure. Several major findings from this study are summarized as follows:

1. Serving as truth in QPE evaluation, sufficient spatial reference rainfall samples are vital for truthfully revealing the performance of QPE as well as various aspects of radar error. The collection of spatial reference rainfall should be based on spatial rainfall variability at a subpixel level.
2. The Stage IV and MRMS QPEs perform fairly well during the three investigated storms: with Stage IV overestimating and MRMS underestimating the hourly rainfall by 2% and 12%, respectively. Both QPEs tend to overestimate very light rainfall.
3. Spatial correlation of radar errors from both QPEs can be described as powered exponential functions of interpixel distance. No significant temporal correlation of radar errors is found in this study for either QPE at any temporal lags.
4. Spatial resolution of QPE determines the estimation of MAP as the inputs to hydrologic simulations.

The insight gained from investigating radar error will enable us to further improve QPE performance in two ways: (1) to improve the rainfall estimation algorithm accounting for CBs; and (2) to model radar error as a spatially and temporally correlated random process. In addition, the sampling approach of spatial reference rainfall is not limited to Stage IV and MRMS as long as two utilized QPE products have different spatial resolutions and overlapping spatial and temporal extents. Therefore, this approach can be applied more broadly. For instance, provided available rain gauge records, 20 years of the multisensor precipitation estimates (MPE, 4×4 km^2) covering the CONUS can be evaluated by utilizing the corresponding reflectivity product (1×1 km^2) converted to rainfall intensity using the radar reflectivity–rainfall rate (Z–R) relationship. Such research will be presented in a forthcoming paper.

Data Availability Statement

The data, models or code that support the findings of this study are available from the corresponding author upon reasonable request. The available data for requesting are listed as follows:

1. QPE radar rainfall data for Hurricane Harvey.
2. MRMS rainfall data for Hurricane Harvey.
3. Rain gauge data for Hurricane Harvey from HCFCD.
4. HEC-HMS model for Brays Bayou.
5. Observed stream data at USGS8074760, USGS8074810, USGS8075000, and USGS807511.

Acknowledgments

This study was supported by funding from the National Science Foundation (#1832065) and the U.S. Army Corps of Engineers (W9126G-17-2-0042). The authors would also like to thank the West Gulf River Forecast Center (WGRFC) for providing radar rainfall data, and the Harris County Flood Control District (HCFCD) for providing the rain gauge data.

References

Barancourt, C., J. D. Creutin, and J. Rivoirard. 1992. "A method for delineating and estimating rainfall fields." *Water Resour. Res.* 28 (4): 1133–1144. https://doi.org/10.1029/91WR02896.

Bass, B., A. Juan, G. Avantika, Z. Fang, and P. B. Bedient. 2016. "2015 memorial day flood impacts for changing watershed conditions in Houston, TX." *J. Nat. Hazards Rev.* 18 (3): 05016007. https://doi.org/10.1061/(ASCE)NH.1527-6996.0000241.

Bedient, P. B., B. C. Hoblit, D. C. Gladwell, and B. E. Vieux. 2000. "NEXRAD radar for flood prediction in Houston." *J. Hydrol. Eng.* 5 (3): 269–277. https://doi.org/10.1061/(ASCE)1084-0699(2000)5:3(269).

Bedient, P. B., A. Holder, J. A. Benavides, and B. E. Vieux. 2003. "Radar-based flood warning system applied to Tropical Storm Allison." *J. Hydrol. Eng.* 8 (6): 308–318. https://doi.org/10.1061/(ASCE)1084-0699(2003)8:6(308).

Bedient, P. B., A. Holder, J. F. Thompson, and Z. Fang. 2007. "Modeling of stormwater response under large tailwater conditions—Case study for the Texas medical center." *J. Hydrol. Eng.* 12 (3): 256–266. https://doi.org/10.1061/(ASCE)1084-0699(2007)12:3(256).

Blake, E. S., and D. A. Zelinsky. 2018. "Hurricane Harvey—National hurricane center—NOAA." Accessed June 1, 2018. https://www.nhc.noaa.gov/data/tcr/AL092017_Harvey.pdf.

Box, G. E. P., and G. M. Jenkins. 1976. *Time series analysis: Forecasting and control*. Revised ed. San Francisco: Holden-Day.

Brassel, K. E., and D. Reif. 1979. "A procedure to generate Thiessen polygons." *Geog. Anal.* 11 (3): 289–303. https://doi.org/10.1111/j.1538-4632.1979.tb00695.x.

Catizone, P. A., S. E. Zell, C. R. Arrington, M. B. Newman, S. F. Weber, and R. J. White. 2014. "Comparative statistical study of hourly precipitation determined by radar-based stage IV and ground-based methods in the North Central United States." *J. Air Waste Manage. Assoc.* 64 (3): 291–308. https://doi.org/10.1080/10962247.2013.872209.

Ciach, G. J., E. Habib, and W. F. Krajewski. 2003. "Zero-covariance hypothesis in the error variance separation method of radar rainfall verification." *Adv. Water Resour.* 26 (5): 573–580. https://doi.org/10.1016/S0309-1708(02)00163-X.

Ciach, G. J., and W. F. Krajewski. 1999. "Radar–rain gauge comparisons under observational uncertainties." *J. Appl. Meteorol.* 38 (10): 1519–1525. https://doi.org/10.1175/1520-0450(1999)038<1519:RRGCUO>2.0.CO;2.

Ciach, G. J., W. F. Krajewski, and G. Villarini. 2007. "Product-error-driven uncertainty model for probabilistic quantitative precipitation estimation with NEXRAD data." *J. Hydrometeorol.* 8 (6): 1325–1347. https://doi.org/10.1175/2007JHM814.1.

Ciach, G. J., M. L. Morrissey, and W. F. Krajewski. 2000. "Conditional bias in radar rainfall estimation." *J. Appl. Meteorol.* 39 (11): 1941–1946. https://doi.org/10.1175/1520-0450(2000)039<1941:CBIRRE>2.0.CO;2.

Cocks, S. B., J. Zhang, S. M. Martinaitis, Y. Qi, B. Kaney, and K. Howard. 2017. "MRMS QPE performance east of the Rockies during the 2014 warm season." *J. Hydrometeorol.* 18 (3): 761–775. https://doi.org/10.1175/JHM-D-16-0179.1.

Dai, Q., D. Han, M. Rico-Ramirez, and P. K. Srivastava. 2014. "Multivariate distributed ensemble generator: A new scheme for ensemble radar precipitation estimation over temperate maritime climate." *J. Hydrol.* 511 (Apr): 17–27. https://doi.org/10.1016/j.jhydrol.2014.01.016.

Entekhabi, D., and P. S. Eagleson. 1989. "Land surface hydrology parameterization for atmospheric general circulation models including subgrid scale spatial variability." *J. Clim.* 2 (8): 816–831. https://doi.org/10.1175/1520-0442(1989)002<0816:LSHPFA>2.0.CO;2.

Fang, Z., P. B. Bedient, J. A. Benavides, and A. L. Zimmer. 2008. "Enhanced radar-based flood alert system and floodplain map library." *J. Hydrol. Eng.* 13 (10): 926–938. https://doi.org/10.1061/(ASCE)1084-0699(2008)13:10(926).

Fang, Z., P. B. Bedient, and B. Buzcu-Guven. 2011. "Long-term performance of a flood alert system and upgrade to FAS3: A Houston Texas case study." *J. Hydrol. Eng.* 16 (10): 818–828. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000374.

Fang, Z., G. Dolan, A. Sebastian, and P. B. Bedient. 2014. "Case study: Flood mitigation and hazard management at the Texas medical center in the wake of tropical storm Allison (2001)." *Nat. Hazards Rev.* 15 (3): 05014001. https://doi.org/10.1061/(ASCE)NH.1527-6996.0000139.

FEMA. 2017. "Historic disaster response to Hurricane Harvey in Texas." Accessed October 10, 2017. https://www.fema.gov/news-release/2017/09/22/historic-disaster-response-hurricane-harvey-texas.

Gao, S., and Z. Fang. 2018. "Using storm transposition to investigate the relationships between hydrologic responses and spatial moments of catchment rainfall." *Nat. Hazards Rev.* 19 (4): 04018015. https://doi.org/10.1061/(ASCE)NH.1527-6996.0000304.

Gourley, J. J., and B. E. Vieux. 2005. "Evaluating the accuracy of quantitative precipitation estimates from a hydrologic modeling perspective." *J. Hydrometeorol.* 6 (2): 115–133. https://doi.org/10.1175/JHM408.1.

Habib, E., A. Aduvala, and E. A. Meselhe. 2008. "Analysis of radar-rainfall error characteristics and implications for streamflow simulations uncertainty." *Hydrol. Sci. J.* 53 (3): 568–587. https://doi.org/10.1623/hysj.53.3.568.

Habib, E., 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 (10): 967–980. https://doi.org/10.1016/j.advwatres.2004.08.003.

Habib, E., and W. F. Krajewski. 2002. "Uncertainty analysis of the TRMM ground-validation radar-rainfall products: Application to the TEFLUN-B field campaign." *J. Appl. Meteorol.* 41 (5): 558–572. https://doi.org/10.1175/1520-0450(2002)041<0558:UAOTTG>2.0.CO;2.

Habib, E., W. F. Krajewski, and G. J. Ciach. 2001a. "Estimation of rainfall interstation correlation." *J. Hydrometeorol.* 2 (6): 621–629. https://doi.org/10.1175/1525-7541(2001)002<0621:EORIC>2.0.CO;2.

Habib, E., W. F. Krajewski, and A. Kruger. 2001b. "Sampling errors of tipping-bucket rain gauge measurements." *J. Hydrol. Eng.* 6 (2): 159–166. https://doi.org/10.1061/(ASCE)1084-0699(2001)6:2(159).

Habib, E., B. F. Larson, and J. Graschel. 2009. "Validation of NEXRAD multisensor precipitation estimates using an experimental dense rain gauge network in South Louisiana." *J. Hydrol.* 373 (3): 463–478. https://doi.org/10.1016/j.jhydrol.2009.05.010.

HADS (Hydrometeorological Automated Data System). 2017a. "Data acquisition and distribution system." Accessed November 15, 2017. https://hads.ncep.noaa.gov/.

HCFCD (Harris County Flood Control District). 2017b. "Projects and studies, Brays Bayou overview." Accessed June 15, 2017. https://www.hcfcd.org/projects-studies/brays-bayou/.

HCFWS (Harris County Flood Warning System). 2017c. "Flood warning and information system." Accessed September 1, 2017. https://www.harriscountyfws.org.

Juan, A., Z. Fang, and P. B. Bedient. 2015. "Developing a radar-based flood alert system for Sugar Land, Texas." *J. Hydrol. Eng.* 22 (5): E5015001. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001194,E5015001.

Kessler, E., and B. Neas. 1994. "On correlation, with applications to the radar and raingage measurement of rainfall." *Atmos. Res.* 34 (1–4): 217–229. https://doi.org/10.1016/0169-8095(94)90093-0.

Kitchen, M., and R. M. Blackall. 1992. "Representativeness errors in comparisons between radar and gauge measurements of rainfall." *J. Hydrol.* 134 (1–4): 13–33. https://doi.org/10.1016/0022-1694(92)90026-R.

Ko, D., T. Lee, and D. Lee. 2018. "Spatio-temporal-dependent errors of radar rainfall estimates in flood forecasting for the Nam River dam Basin." *Meteorol. Appl.* 25 (2): 322–336. https://doi.org/10.1002/met.1700.

Krajewski, W. F., and J. A. Smith. 2002. "Radar hydrology: Rainfall estimation." *Adv. Water Resour.* 25 (8–12): 1387–1394. https://doi.org/10.1016/S0309-1708(02)00062-3.

Lin, Y., and K. E. Mitchell. 2005. "1.2 the NCEP stage II/IV hourly precipitation analyses: Development and applications." In *Proc., 19th Conf. Hydrology, American Meteorological Society*. Boston: American Meteorological Society.

Mandapaka, P. V., G. Villarini, B. C. Seo, and W. F. Krajewski. 2010. "Effect of radar-rainfall uncertainties on thespatial characterization of rainfall events." *J. Geophys. Res.* 115 (D17110). https://doi.org/10.1029/2009JD013366.

Moran, P. A. 1950. "Notes on continuous stochastic phenomena." *Biometrika* 37 (1–2): 17–23. https://doi.org/10.1093/biomet/37.1-2.17.

NCEI (National Center for Environmental Information). 2018. "NEXRAD." Accessed February 18, 2018. https://www.ncdc.noaa.gov/data-access/radar-data/nexrad.

Nelson, B., O. Prat, D. Seo, and E. Habib. 2016. "Assessment and implications of NCEP stage IV quantitative precipitation estimates for product comparisons." *Weather Forecasting* 31 (2): 371–394. https://doi.org/10.1175/WAF-D-14-00112.1.

NOAA (National Oceanic and Atmospheric Administration). 2016. "National water model: Improving NOAA's water prediction services." Accessed March 16, 2018. http://water.noaa.gov/documents/wrn-national-water-model.pdf.

Peleg, N., M. Ben-Asher, and E. Morin. 2013. "Radar subpixel-scale rainfall variability and uncertainty: Lessons learned from observations of a dense rain-gauge network." *Hydrol. Earth Syst. Sci.* 17 (6): 2195.

Seo, D. J., and J. A. Smith. 1996. "Characterization of the climatological variability of mean areal rainfall through fractional coverage." *Water Resour. Res.* 32 (7): 2087–2095. https://doi.org/10.1029/96WR00486.

Smith, J. A., M. L. Baeck, J. E. Morrison, and P. Sturdevant-Rees. 2002. "The regional hydrology of extreme floods in an urbanizing drainage basin." *J. Hydrometeorol.* 3 (3): 267–282. https://doi.org/10.1175/1525-7541(2002)003<0267:TRHOEF>2.0.CO;2.

Smith, J. A., M. L. Baeck, Y. Zhang, and C. A. Doswell III. 2001. "Extreme rainfall and flooding from supercell thunderstorms." *J. Hydrometeorol.* 2 (5): 469–489. https://doi.org/10.1175/1525-7541(2001)002<0469:ERAFFS>2.0.CO;2.

Torres, J., B. Bass, N. Irza, Z. Fang, J. Proft, C. Dawson, M. Kiani, and P. B. Bedient. 2015. "Characterizing the hydraulic interactions of hurricane storm surge and rainfall–runoff for the Houston–Galveston region." *J. Coastal Eng.* 106 (Dec): 7–19. https://doi.org/10.1016/j.coastaleng.2015.09.004.

Vieux, B. E., and P. B. Bedient. 1998. "Estimation of rainfall for flood prediction from WSR-88D reflectivity: A case study." *Weather Forecasting* 13 (2): 407–415. https://doi.org/10.1175/1520-0434(1998)013<0407:EORFFP>2.0.CO;2.

Wang, X., H. Xie, H. Sharif, and J. Zeitler. 2008. "Validating NEXRAD MPE and stage III precipitation products for uniform rainfall on the upper Guadalupe River Basin of the Texas Hill country." *J. Hydrol.* 348 (1–2): 73–86. https://doi.org/10.1016/j.jhydrol.2007.09.057.

Yilmaz, K., T. Hogue, K. L. Hsu, S. Sorooshian, H. Gupta, and T. Wagener. 2005. "Intercomparison of rain gauge, radar, and satellite-based precipitation estimates with emphasis on hydrologic forecasting." *J. Hydrometeorol.* 6 (4): 497–517. https://doi.org/10.1175/JHM431.1.

Zhang, J., et al. 2014. "Initial operating capabilities of quantitative precipitation estimation in the Multi-Radar Multi-Sensory system." In *Proc., 28th Conf. of Hydrology*. Boston: American Meteorological Society. https://ams.confex.com/ams/94Annual/webprogram/Paper240487.html.

Zhang, J., et al. 2016. "Multi-radar multi-sensor (MRMS) quantitative precipitation estimation: Initial operating capabilities." *Bull. Am. Meteorol. Soc.* 97 (4): 621–638. https://doi.org/10.1175/BAMS-D-14-00174.1.

Zhang, Y., and D. J. Seo. 2017. "Recursive estimators of mean-areal and local bias in precipitation products that account for conditional bias." *Adv. Water Resour.* 101 (Mar): 49–59. https://doi.org/10.1016/j.advwatres.2017.01.002.

Zhang, Y., and J. A. Smith. 2003. "Space-time variability of rainfall and extreme flood response in the Menomonee River Basin, Wisconsin." *J. Hydrometeorol.* 4 (3): 506–517. https://doi.org/10.1175/1525-7541(2003)004<0506:SVORAE>2.0.CO;2.