# RAINFALL SIMULATION FOR INVESTIGATNG HYDROLOGIC RESPONSES TO STORM CHARACTERISTICS AND FACILITATING FREQUENCY ANALYSIS

by

Shang Gao

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Arlington

in Partial Fulfillment of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT ARLINGTON

August 2019

# Abstract

RAINFALL SIMULATION FOR INVESTIGATNG HYDROLOGIC RESPONSES TO STORM CHARACTERISTICS AND FACILITATING FREQUENCY ANALYSIS

Shang Gao

The University of Texas at Arlington, 2019

Hydrologic and Hydraulic practices depend on the characterization of extreme precipitation that is rare in nature. Despite the advancement in sensing technology, the observation record of extreme rainfall with good spatio-temporal resolution has not reached a sufficient length. Furthermore, rainfall observation may never be sufficient in a non-stationary climate. Towards solutions to these issues, rainfall simulation was explored in this study (1) to investigate hydrologic responses to storm characteristics and (2) to facilitate rainfall frequency analysis. To that end, this study initially employed storm transposition plus hydrologic simulation for generating possible rainfall-runoff scenarios which further reveal relationships between storm characteristics and hydrologic responses from an urban watershed, Brays Bayou in Houston. Consequently, this study generated meaningful findings highlighting the role of Brays Bayou's flat terrain in affecting flood peaks. Meanwhile, limitations in storm transposition were exposed and led to the development of a Dynamic Moving Storm (DMS) generator in the second part of this study. The DMS generator consists of three fundamental components, storm movement, spatial and temporal variabilities of rainfall intensity. By parameterizing each component, a vast variety of rainfall scenarios could be generated, simulated, and then analyzed again using Brays Bayou as the testbed. Hydrologic responses were revealed as results of not only individual parameters but also parameter interaction. The findings confirmed and

augmented those based on the storm transposition approach. In accord with previous studies, storms moving downstream at the same velocity as streamflow were found to cause the resonance effect on flood peaks. Brays Bayou's flat terrain was the key in explaining the patterns in various flood peaks and storm parameters. As the last part of this study, a stochastic storm generator was developed featuring optimal estimation for spatio-temporal modeling of rain fields and a non-parametric approach for generating model parameters. A case study in Dallas-Fort Worth Metroplex was conducted to generate possible realizations of 100 most rainy days at high spatio-temporal resolution. As the first step towards applying a long-term stochastic rainfall simulation, this approach was proved to offer many attractive traits, as statistical properties of observed rainfall were well preserved in simulations requiring only a small number of parameters. Collectively, this study offers innovative pathways for better utilizing rainfall information in flood modeling and rainfall frequency analysis with reduced uncertainties from the limited records of extreme storm events that make them feasible alternatives to traditional approaches.

# Acknowledgements

First and foremost, I would like to acknowledge Dr. Nick Z. Fang as my advisor who has always been supportive for my research ideas even when prospects of success seemed uncertain. More than just an advisor, Nick accepted me into the Fang Research Group as his first Ph.D. student and showed me what it takes to be a trailblazer. In the past five years, I have witnessed the growth of Fang Research Group alongside with each group member as my lifelong colleagues. Nick has exposed me to a vast variety of tasks, meetings, and conferences, which totally defied my stereotype about Ph.D. work. Together with Nick, I learnt that this pathway to a Ph.D. is like a stream finding its way to the ocean. The path can twist and turn; sometimes seen by people but sometimes hidden underground. Nonetheless, the gravitational pull from the true knowledge always leads you to the destination.

I want to sincerely thank my committee members, Professor Dong-jun Seo, Habib Ahmari, and Philip Bedient. Their constructive suggestions throughout my Ph.D. cleared my confusions and guided onto the right track. Most importantly, I learnt one thing from them – knowledge and experience always triumph over smart thoughts. I also want to thank every group member in the Fang Research Group, who I have been so fortunate to collaborate with. My Ph.D. wouldn't have been complete without their passion and intelligence.

To my family, thank you for encouraging me in this pursuit of Ph.D. I am especially grateful to my parents, who supported me emotionally and financially to follow my dreams studying in the U.S. Even from halfway across the world, I always knew that you believed in me and wanted the best for me. Thank you for the freedom to indulge in my interests throughout my childhood. Thank you for teaching me the values in receiving the best education to one's life. This last word

of acknowledgement I have saved is for my wife, Jiaqi Zhang. This journey would have been impossible without your support. I see this Ph.D. as a testimony to how great a team we make. In the end, our marriage grew deeper and has made this Ph.D. the best years of my life.

July 01, 2019

To my parents

For my care-free childhood full of colorful memories

To my cousin Bing Liu

For the brotherly love we shared as two only children

I cherish them forever…

To my wife and sweetest friend, Jiaqi

For life's perfect arrangement putting you in this journey

These are the best days of my life…

*Thank you.*

# Table of Contents

# Chapter 1: Introduction

## 1.1 DESCRIPTION OF PROBLEM

Rainfall plays an essential part in the hydrologic cycle and is the driving phenomenon of runoff mechanisms. Many hydrologic/hydraulic (H/H) applications are dependent on accurate information of rainfall at high spatio-temporal resolution with a long record-length (Cunnane, 1988; Faurès et al., 1995). However, rain gage networks as a major data source for decades of H/H practices inherently lack the necessary density to capture spatial variability (Michaud and Sorooshian, 1994). It was not until the applications of weather radar systems that a detailed understanding of rainfall with high spatial resolution was acquired via radar rainfall, also known as quantitative precipitation estimate (QPE). Along with the advent of QPE, spatio-temporal structure of storms has drawn considerable interest from the research community but still requires deeper understanding, especially on the corresponding hydrologic effects.

Another challenge in applying radar rainfall in hydrologic/hydraulic designs comes from its relative short record-length. As the centerpiece of the wide range of H/H applications, rainfall frequency analysis (RFA) reveals the relationship between rainfall magnitude and likelihood. Any RFA becomes more valid with longer record of rainfall observation, since the return period of infrequent severe storms are better captured. In the absence of a long data record, extrapolation is inevitable which adds uncertainty to RFA.

## 1.2 RESEARCH GOALS & QUESTIONS

The overarching goals of this Ph.D. work are 1) *to improve our understanding of the roles of storm characteristics in flood generation* and 2) *to better incorporate spatio-temporal*

*variabilities of rainfall into design storms*. The availability of hydro-meteorological data and the frequency of flood hazards of two urban areas in Texas and, more specifically, the Dallas-Fort Worth metroplex and Houston led to the selection of these two regions as test beds for this research. The results of this research are timely, as flood hazards induced by heavy rainfall (e.g., 2015 Memorial Day storm, 2016 Tax Day storm, and 2017 Hurricane Harvey) hit these areas in recent years with increased intensity and frequency.

In view of the realistic concern for heavy cost of lives and properties, this study seeks to simulate rainfall through the use of advanced stochastic techniques as well as hydrologic modeling in order to answer the following research questions:

1. How do characteristics of storms (e.g. storm locations, storm movements, and spatio-temporal variabilities of rainfall intensity) affect flood peaks as individual factor or via interaction with other storm characteristics?
2. How can synthetic design storms better incorporate spatio-temporal variability of rainfall intensity and storm movements so that the corresponding hydrologic responses can be better captured?
3. How can storms be stochastically simulated at desired spatio-temporal resolution as a way to extend the limited length of QPE record?

In answering these questions, the research will be conducted to achieve the following objectives:

1. Investigate hydrologic responses to storm characteristics based on rainfall-runoff scenarios created through deterministic storm transposition.
2. Develop a deterministic storm generator to comprehensively study the hydrologic responses to storm characteristics.

3. Develop a stochastic storm generator that is capable of reproducing key statistics of observed rainfall in order to facilitate rainfall frequency analysis.

## 1.3 DISSERTATION STRUCTURE

This dissertation consists of five chapters. Chapter 1 provides a brief background of the research, identifies the study areas, presents the research questions and objectives, and outlines the structure of this dissertation. Subsequent chapters contain manuscripts focusing on the following topics: investigating hydrologic responses to catchment rainfall from transposed QPE (Chapter 2); developing a synthetic storm generator and analyzing hydrologic responses to storm characteristics (Chapter 3); and developing a stochastic storm generator (Chapter 4). Chapter 5 presents a conclusions of the whole dissertation.

Chapter 2 is exploratory in nature and seeks to link characteristics of storms to hydrologic responses via catchment rainfall. This study featured transposing radar data of historical storms in order to generate a large number of rainfall-runoff scenarios. The patterns exhibited in the hydrologic responses from the rainfall-runoff scenarios were interpreted in relation to spatial organization of catchment rainfall. The results highlighted the role of spatial concentration of catchment rainfall in affecting the correlation between mean areal precipitation (MAP) and peak discharge. Chapter 2 revealed limitations of the approach and led to the development of a deterministic storm generator in Chapter 3 as a better tool to investigate hydrologic responses. The hydrologic models validated in Chapter 2 were also used in Chapter 3. The hydrologic metrics successfully used in Chapter 3 for analyzing the results were also applied in Chapter 3.

Chapter 3 involves developing a deterministic storm generator – Dynamic Moving Storm (DMS) generator. Hydrologic responses to individual parameters in DMS generator as well as

their interactions were extensively investigated using global sensitivity analysis and pairwise sensitivity analysis. The results demonstrated the necessity in considering all three components in the DMS framework, i.e. 1) spatial variability of rainfall intensity, 2) temporal variability of rainfall intensity, and 3) storm movement, since interactions among parameters significantly affect hydrologic responses. The DMS generator served as a stepping stone for the stochastic storm generator in Chapter 5.

The stochastic storm generator in Chapter 5 was mainly based on the theory of optimal estimation. Compared to other existing methods, optimal estimation is attractive because the key rainfall statistics are prescribed as model parameters and are inherently preserved throughout the modeling framework. Innovative methods were also developed in Chapter 5 to ensemble model parameters by borrowing rainfall information from adjacent homogenous areas. All the manuscripts in this dissertation have been published or are intended for publication. Their full references are as follows:

Gao, S., and Fang, Z. (2018). “Using Storm Transposition to Investigate the Relationships between Hydrologic Responses and Spatial Moments of Catchment Rainfall”. ASCE Journal of Natural Hazards Review, ASCE 2018 19(4):04018015 DOI:10.1061/(ASCE)NH.1527-6996.0000304

Gao, S., & Fang, Z. N. Investigating Hydrologic Responses to Spatio‐temporal Characteristics of Storms Using a Dynamic Moving Storm (DMS) Generator. Hydrological Processes.

Gao, S., and Fang, Z. (2019). “Stochastic Storm Simulation Using Optimal Estimation and Non-Parametric Generation of Ensemble Parameters”. (In Preparation).

## 1.4 REFERENCES

Cunnane, C. (1988). Methods and merits of regional flood frequency analysis. Journal of Hydrology, 100(1-3), 269-290.

Faurès, J. M., Goodrich, D. C., Woolhiser, D. A., & Sorooshian, S. (1995). Impact of small-scale spatial rainfall variability on runoff modeling. Journal of hydrology, 173(1-4), 309-326.

Michaud, J. D., & Sorooshian, S. (1994). Effect of rainfall‐sampling errors on simulations of desert flash floods. Water Resources Research, 30(10), 2765-2775.

# Chapter 2:

# Using Storm Transposition to Investigate the Relationships between Hydrologic Responses and Spatial Moments of Catchment Rainfall[1]

Shang Gao and Zheng Fang

Gao, S., and Fang, Z. (2018). "Using Storm Transposition to Investigate the Relationships between Hydrologic Responses and Spatial Moments of Catchment Rainfall". ASCE Journal of Natural Hazards Review, ASCE 2018 19(4):04018015 DOI:10.1061/(ASCE)NH.1527-6996.0000304

[1] Used with Permission from ASCE, the publisher of Natural Hazards Review, 2019

## ABSTRACT

The dependence between spatiotemporally varied rainfall and watershed hydrology has long been recognized and studied using either synthetic or real storms/watersheds. To examine the relationship, the authors apply the storm transposition technique to reposition three historical storm events over the Brays Bayou watershed in Houston, Texas. Transposing various storms not only enables the researchers to pinpoint the worst-case hydrologic impact but also generate a large variety of rainfall-runoff scenarios that exhibit distinct patterns in peak discharge and runoff volume. Furthermore, the authors utilize spatial statistics (termed spatial moments) to characterize catchment rainfall and interpret the identified patterns in hydrologic responses. Multivariate linear regression is particularly employed in this study to quantitatively represent the relationship between peak discharge (dependent variable) and spatial moment of catchment rainfall (independent variables). During the study, storm dimension is found to be a primary variable to both spatial organizations of catchment rainfall and spatial patterns in peak discharge rate; small-sized storms tend to generate spatially-concentrated catchment rainfall with large displacements between the centroids of catchment rainfall and drainage network; flow distance from the storm core to the catchment outlet is inversely correlated to peak flow. The insights gained from this study will benefit future research in hydrologic analyses and prediction of rainfall-induced hazards, especially determining critical factors in the hydrologic responses during natural disasters like Hurricane Harvey (2017).

## KEY TERMS

Storm Transposition; Spatial Moments of Catchment Rainfall; Multivariate Linear Regression; Multi-Radar Multi-Sensor Rainfall

# INTRODUCTION

Intense storms with high spatiotemporal variability are often the main causes for serious flooding occurring in urbanized areas. Many previous studies have been conducted focusing on different aspects of the rainfall characteristics like spatiotemporal variability and storm movement. Ball (1994) demonstrates the significance of varied temporal distribution of rainfall to peak discharge and hydrograph shape. Morin et al. (2006) found that spatial variability within thunderstorm rain cells have great influence on both rainfall volume and peak discharge of a semi-arid catchment. Many researchers also investigated storm movement, i.e. starting location, direction and velocity, as an important factor to the resulted hydrologic response using experimental (Yen and Chow, 1969), numerical (Surkan 1974, Stephenson 1984, Richardson and Julien 1989, Ogden et al., 1995, Lee and Huang 2007) and analytical methods (Singh 1998, de Lima and Singh 2002, Viglione et al., 2010a). The authors are thereby motivated to better understand the severity of flooding in terms of peak discharge and runoff volume with respect to storm centering locations in relative to a watershed, which involves repositioning or transposing spatial rainfall data of historical storms and then conducting hydrologic simulation with accuracy.

Storm transposition is a method to move an observed storm from its original location to other places where the storm could have possibly occurred. In previous research, the term “storm transposition” is often found associated with the word “stochastic” or “probabilistic” (Wang, 1987; Foufoula-Georgiou, 1989; Franchini et al., 1996). Stochastic storm transposition, or SST, has been primarily used to extend rainfall records for rainfall and flood frequency studies (e.g. Wright et al., 2013; Wright et al., 2014; England et al., 2014). The spatial limits for SST are prescribed by a “meteorologically homogenous region” (Wang, 1987) where a storm occurs following a uniform Poisson process (Foufoula-Georgiou, 1989). On the other hand, storm

transposition, when applied without the stochastic framework, can be referred to as deterministic storm transposition (DST). In the field of hydrologic design, such a deterministic approach is often used to validate spatial and temporal features of frequency storms. For instance, Benson (2014) utilized radar rainfall data for storm transposition and compared the hydrologic impacts of design storms and the transposed historical storms. Extreme storms were transposed deterministically to compensate the lack of database for possible maximum precipitation (PMP) in many previous studies (Rakhecha and Kennedy, 1985; Hansen, 1987; Noriah and Rakhecha, 2001). In this study, the deterministic storm transposition is mainly utilized to create an ensemble of scenarios which not only include the worst case in terms of hydrologic response but also yield meaningful patterns between rainfall and runoff for a highly urbanized watershed.

Catchment or watershed, as the receiving body of rainfall, plays a crucial role in rainfall-runoff process. Therefore, this study focuses on rainfall received by the catchment, or the "catchment rainfall". Previous researchers have developed measures not only for the characteristics of rainfall itself but also for rainfall features in relative to geometric properties of the catchment. For instance, Syed et al., (2003) utilized the Euclidean distance between watershed outlet and the storm core which was defined as the centroid of rainfall with intensity over 25 mm/hr (1 in/hr). Moreover, various measures of catchment rainfall were developed based on the distance to the watershed outlet along flow path, or flow distance. Smith et al., (2002, 2005) employed a series of scaled metrics called normalized flow distance and normalized dispersion to examine rainfall variability. Zoccatelli et al., (2011) derived a set of spatial statistics of catchment rainfall by rearranging the analytical solutions developed by Woods and Sivapalan (1999) and Viglione et al., (2010a, b). In the authors' opinion, flow distance is a significant parameter among watershed characteristics and should be incorporated

into catchment rainfall measurement because the river/channel network plays an important part in attenuating and translating the runoff through routing.

Fully investigating the rainfall characteristics requires data with high spatial and temporal resolution. However, for decades, hydrologic models utilized rain gage networks as a major data source (Michaud and Sorooshian, 1994), which inherently lacks the necessary density to capture spatial variability. It was not until the applications of weather radar systems that a detailed understanding of rainfall with high spatial resolution was acquired. Numerous recent studies took advantage of this radar rainfall technology and generated meaningful outcomes about spatial and temporal structures of rainfall (e.g. Feral et al., 2003; von Hardenberg 2003; Wealands et al., 2005; Paschalis et al., 2013; Peleg and Morin, 2014). In this study, Multi-Radar Multi-Sensor (MRMS) rainfall data are utilized. The MRMS quantitative precipitation system currently integrates about 180 operational radars and creates a seamless 3D radar mosaic across the conterminous United States (CONUS) and southern Canada at very high spatial (1 km) and temporal (2 min) resolution (Cocks et al., 2016). The radar-base data are integrated with atmospheric environmental data, satellite data, and lightning and rain gauge observations to generate a suite of severe weather and QPE products (Cocks et al., 2016). The authors utilize the MRMS gage-corrected quantitative precipitation estimate (QPE) product (ID: GaugeCorrQPE01H) with resolution of 1 km and 1 hour. This product is generated by first mosaicking reflectivity scans from multiple radars into the so-called seamless hybrid scan reflectivity (SHSR), then applying correction algorithm to mitigate estimation errors, and conducting bias-adjustment with rain gage data (Zhang et al., 2016). Evaluation works have been performed in recent years regarding various MRMS products in different regions and different seasons (Chen et al., 2016; Cocks et al., 2016; Martinaitis et al., 2016). In particular, Willie et al.,

(2016) compared different rainfall products for the coastal rainfall events and concluded that MRMS gage-corrected QPE product gives the best overall performance. From the hydrologic perspective, the authors validate the performance of MRMS gage-corrected QPE using three historical storm events that occurred over Houston in 2015 and 2016.

Houston, Texas has undergone a rapid urban expansion and is prone to intense tropical storms with a history of major flood events. One of the flood-prone urban watersheds, Brays Bayou (**Figure 1**) with 334 $km^2$ (129 $mi^2$) of area, is selected for this study. The watershed 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, 2017). Brays Bayou is prone to flooding due to its flat slopes, impermeable land surface and clay soils, and the explosive rainfall that characterizes the region (Bedient et al., 2003). Throughout the decades, the Brays Bayou watershed has a history of flood mitigation projects. In the late 1960s, a major concrete channelization project was completed to accommodate storms at the 100-year level. This project included 40.9 km (25.4 mi.) of channel improvements along Brays Bayou that involved widening and deepening of the channel. In addition, 22 km (13.7 mi.) of the channel was lined with concrete from 9.5 km (5.9 mi.) upstream of Brays Bayou outlet to US-59. The project sought to provide enough capacity within the channel banks to handle 100-year flows based on full development of the watershed back then. However, by the year of 1983, it was found that a 10-year storm would bring the channel to its full capacity due to the rapid development throughout the watershed. Triggered by a September 1983 storm which flooded 1,000 homes, a regional flood control plan was developed and adopted in 1988 for the Brays Bayou watershed, including channel improvements, regional detention basins, and a diversion project. By the year of 1993 the watershed had reached over 90% development, leading to bankfull conditions at Main St. in

response to only a 5 to 10-year storm (Bedient and Huber, 2002). After the devastation of Tropical Storm Allison in 2001, major efforts from the City of Houston, the Federal Emergency Management Agency (FEMA), the Harris County Flood Control District (HCFCD), and the Texas Medical Center (TMC) were collectively invested to resolve the flooding issues (Fang et al., 2011; Bass et al., 2016). Given the flooding issue of Brays Bayou, this study is conducted in the hope of enhancing the decision making process for emerging intense storms.

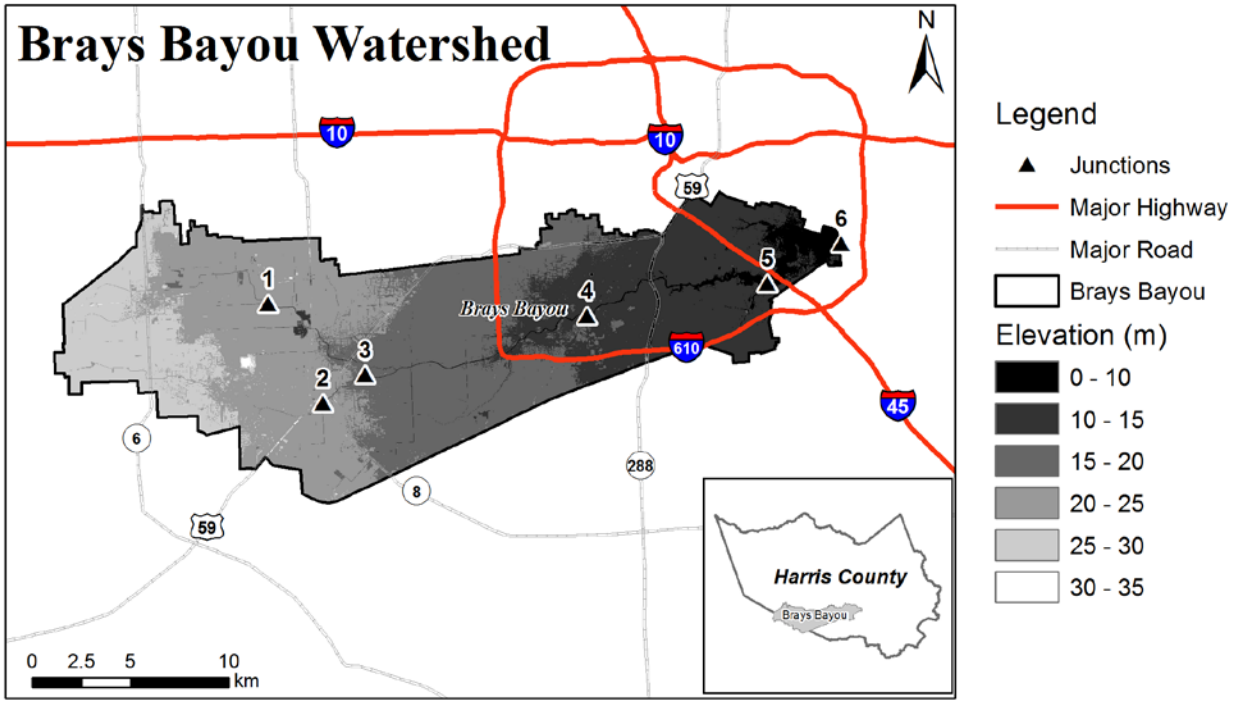



**Figure 1:** Study area − the Brays Bayou watershed, Houston, Texas.

In this analysis, three historical storms occurring in Brays Bayou during 2015 and 2016 were selected: the 2015 Memorial Day storm (May 25$^{th}$ of 2015), the 2015 October storm (October 30$^{th}$ of 2015) and the 2016 Tax Day storm (April 17$^{th}$ of 2016). **Table 1** summarizes the starting and ending times for the three historical storms. **Figures 2A**, **2B** and **2C** respectively show the cumulated rainfall for the three storms based on the MRMS data (1 km × 1 km, 1 hour).

It is found that, for the 2015 Memorial Day storm, the maximum cumulated rainfall (indicated by the highest contour) occurred near the midstream of Brays Bayou; while for the 2015 October storm and the 2016 Tax Day storm, the highest rainfall amount is found further away from the watershed. **Figure 3A** shows the intensity–duration information of the three historical storms derived from the MRMS data pixel (1 km × 1 km) with the maximum intensity, as well as the values of frequency storms (Intensity-Duration-Frequency curves). One can see that the 2015 October storm exceeds a 100-year storm, while the 2015 Memorial Day storm and the 2016 Tax Day storm are both close to a 50-year storm. On the other hand, **Figure 3B** shows the peak discharges values measured by 5 junctions (Junctions 1 to 5 in **Figure 1**) with USGS gages (**Junction 1**/USGS8074760@Belle Park Dr., **Junction 2**/USGS8074800@Keegans, **Junction 3**/USGS8074810@S. Gessner Rd., **Junction 4**/USGS8075000@Main St., and **Junction 5**/USGS8075110@MLK Blvd.) along Brays Bayou during the three historical storms, compared with the values from frequency storms. None of the three storms is found to generate greater peak discharges than a 50-year storm in Brays Bayou, and the 2015 October storm yielded even lower peak flow rates than a 10-year storm. In comparison of the observations in **Figures 3A** and **3B**, if the most intense portion of these three historical storms had been transposed to within the Brays Bayou, much greater peak discharge should have occurred in Brays Bayou,.

**Table 1:** Starting and Ending Times for the Three Selected Historical Storms

| Storm | Start Time (CDT) | End Time (CDT) | Duration (hours) |
|---|---|---|---|
| 2015 Memorial Day Storm | 5/25/2015 19:00 | 5/27/2015 7:00 | 36 |
| 2015 October Storm | 10/30/2015 19:00 | 10/31/2015 19:00 | 24 |
| 2016 Tax Day Storm | 4/17/2016 13:00 | 4/19/2016 6:00 | 41 |

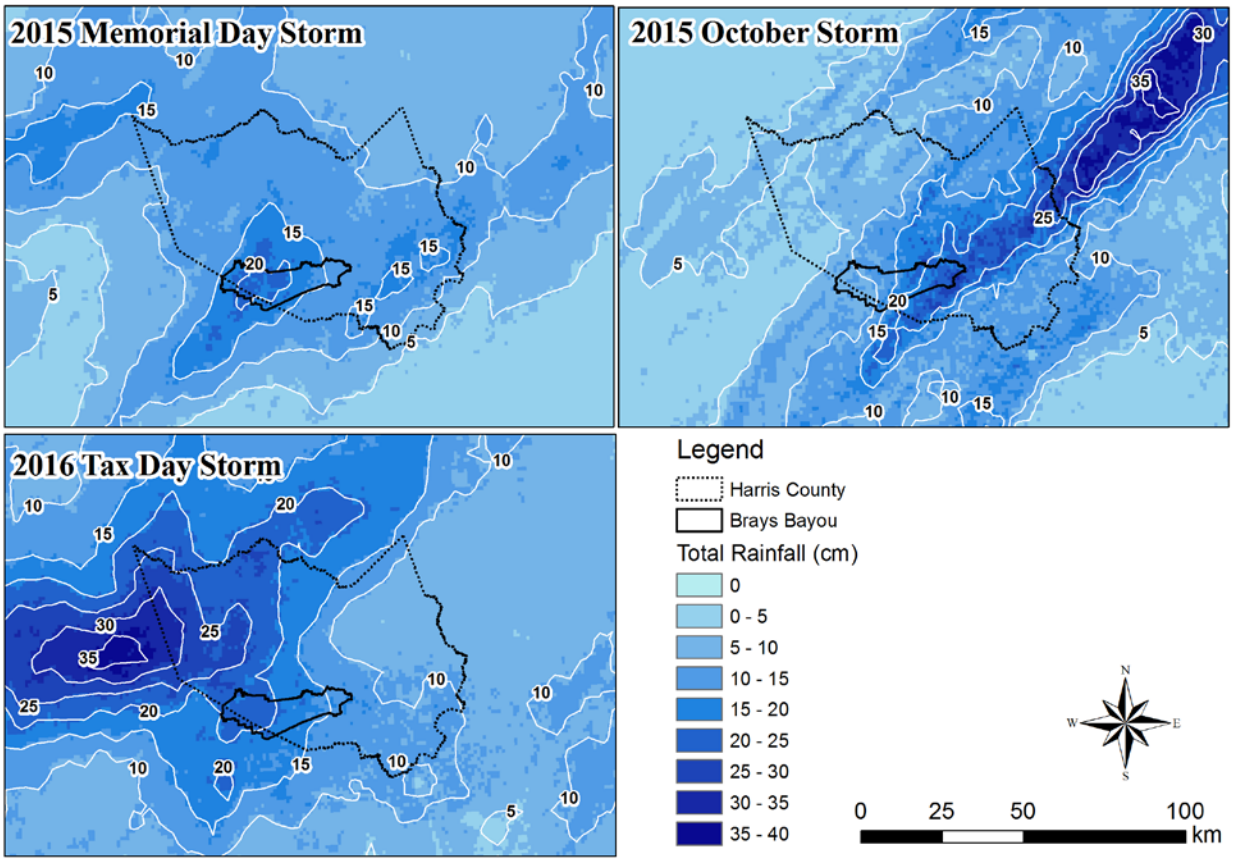



**Figure 2:** Cumulative rainfall based on the MRMS data of the three historical storms on (A) May 25th of 2015, (B) October 30th of 2015, and (C) April 17th of 2016, respectively.

In this regard, the authors are motivated to investigate the critical storm location and the consequent hydrologic impacts by transposing the MRMS rainfall data over Brays Bayou, as well as the patterns between rainfall and runoff in the considerable scenarios generated by storm transposition. By linking spatial statistics of catchment rainfall to the computed runoffs from hydrologic models, a better understanding of the dependency of hydrologic responses on spatial characteristics of rainfall can then be achieved via the following objectives:

1. To validate the hydrologic simulation in Brays Bayou using MRMS radar rainfall data along with USGS stream gage data from three historical storms;

2. To investigate the hydrologic responses from various storm centering locations of the three historical storms in Brays Bayou and pinpoint the worst-case scenarios;

3. To characterize catchment rainfall generated by storm transposition based on spatial statistics taking into account flow distance;

4. To interpret the patterns in hydrologic responses based on the spatial statistics from Objective 3.

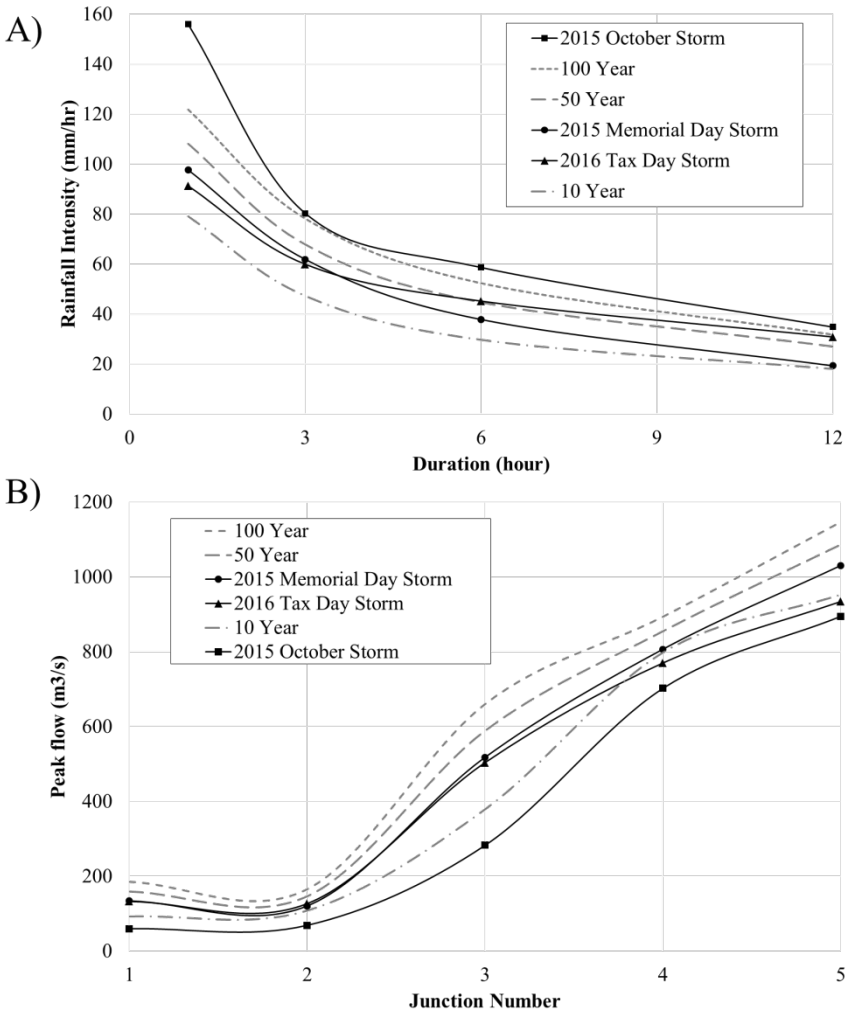



**Figure 3:** Comparison of the three historical storms with design events in terms of (A) intensity-duration information of rainfall and (B) peak discharge measured at 5 junctions.

# METHODOLOGY

The overarching methodology of this study consists of the following four parts, i.e. 1) calibration and validation of hydrologic model, 2) storm transposition, 3) spatial moments of catchment rainfall and flow distance, and 4) multivariate linear regression. Specifically, a developed hydrologic model is first calibrated and validated using MRMS radar rainfall data and USGS gages for three historical storms over Brays Bayou (2015 Memorial Day storm, 2015 October storm and 2016 Tax Day storm). An algorithm is then developed in this study to automatically transpose the three historical storms from their original locations to various spots within the watershed boundary and simulate hydrologic responses from the various storm centering locations. In addition, spatial statistics of both rainfall and drainage network (termed "spatial moments" of catchment rainfall and flow distance) are applied to interpret their relationships with flood response via multivariate linear regression. Details of these four components are introduced in the following sections.

### *Calibration and Validation of Hydrologic Model*

To best represent the watershed condition during 2015 to 2016, a hydrologic (HEC-HMS) model is calibrated using the MRMS rainfall data of three historical storms and five junctions with USGS stream gages Back in 2001, the devastating impact from Tropical Storm Allison (TSA) triggered the Tropical Storm Allison Recovery Project (TSARP), which led to the development of updated hydrologic (HEC-HMS) and hydraulic (HEC-RAS) models. The HMS model used in this study originated from the TSARP products and was modified by Bass et al., (2016) to best represent the watershed condition during the 2015 Memorial Day storm. Bass et al., (2016) reported the satisfactory match between the simulated and the observed hydrographs

at Junctions 2 to 5; and suggested the overestimated detention effect upstream of Junction 1 at Belle Park Drive. Therefore, the calibration effort in this study is first invested to improve the simulation performance at Junction 1. Additionally, the antecedent watershed condition is modified to fully saturated, meaning that infiltration process is nullified in the HEC-HMS simulations. **Figure 4** shows the comparison of simulated and observed hydrographs at Junction 1 to 5 (columns 1 to 5) for the three historical storms (three rows), and the overall close match indicates that the Brays Bayou HEC-HMS model is well calibrated and suitable to be used in further analyses.

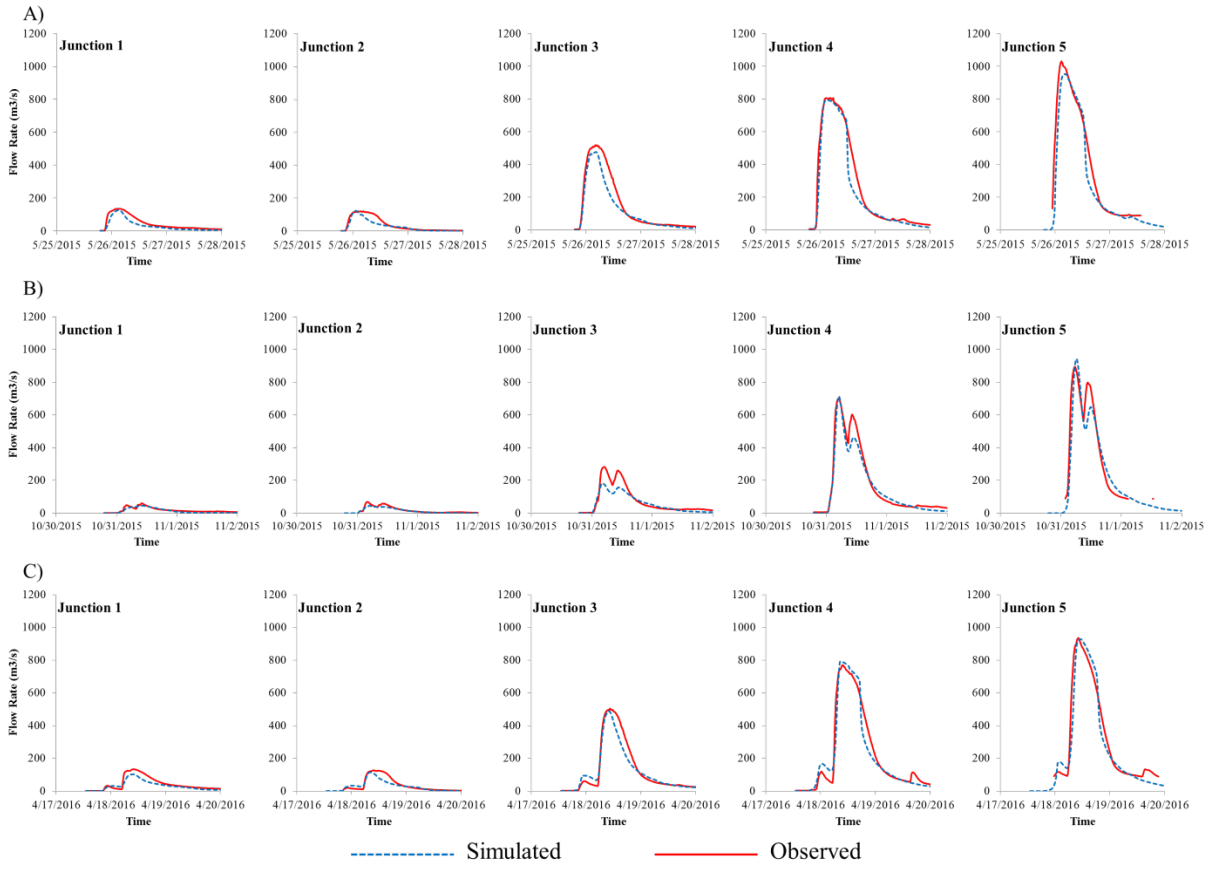



**Figure 4:** Comparison of simulated and observed hydrographs at Junction 1 to 5 for (A) the 2015 Memorial Day storm, (B) the 2015 October storm, and (C) the 2016 Tax Day storm.

***Storm Transposition***

As mentioned earlier, storm transposition is conducted using MRMS rainfall data. **Figure 5A** shows the storm transposition end points, which are the centroids of MRMS grids (1 km × 1 km) within the watershed boundary; whereas the transposition start points are the locations of maximum total rainfall for the three historical storms as shown in **Figure 5B**. Consequently, there are a total of 309 transposition scenarios corresponding to the 309 transposition end points for each storm event. In all transposition scenarios, the intense portion of the storm is repositioned into the watershed to identify the possible "worst-case" hydrologic impacts.

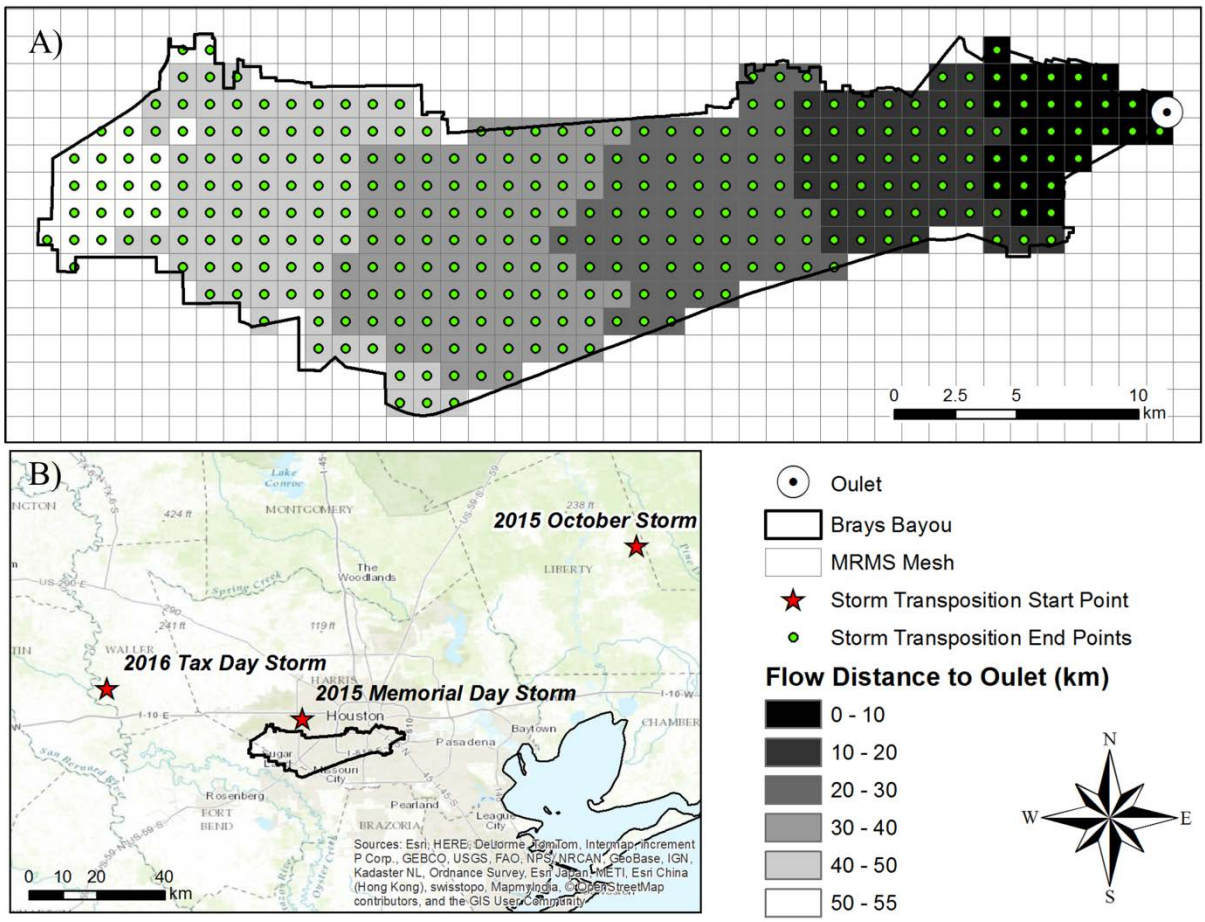



**Figure 5:** Storm transposition scheme with the transposition end points (A) and start points (B).

In order to make subsequent hydrologic simulations, hyetographs are generated through the schematic flow as shown in **Figure 6**. First, time series of MRMS radar data (original for the whole U.S.) are clipped to a smaller domain in order to expedite following processing steps. Second, the "Shift" tool is used to move the rainfall raster data along a defined vector ($\Delta x, \Delta y$).The shifted raster rainfall data with resolution of (1 km × 1 km) is then resampled to a finer resolution (250 m × 250 m), which increases the accuracy of basin-averaged rainfall calculated by the following "Zonal Statistics" tool. Based on the shapefile of the subbasins in Brays Bayou, the areal average rainfall values over each subbasin can be calculated via Zonal Statistics tool. The end products from ArcGIS processing are hyetographs for all subbasins. Additionally, a data conversion tool called DSS-UTL is applied (1) to import hyetographs into a HEC-HMS model and (2) to export hydrographs from desired junctions. The processes above are integrated and automated using Python scripting language for this study.

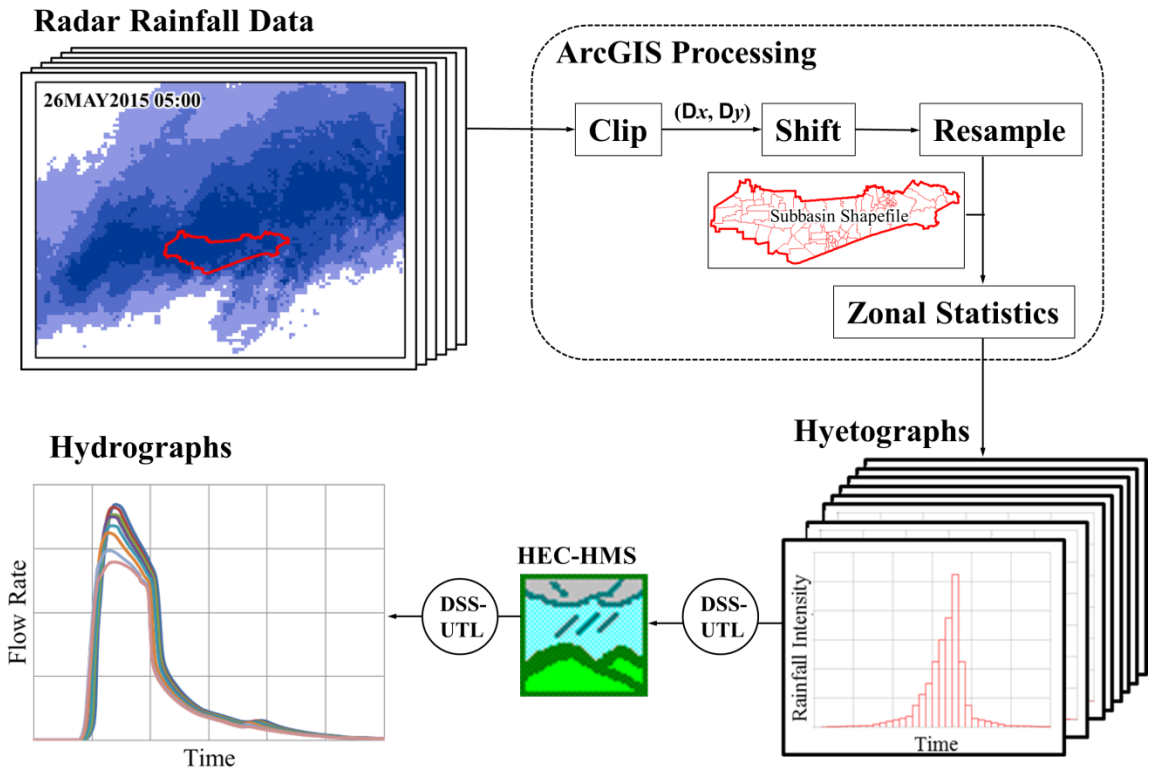


**Figure 6:** Schematic flow chart for generating hyetographs and conducting hydrologic simulations.

***Spatial Moments of Catchment Rainfall and Flow Distance***

The authors focus on the spatial statistics of cumulated catchment rainfall with respect to flow distance which are similar to the mathematical formulas characterizing instantaneous rainfall intensity in previous research (Zoccatelli et al., 2011). The *n*-th spatial moment of cumulated catchment rainfall $P_n$ is expressed in **Equation 1**:

$$P_n = |A|^{-1} \int_A r_t(x,y) d(x,y)^n \, dA \tag{1}$$

where $r_t(x,y)$ is the cumulated rainfall at location (*x, y*). The zero-th order of $P_n$ is equivalent of the cumulated mean areal precipitation (MAP); *d(x, y)* is the distance between position (*x, y*) and the catchment outlet measured along the flow path. **Figure 5A** shows the flow distance value from each pixel (1 km × 1 km) in the watershed boundary to the watershed outlet, which is generated using the "Flow Distance" tool in ArcGIS. The *n*-th moment of flow distance $g_n$ is based on **Equation 2**:

$$g_n = |A|^{-1} \int_A d(x,y)^n \, dA \tag{2}$$

where $r_t(x,y)$ is the mean value of cumulated rainfall at location (*x, y*). In order to represent the spatial location of rainfall in relative to the drainage network, dimensionless spatial moments of catchment rainfall ( $\Delta_n$ ) is obtained by taking the ratio between the spatial moments of cumulated rainfall and moments of the flow distance. The first- and second-order dimensionless moments, $\Delta_1$ and $\Delta_2$, are computed based on **Equations 3** and **4**:

$$\Delta_1 = \frac{P_1}{P_0 g_1} \tag{3}$$

$$\Delta_2 = \frac{1}{g_2 - g_1}\left[\frac{P_2}{P_0} - (\frac{P_1}{P_0})^2\right] \tag{4}$$

$\Delta_1$ describes the location of centroid of catchment rainfall with respect to the location of centroid of drainage network (i.e. the average value of flow distance). Values of $\Delta_1$ being one reflect a rainfall distribution concentrated at the watershed centroid, with values less than one indicating that rainfall is concentrated near the watershed outlet, and values greater than one indicating that rainfall is concentrated near the watershed headwaters (Zoccatelli et al., 2011). $\Delta_2$ depicts the dispersion of the location of catchment rainfall with respect to the variance of flow distances. Values of $\Delta_2$ being one reflect a uniform rainfall distribution, with values less than one indicating that rainfall is concentrated around a single core, and values greater than one indicating that rainfall is dispersed with multiple cores (Zoccatelli et al., 2011).

***Multivariate Linear Regression***

Based on the results from the storm transposition analysis, multivariate linear regression is further performed to interpret the importance of the spatial characteristics of rainfall (independent variables) in influencing hydrologic responses (dependent variables). To decipher the relatively importance of each regressor, the authors utilize the standardized beta coefficients which are associated with one standard deviation of change in the independent and dependent variables. For instance, a bivariate linear regression using standardized beta coefficients can be expressed using **Equation 5**:

$$\hat{y} = \beta_1 \hat{x}_1 + \beta_2 \hat{x}_2 \tag{5}$$

where $\hat{y}$ is the standardized dependent variable; $\hat{x}_1$ and $\hat{x}_2$ are the standardized independent variables with $\beta_1$ and $\beta_2$ being their corresponding standardized beta coefficients. And $\beta_1$ and $\beta_2$ can be calculated using the simple correlation coefficients between dependent and independent variables as shown by **Equations 6** and **7**:

$$\beta_1 = \frac{r_{y1} - r_{12} r_{y2}}{1 - {r_{12}}^2} \tag{6}$$

$$\beta_2 = \frac{r_{y2} - r_{12} r_{y1}}{1 - {r_{12}}^2} \tag{7}$$

where $r_{y1}$, $r_{y2}$ and $r_{12}$ are respectively the simple correlation coefficients of $\hat{y}$ and $\hat{x}_1$, $\hat{y}$ and $\hat{x}_2$, and $\hat{x}_1$ and $\hat{x}_2$.

## RESULTS AND ANALYSES

### *Storm Transposition*

In the storm transposition analysis, hydrologic responses are first examined for two Junctions (Junction 3 and Junction 6) with the three historical storms, forming a total of six simulation scenarios. Each scenario includes 309 times of deterministic storm transposition which serve as data points for subsequent multivariate regression analyses. The two junctions (3 and 6) are particularly selected to demonstrate how catchment size causes variation in hydrologic responses at the midstream and downstream, respectively. **Figures 7, 8,** and **9** show the resulted peak flow (upper two panels) and runoff volume (lower two panels) at the Junction 3 (S. Gessner Rd.) and Junction 6 (watershed outlet) for the 2015 Memorial Day storm, the 2015 October storm, and the 2016 Tax Day storm, respectively. The dimensionless values shown in these figures are the ratios between the values from the transposed storms and that from the original

storm. The contours with the highest values in these maps indicate the magnitudes of the worst-case flood response with the associated storm centering locations. **Table 2** summarizes the original peak discharge and runoff volume values along with the ranges of dimensionless values in the six simulation scenarios. It is found that the transposed 2015 October storm can generate the worst flood impact with the maximum peak flow and runoff volume being 366% and 393% of the original values, respectively. It should be noted that for the 2015 Memorial Day event, transposed storms do not show increased flood impact as much as in the cases of the other two storms, which means the original centering location of the 2015 Memorial Day storm is already close to the worst-case scenario.

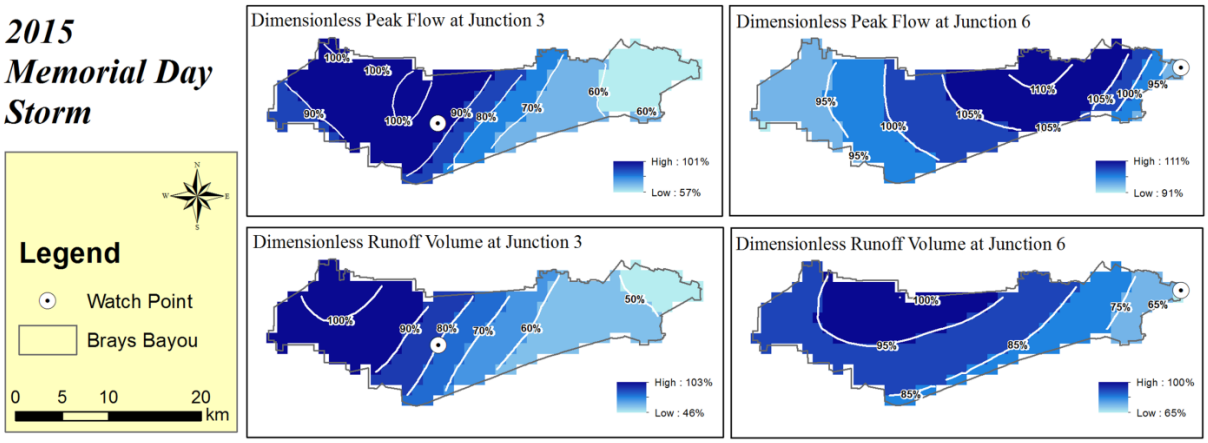



**Figure 7:** Peak flow and runoff volume values at the Junction 3 (S. Gessner Rd.) and Junction 6 (watershed outlet) for the 2015 Memorial Day storm.

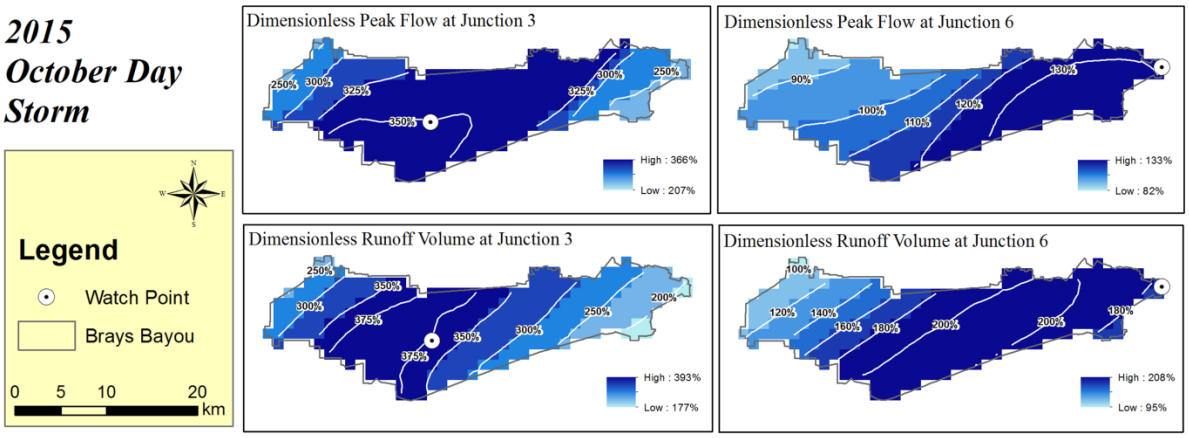



**Figure 8:** Peak flow and runoff volume values at the Junction 3 (S. Gessner Rd.) and Junction 6 (watershed outlet) for the 2015 October storm.

**Table 2:** Original Peak Discharge and Runoff Volume Values and the Ranges of Dimensionless Values for 6 Simulation Scenarios

| ***2015 Memorial Day Storm*** | **Scenario S1** | | **Scenario S2** | |
|---|---|---|---|---|
| | $Q_{peak}$ @ Junc 3 | $V_{runoff}$ @ Junc 3 | $Q_{peak}$ @ Junc 6 | $V_{runoff}$ @ Junc 6 |
| Original Value | 477 $m^3/s$ | 571 $m^3$ | 1,047 $m^3/s$ | 1,299 $m^3$ |
| Min Dimensionless Values | 57% | 46% | 91% | 65% |
| Max Dimensionless Values | 101% | 103% | 111% | 100% |
| ***2015 October Storm*** | **Scenario S3** | | **Scenario S4** | |
| | $Q_{peak}$ @ Junc 3 | $V_{runoff}$ @ Junc 3 | $Q_{peak}$ @ Junc 6 | $V_{runoff}$ @ Junc 6 |
| Original Value | 184 $m^3/s$ | 268 $m^3$ | 1,072 $m^3/s$ | 1,099 $m^3$ |
| Min Dimensionless Values | 207% | 177% | 82% | 95% |
| Max Dimensionless Values | 366% | 393% | 133% | 208% |
| ***2016 Tax Day Storm*** | **Scenario S5** | | **Scenario S6** | |
| | $Q_{peak}$ @ Junc 3 | $V_{runoff}$ @ Junc 3 | $Q_{peak}$ @ Junc 6 | $V_{runoff}$ @ Junc 6 |
| Original Value | 488 $m^3/s$ | 699 $m^3$ | 1,026 $m^3/s$ | 1,327 $m^3$ |

| | | | | |
|---|---|---|---|---|
| Min Dimensionless Values | 99% | 122% | 118% | 150% |
| Max Dimensionless Values | 132% | 164% | 133% | 192% |

In addition to the ranges of the dimensionless values, the general spatial patterns in the hydrologic responses are found different for the three storms as shown in **Figures 7, 8,** and **9**. For the two storms in 2015, the simulated peak discharges at both junctions (3 and 6) appear to increase as the storm is centered closer to the receiving junctions, while the opposite pattern can be observed in the case of the 2016 Tax Day storm where higher peak flow values are generated by storms centered at headwaters. A clearer insight can be obtained by comparing and correlating the patterns of peak flow and runoff volume (upper and lower panels in **Figures 7**, **8** and **9**). For the 2016 Tax Day storm, very similar patterns in the upper and lower panels can be found with the correlation coefficients $\rho$ being 0.96 and 0.95 respectively; while for Scenario S2 and S4 (2015 Memorial Day storm @ Junction 6 and the 2015 October storm @ Junction 6), there is less similarity ( $\rho$ = -0.11 for S2 and 0.8 for S4) between the two panels. The authors found that while the positive correlation exists between peak flow and runoff volume as shown in the case of the 2016 Tax Day storm, it can be affected by the storm centering location in relative to the channel network. As a matter of fact, a storm centered near the receiving junction can result in higher peak flow rate but lower runoff volume than that of the same storm if centered at headwaters. For this very reason, for the two storms in 2015, the peak discharge at both junctions (3 and 6) exhibits different behaviors than the runoff volume with the highest values generated from storm cores if centered closer to the receiving junctions. It should be also noted that the Brays Bayou watershed is fully saturated in the HEC-HMS model, meaning that runoff volume is equivalent of the cumulated mean areal precipitation (MAP). Therefore, based

on the storm transposition results, it is found that the correlation between cumulative MAP and peak flow is affected by the storm centering location in relative to the channel network.

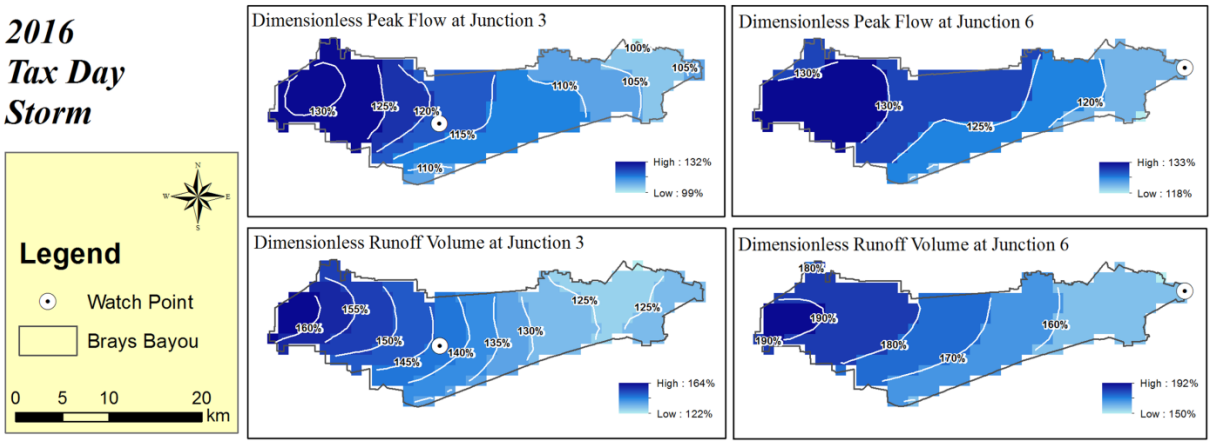



**Figure 9:** Peak flow and runoff volume values at the Junction 3 (S. Gessner Rd.) and Junction 6 (watershed outlet) for the 2016 Tax Day storm.

### *Spatial Characteristics of Rainfall*

In order to better explain how the three storms yield different hydrologic responses in Brays Bayou, their spatial characteristics are further investigated. **Figures 10A** and **10B** show the areal reduction factors (ARF) of the 12-hour maximum rainfall of the three investigated storms for area ranging from 0 to 10,000 km$^2$ and from 0 to 500 km$^2$, and the vertical line marks the dimension of Brays Bayou, i.e. 334 km$^2$. The ARF of the 2016 Tax Day storm is found to have a milder decaying slope than those of the two storms in 2015 within both 500 km$^2$ and 10,000 km$^2$ scales. In other words, the 2016 Tax Day storm is generally larger in dimension than those two storms in 2015. Also the 2015 Memorial Day storm and the 2015 October storm have a similar dimension within the scale of Brays Bayou (the vertical line).

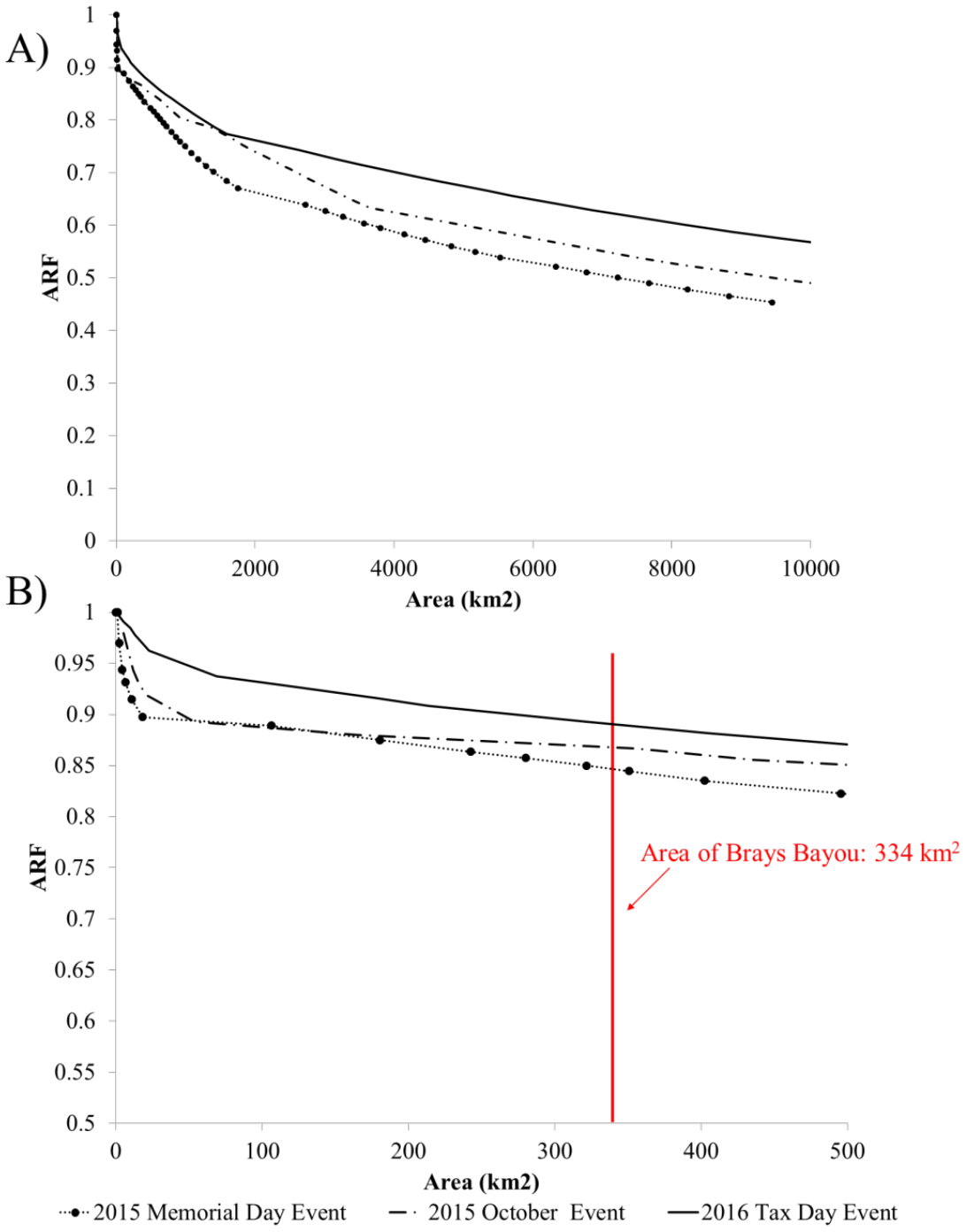



**Figure 10:** Areal reduction factors (ARFs) of the 12-hour maximum rainfall of the three investigated storms for area ranging (A) from 0 to 10,000 km$^2$ and (B) from 0 to 500 km$^2$

Moreover, insight is obtained by examining the spatial moments of catchment rainfall. **Figure 11** shows are a panel of histograms for the first- (1st row) and second-order (2nd row) dimensionless spatial moments of catchment rainfall, $\Delta_1$ and $\Delta_2$, for the 2015 Memorial Day storm (1st column), the 2015 October storm (2nd column) and the 2016 Tax Day storm (3rd column). As introduced earlier in this paper, $\Delta_1$ measures the location of the storm core in relative to the drainage network, and $\Delta_2$ measures the dispersion of catchment rainfall. It can be found that the two storms in 2015 exhibit a greater spread in $\Delta_1$ (**Figures 11A** and **11B**) than the 2016 Tax Day storm yielding concentrated $\Delta_1$ values (**Figure 11C**). This is due to the difference in their dimensions: rainfall from large-sized storms has little spatial variation within the watershed scale, thus centroids of the catchment rainfall in all (309) storm transposition scenarios differ very little from each other. For the same reason, catchment rainfall from a small-sized storm is spatially more concentrated than that from a large-sized storm. As shown in **Figures 11D** and **11E**, the majority of $\Delta_2$ values from the two storms in 2015 are smaller than one (concentrated), which is opposite to **Figure 11F** where considerable scenarios generate $\Delta_2$ values near or greater than one (dispersed). In summary, the 2015 Memorial Day and the 2015 October storms are smaller in dimension, but can cause more spatially concentrated catchment rainfall than that of the 2016 Tax Day storm.

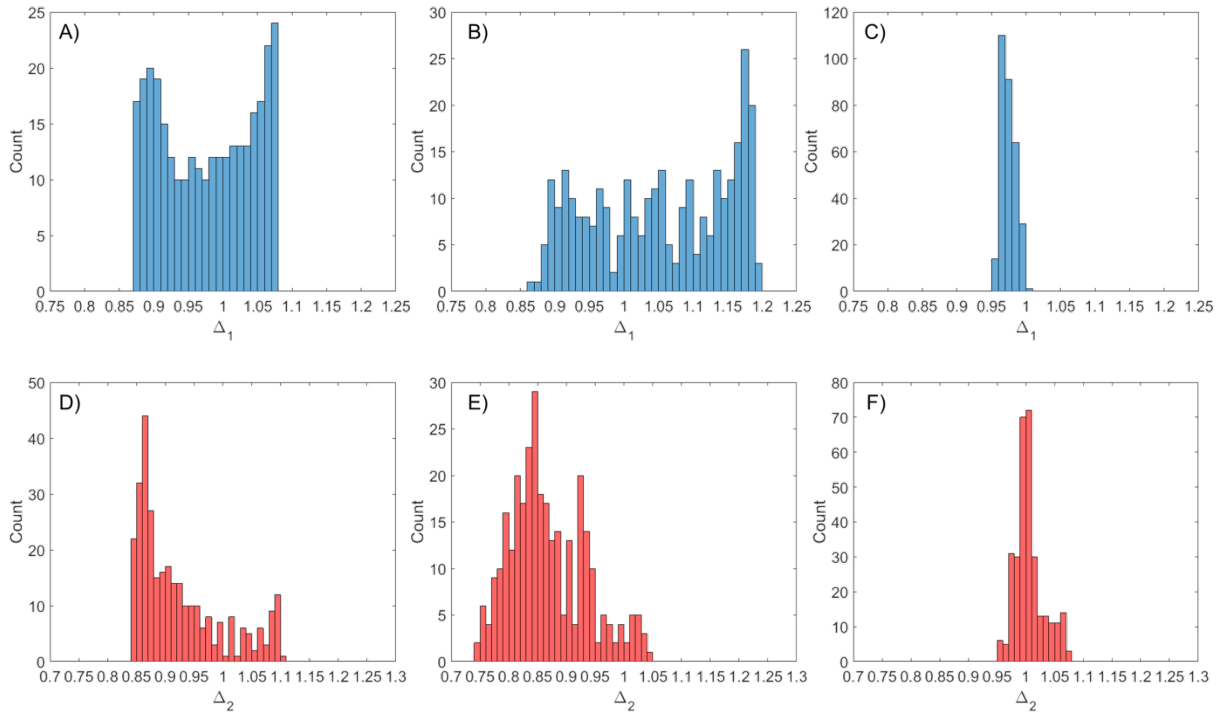



**Figure 11:** Histograms for the first- and second-order dimensionless spatial moments of catchment rainfall, $\Delta_1$ and $\Delta_2$, for the 2015 Memorial Day storm (A and D), the 2015 October storm (B and E) and the 2016 Tax Day storm (C and F).

***Multivariate Linear Regression Analysis***

From a quantitative perspective, the authors further utilize multivariate linear regression to interpret the importance of spatial variability of catchment rainfall in affecting the peak discharge rate at the catchment outlet. Peak discharge rates received at the junctions of interest (3 and 6) are employed as dependent variables, while the cumulated MAP ($P_0$) and the first-order dimensionless spatial moment of catchment rainfall $\Delta_1$ are independent variables. The independent and dependent variables are standardized and the multivariate linear regression simply follows **Equation 8** as below:

$$\hat{Q}_{peak} = \beta_1 \hat{P}_0 + \beta_2 \hat{\Delta}_1 \tag{8}$$

where $\hat{Q}_{peak}$ is the standardized peak flow rate; $\beta_1$ and $\beta_2$ are the standardized beta coefficients associated with standardized $P_0$ and $\Delta_1$ respectively. Six sets of bivariate linear regression analyses are conducted with respect to the six storm transposition scenarios. The sample size of each analysis is 309.

**Figure 12** shows a panel of 3-D plots of the data (scattered dots) and the resulted bivariate linear models (meshed plain) for the two junctions (columns in the panel) and the three storms (rows in the panel). The mathematical equation of the model is presented in each panel along with the corresponding coefficient of determination ($R^2$). As indicated by the $R^2$ values, all the bivariate linear regression analyses result in good fitness. In comparison, **Table 3** shows the $R^2$ values from simple linear regression using only $P_0$ as the independent variable along with those from the bivariate linear regression. The bivariate linear model is found to improve $R^2$ for the scenarios **S1**, **S2**, **S3**, and **S4**, especially for **S2** (the 2015 Memorial Day storm at Junction 6) seeing that $R^2$ increase from 0.01 to 0.92. However, in the case of the 2016 Tax Day storm (S5 and S6), the improvement in $R^2$ value is not explicit because the simple linear regression already generates very good fitness. Such findings indicate a strong correlation between cumulated MAP and peak discharge rate for the 2016 Tax Day storm. In addition, for the two smaller-sized storms in 2015, location of the storm core, represented by $\Delta_1$, is a critical variable to interpret the patterns in peak discharge.

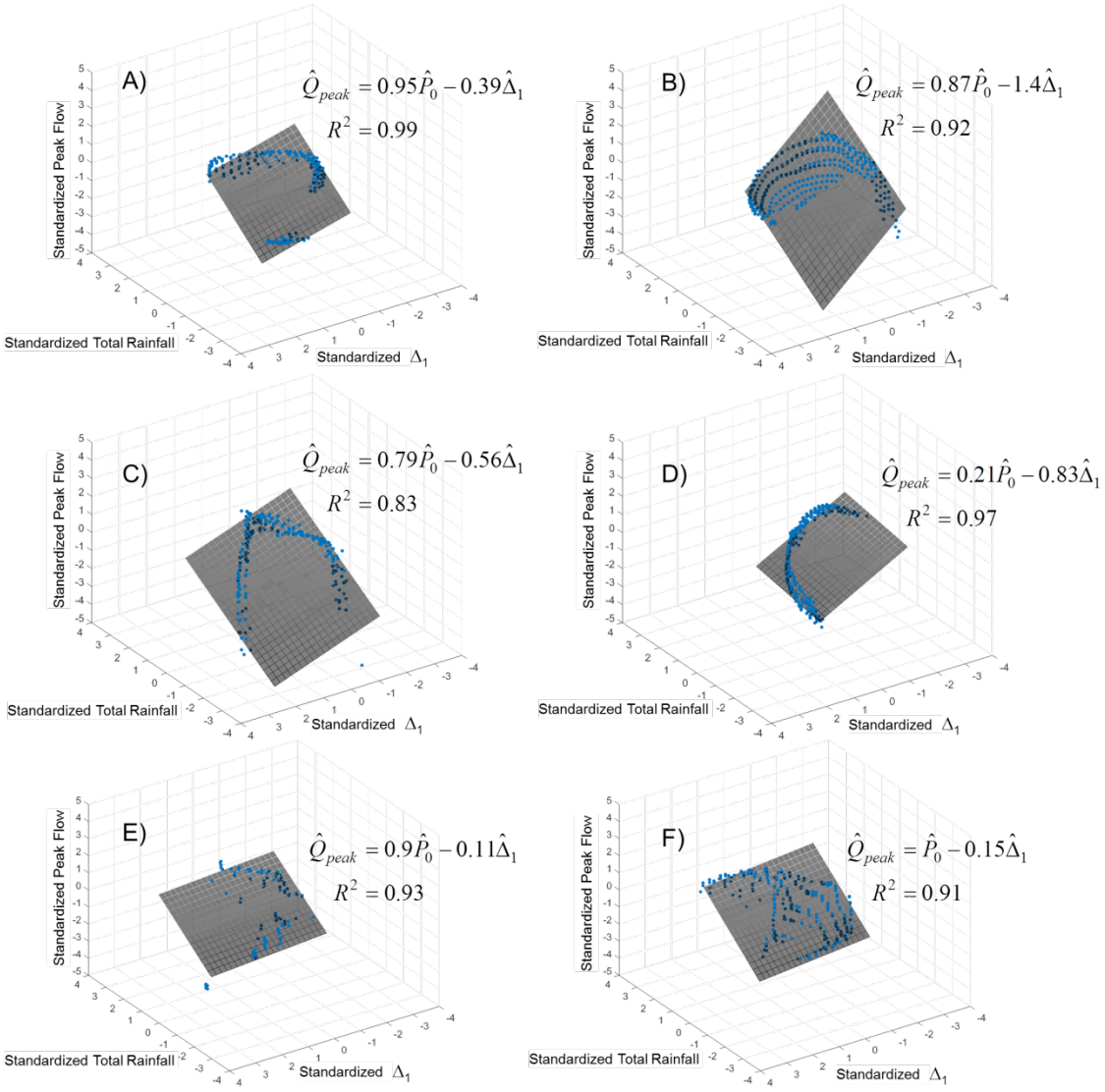



**Figure 12:** 3-D plots of the data (scattered dots) and the resulted bivariate linear models (meshed plain) at the Junction 3 and Junction 6 for the 2015 Memorial Day storm (A and B), the 2015 October storm (C and D) and the 2016 Tax Day storm (E and F).

**Table 3:** $R^2$ Values for Simple Linear Regression and Bivariate Linear Regression Analyses

| | **2015 Memorial Day Storm @ Junction 3 (S1)** | **2015 Memorial Day Storm @ Junction 6 (S2)** |
|---|---|---|
| $\boldsymbol{R^2}$ **using only** $\hat{P}$ | 0.88 | 0.01 |
| $\boldsymbol{R^2}$ **using both** $\hat{P}$ **and** $\hat{\Delta}_1$ | 0.99 | 0.92 |

| | 2015 October 31 Storm @ Junction 3 (S3) | 2015 October 31 Storm @ Junction 6 (S4) |
|---|---|---|
| $R^2$ using only $\hat{P}$ | 0.66 | 0.65 |
| $R^2$ using both $\hat{P}$ and $\hat{\Delta}_1$ | 0.84 | 0.97 |
| | **2016 Tax Day Storm @ Junction 3 (S5)** | **2016 Tax Day Storm @ Junction 6 (S6)** |
| $R^2$ using only $\hat{P}$ | 0.92 | 0.90 |
| $R^2$ using both $\hat{P}$ and $\hat{\Delta}_1$ | 0.93 | 0.91 |

The standardized Beta coefficients for the six simulation scenarios are summarized in **Table 4**. $\beta_1$ is found positive for all the scenarios, as greater MAP usually generates greater peak flow rates. All $\beta_2$ values are however negative because greater $\Delta_1$ values means longer translation distance from the storm core to the receiving junction, causing more attenuation on the peak flow rate. As $\beta_1$ and $\beta_2$ respectively measure the importance of $\Delta_1$ and $\Delta_2$, $|\beta_1/\beta_2|$ simply indicates the relative importance between the two regressors. In comparison of the results from the three storms in **Table 4**, $|\beta_1/\beta_2|$ is found smaller for the two storms in 2015 than for the 2016 Tax Day storm at both Junctions (3 and 6), which means that $\Delta_1$ is more influential to the peak flow rate for these two storms of smaller dimension. Comparing the $|\beta_1/\beta_2|$ results at the two junctions (3 and 6), for all three storms, the authors find that values at Junction 6 are smaller than those at Junction 3 indicating greater importance of $\Delta_1$ for a larger-sized catchment. Therefore, $\Delta_1$ becomes more influential to peak flow rate as the storm dimension decreases or as the catchment dimension increases. Both findings are consistent because increasing the storm dimension essentially has the same effect as decreasing the catchment dimension. The similar findings are confirmed by Syed et al., (2003) that the position of storm core relative to the

catchment outlet becomes more important to rainfall-runoff process as the catchment dimension increases.

**Table 4:** Standard Beta Coefficients for Bivariate Linear Regression Analysis

| | **2015 Memorial Day Storm @ Junction 3** | **2015 Memorial Day Storm @ Junction 6** |
|---|---|---|
| $\beta_1$ | 0.95 | 0.87 |
| $\beta_2$ | -0.38 | -1.4 |
| $\lvert\beta_1 / \beta_2\rvert$ | 2.50 | 0.62 |
| | **2015 October 31 Storm @ Junction 3** | **2015 October 31 Storm @ Junction 6** |
| $\beta_1$ | 0.93 | 0.21 |
| $\beta_2$ | -0.39 | -0.83 |
| $\lvert\beta_1 / \beta_2\rvert$ | 2.36 | 0.25 |
| | **2016 Tax Day Storm @ Junction 3** | **2016 Tax Day Storm @ Junction 6** |
| $\beta_1$ | 0.9 | 1 |
| $\beta_2$ | -0.11 | -0.15 |
| $\lvert\beta_1 / \beta_2\rvert$ | 8.38 | 6.87 |

## CONCLUSIONS AND FUTURE WORK

The authors investigate the potential hydrologic impact from possible storm centering locations via deterministic storm transposition. Spatial moments of catchment rainfall are employed to characterize the spatial organization of catchment rainfall in relative to drainage network. Furthermore, multivariate linear regression is applied to interpret the patterns exhibited in the hydrologic responses from the considerable storm transposition scenarios. The numerical techniques utilized in this study can be replicated for other areas given available data and models. The insights gained from this study would benefit the decision making and emergency management processes during severe storm events. Several major findings from this study are summarized as below:

1. Storm dimension is a significant variable to the spatial organization of catchment rainfall: small-sized storms tend to generate spatially concentrated catchment rainfall with a large displacement between the centroids of catchment rainfall and drainage network; and large-sized storms exhibit opposite patterns.
2. Peak discharge is positively correlated to cumulated mean areal precipitation (MAP); the importance of cumulated MAP to peak discharge becomes greater as storm dimension increases or as catchment dimension decreases.
3. Peak discharge is inversely correlated to the flow distance from a storm core to the catchment outlet; such relationship becomes stronger as storm dimension decreases or as catchment dimension increases.

The findings by this study are based upon the hypothesis of negligible loss, thus are most meaningful to urban areas. However, the authors think that similar approach can still be applied to less developed areas in future studies, by linking peak discharge with only the intense portion of precipitation (e.g. with intensity greater than 25 mm/hr) instead of total rainfall. It is worth emphasizing that multivariate linear regression is used for interpretation rather than prediction. Extra caution should be taken when applying regression techniques in flood prediction, since this study shows how various storms would exhibit different linear relationships with hydrologic responses. While a lumped hydrologic model is utilized in this study, the authors are motivated to validate the findings herein using a distributed hydrologic model in a future study, which will enable us to gain a deeper insight of the relationship between hydrologic responses and flow distance/flow direction. In addition to spatial variability of rainfall, temporal features of rainfall and runoff will be potential focuses for future studies using spatial moments of catchment rainfall.

## ACKNOWLEDGMENT

This research was partially funded by the Tarrant Reginal Water District (TRWD). The authors would like to thank the West Gulf River Forecast Center (WGRFC) for providing radar rainfall data, and the United States Army Corps of Engineers (USACE) for providing the HEC-MetVUE software. In addition, the authors would especially like to thank, Dr. Philip B. Bedient from the Severe Storm Prediction, Education and Evacuation from Disasters (SSPEED) Center and Dr. Baxter Vieux at Vieux & Associates, Inc. for their valuable comments during the period of this study.

## REFERENCES

Ball, J. E. (1994). The influence of storm temporal patterns on catchment response. *Journal of Hydrology*, *158*(3), 285-303.

Bass, B., Juan, A., Gori, A., Fang, Z., & Bedient, P. (2016). 2015 Memorial Day Flood Impacts for Changing Watershed Conditions in Houston. *Natural Hazards Review*, 05016007.

Bedient, P. B., Holder, A., Benavides, J. A., & Vieux, B. E. (2003). Radar-based flood warning system applied to Tropical Storm Allison. *Journal of Hydrologic Engineering*, *8*(6), 308-318.

Bedient P.B., Huber W. C. (2002). *Hydrology and Floodplain Analysis*, 3rd Ed. Prentice-Hall, Upper Saddle River NJ: 763.

Benson, S. A. (2014). Comparative Analysis Of Flood Frequency Based On Radar-based Precipitation Data And Precipitation Trends.

Chen, S., Gourley, J. J., Hong, Y., Cao, Q., Carr, N., Kirstetter, P. E., ... & Flamig, Z. (2016). Using citizen science reports to evaluate estimates of surface precipitation type. *Bulletin of the American Meteorological Society*, *97*(2), 187-193.

Cox, D.R. & Isham, V. (1988). A simple spatial-temporal model of rainfall. *Proceedings of the Royal Society, London Series A – Mathematical, Physical and Engineering Sciences*, 415, 31.

Cocks, S. B., Martinaitis, S. M., Kaney, B., Zhang, J., & Howard, K. (2016). MRMS QPE Performance during the 2013/14 Cool Season. *Journal of Hydrometeorology*, *17*(3), 791-810.7–328.

de Lima, J. L. M. P., & Singh, V. P. (2002). The influence of the pattern of moving rainstorms on overland flow. *Advances in Water Resources*, *25*(7), 817-828.

England, J. F., Julien, P. Y., & Velleux, M. L. (2014). Physically-based extreme flood frequency with stochastic storm transposition and paleoflood data on large watersheds. *Journal of hydrology*, *510*, 228-245.

Fang, Z., Bedient, P. B., & Buzcu-Guven, B. (2011). Long-term performance of a flood alert system and upgrade to FAS3: A Houston, Texas, case study. *Journal of Hydrologic Engineering*, *16*(10), 818-828.

Féral, L., Sauvageot, H., Castanet, L., & Lemorton, J. (2003). HYCELL—A new hybrid model of the rain horizontal distribution for propagation studies: 1. Modeling of the rain cell. *Radio Science*, *38*(3), 22-1.

Foufoula-Georgiou, E. (1989). A probabilistic storm transposition approach for estimating exceedance probabilities of extreme precipitation depths. *Water Resources Research*, *25*(5), 799-815.

Franchini, M., Helmlinger, K. R., Foufoula-Georgiou, E., & Todini, E. (1996). Stochastic storm transposition coupled with rainfall—runoff modeling for estimation of exceedance probabilities of design floods. *Journal of hydrology*, *175*(1-4), 511-532.799-815.

Hansen, E. M. (1987). Probable maximum precipitation for design floods in the United States. Journal of Hydrology, 96(1-4), 267-278.

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

Lee K. T., & Huang J. K., (2007). Effect of moving storms on attainment of equilibrium discharge. *Hydrol Process.*, 21(24), 3357–3366.

Martinaitis, S. M., Gourley, J. J., Flamig, Z. L., Argyle, E. M., Clark III, R. A., Arthur, A., ... & Albright, B. (2016). The HMT multi-radar multi-sensor hydro experiment. *Bulletin of the American Meteorological Society*, (2016).

Michaud, J. D., & Sorooshian, S. (1994). Effect of rainfall-sampling errors on simulations of desert flash floods. *Water Resources Research*, *30*(10), 2765-2775.

Morin, E., Goodrich, D. C., Maddox, R. A., Gao, X., Gupta, H. V., & Sorooshian, S., (2006). Spatial patterns in thunderstorm rainfall events and their coupling with watershed hydrological response. *Advances in Water Resources*, 29(6), 843-860.

Noriah, A. B., & Rakhecha, P. R. (2001). Probable maximum precipitation for 24 h duration over southeast Asian monsoon region—Selangor, Malaysia. *Atmospheric research*, *58*(1), 41-54.

Ogden, F. L., Richardson, J. R., & Julien, P. Y., (1995). Similarity in catchment response: 2. Moving rainstorms, *Wat. Resour. Res.*, 31, 1543-1547.

Paschalis, A., Molnar, P., Fatichi, S., & Burlando, P. (2013). A stochastic model for high-resolution space-time precipitation simulation. *Water Resources Research*, *49*(12), 8400-8417.

Peleg, N., & Morin, E. (2014). Stochastic convective rain-field simulation using a high-resolution synoptically conditioned weather generator (HiReS-WG). *Water Resources Research*, *50*(3), 2124-2139.

Rakhecha, P. R., & Kennedy, M. R. (1985). A generalized technique for the estimation of probable maximum precipitation in India. *Journal of Hydrology*, *78*(3-4), 345-359.

Richardson, J. R., & Julien, P. Y. (1989). One-dimensional modeling of moving rainstorms, in Catchment Runoff and Rational Formula, Yen, B.C., ed., Water Resources Publications, Littleton, CO, 155-167.

Smith, J. A., Baeck, M. L., Morrison, J. E., Sturdevant-Rees, P., Turner-Gillespie, D. F., & Bates, P. D. (2002). The regional hydrology of extreme floods in an urbanizing drainage basin. *Journal of Hydrometeorology*, *3*(3), 267-282.

Smith, J. A., Baeck, M. L., Meierdiercks, K. L., Nelson, P. A., Miller, A. J., & Holland, E. J. (2005). Field studies of the storm event hydrologic response in an urbanizing watershed. *Water Resources Research*, *41*(10).

Stephenson, D. (1984). Kinematic study of effects of storm dynamics on runoff hydrographs. *Water SA*, 10, 189-196.

Singh, V.P. (1998). Effect of the direction of storm movement on planar flow. *Hydrological Processes*, 12, 147-170.

Surkan, A. J. (1974). Simulation of storm velocity effects on flow from distributed channel networks. *Water Resour. Res.*, 10(6), 1149–1160.

Syed, K. H., Goodrich, D. C., Myers, D. E., & Sorooshian, S. (2003). Spatial characteristics of thunderstorm rainfall fields and their relation to runoff. *Journal of Hydrology*, *271*(1), 1-21.

Viglione, A., Chirico, G. B., Woods, R., & Blöschl, G. (2010 a). Generalized synthesis of space–time variability in flood response: an analytical framework. *Journal of Hydrology*, *394*(1), 198-212.

Viglione, A., Battista Chirico, G., Komma, J., Woods, R., Borga, M., & Blöschl, G. (2010 b). Generalized synthesis of space-time variability in flood response: Dynamics of flood event types. In *EGU General Assembly Conference Abstracts* (Vol. 12, p. 5707).

Von Hardenberg, J., Ferraris, L., & Provenzale, A. (2003). The shape of convective rain cells. *Geophysical research letters*, *30*(24).

Wealands, S. R., Grayson, R. B., & Walker, J. P. (2005). Quantitative comparison of spatial fields for hydrological model assessment—some promising approaches. *Advances in Water Resources*, *28*(1), 15-32.

Wright, D. B., Smith, J. A., Villarini, G., & Baeck, M. L. (2013). Estimating the frequency of extreme rainfall using weather radar and stochastic storm transposition. *Journal of hydrology*, *488*, 150-165.

Wright, D. B., Smith, J. A., & Baeck, M. L. (2014). Flood frequency analysis using radar rainfall fields and stochastic storm transposition. *Water Resources Research*, *50*(2), 1592-1615.

Wang, W. (1987). A research on statistical estimation of storm transposition. *Journal of Hydrology*, *96*(1-4), 345-354.

Willie, D., Chen, H., Chandrasekar, V., Cifelli, R., Campbell, C., Reynolds, D., ... & Zhang, Y. (2016). Evaluation of Multisensor Quantitative Precipitation Estimation in Russian River Basin. *Journal of Hydrologic Engineering*, E5016002.

Woods, R., & Sivapalan, M. (1999). A synthesis of space-time variability in storm response: Rainfall, runoff generation, and routing. *Water Resources Research*, *35*(8), 2469-2485.

Yen, B. C., & V. T. Chow, (1969). A laboratory study of surface runoff due to moving rainstorms. *Water Resour. Res.*, 5(5), 989–1006.

Zhang, J., Howard, K., Langston, C., Kaney, B., Qi, Y., Tang, L., ... & Arthur, A. (2016). Multi-Radar Multi-Sensor (MRMS) quantitative precipitation estimation: Initial operating capabilities. *Bulletin of the American Meteorological Society*, *97*(4), 621-638.

Zoccatelli, D., Borga, M., Viglione, A., Chirico, G. B., & Blöschl, G. (2011). Spatial moments of catchment rainfall: rainfall spatial organization, basin morphology, and flood response. *Hydrology and Earth System Sciences*, *15*(12), 3767-3783.

# Chapter 3:

# Investigating Hydrologic Responses to Spatio‐temporal Characteristics of Storms Using a Dynamic Moving Storm (DMS) Generator[2]

Shang Gao and Zheng Fang

Gao, S., & Fang, Z. N. Investigating Hydrologic Responses to Spatio-temporal Characteristics of Storms Using a Dynamic Moving Storm (DMS) Generator. Hydrological Processes, WILEY, 2019 DOI: 10.1002/hyp.13524

[2] Used with Permission from WILEY, the publisher of Hydrological Processes, 2019

## ABSTRACT

A synthetic storm generator – Dynamic Moving Storm (DMS) – is developed in this study to represent spatio-temporal variabilities of rainfall and storm movement in synthetic storms. Using an urban watershed as the testbed, the authors investigate the hydrologic responses to the DMS parameters and their interactions. In order to reveal the complex nature of rainfall-runoff processes, previously simplified assumptions are relaxed in this study regarding 1) temporal variability of rainfall intensity and 2) time-invariant flow velocity in channel routing. The results of this study demonstrate the significant contribution of storm moving velocity to the variation of peak discharge based on a global sensitivity analysis (GSA). Furthermore, a pairwise sensitivity analysis (PSA) is conducted to elucidate not only the patterns in individual contributions from parameters to hydrologic responses but also their interactions with storm moving velocity. The intricacies of peak discharges resulting from sensitivity analyses are then dissected into independent hydrologic metrics, i.e. runoff volume and standard deviation of runoff timings, for deeper insights. It is confirmed that peak discharge is increased when storms travel downstream along the main channel at the speed that corresponds to a temporal superposition of runoff. Spatial concentration of catchment rainfall is found to be a critical linkage through which characteristics of moving storms affect peak discharges. In addition, altering peak timing of rainfall intensity in conjunction with storm movement results in varied storm core locations in the channel network, which further changes the flow attenuation effects from channel routing. For future directions, the DMS generator will be embedded in a stochastic modeling framework and applied in rainfall/flow frequency analysis.

## KEYWORDS:

Storm Movement, Spatial Variability, Temporal Variability, Hydrologic Response, Spatial Concentration of Catchment Rainfall, Global Sensitivity Analysis, Synthetic Design Storm, and Urban Watershed.

# 1. INTRODUCTION

With rapid urbanization and climate change, water-related hazards increasingly place human lives and the economy at risk. To make informed decisions in proactive response to water-related hazards and risks, it is imperative to fully understand the sensitive nature of hydrologic responses from the spatio-temporal characteristics of storms. This subject has motivated tremendous research effort undertaken by hydrologists and meteorologists (e.g. Smith et al., 2005b; Borga et al., 2007; Segond et al., 2007; Rozalis et al., 2010; Tarolli et al., 2013; Furl et al., 2015; Zhou et al., 2018). Therefore, in hydrologic modeling, identifying key elements in rainfall input, such as storm movement and spatio-temporal variability, is paramount to runoff calculation. Storm movement is a critical component in the rainfall process, as most storms exhibit preferential moving velocities and directions unique to different seasons (Marshall, 1980). In the past decades, researchers invested varieties of efforts on exploring the hydrologic responses from storm movement using physical, numerical and analytical approaches (e.g. Yen and Chow, 1969; Surkan, 1974; Ogden, et al., 1995; de Lima and Singh, 2002; Lee and Huang, 2007; Fang et al., 2019).

In the early age of storm movement studies, the typical approach was to conduct experimental techniques with physical models. Pioneers in this regard include Yen and Chow (1969) who established the Water Experimental System (WES) to investigate the effect of storm movement on surface runoff. WES is essentially a collection of simplified V-shaped catchments and water distribution facilities with a rain simulator over synthetic watersheds. They introduced

a new concept of “equivalent storm” defining that a storm moving upstream generates a smaller peak discharge than that from an equivalent storm moving downstream at the same velocity. This physical approach provided visualization and insights for early researchers and was further used as a tool of verification for later-developed numerical approaches. However, the approach of using physical models in this research was expensive and time-consuming, thus limiting its accessibility to the majority of researchers (Liang, 2010).

Mathematical rainfall-runoff models were extensively utilized in the past several decades to examine the magnitude and timing of peak discharge in response to storm movement. For example, Surkan (1974) performed a study on storm movement using a simplified distributed model to conduct a sensitivity analysis to understand how peak flow rates change with respect to storm moving direction and velocity. Stephenson (1984) was the first to apply kinematic-wave approximation to explore the relationship between storm movements and overland flows, and found that for a storm moving upstream, the peak flow was less than that for a stationary storm. Richardson and Julien (1989) used the CASC numerical model to verify the laboratory results previously reported by Yen and Chow in 1968. Ogden et al. (1995) utilized the CASC finite element runoff model for one- and two-dimensional numerical runoff simulations; and their studies showed that the equilibrium discharge can only be attained with storm lengths greater than the watershed length. In regard to channelized catchments, Niemczynowicz utilized the Storm Water Management Model (SWMM) to investigate hydrologic responses from multiple parameters of storm movement using initially a conceptual watershed (1984a) and later a real urbanized area in Sweden (1984b); and the researcher reported that the maximum discharge occurs when a storm moves downstream at a speed equal to the average channel-flow velocity in an urban storm-sewer system. Lee and Huang (2007) utilized kinematic-wave approximation and

finite-difference technique to simulate the effect of moving storms on the attainment of equilibrium discharge for two conceptual watersheds, an overland plane and a V-shaped catchment. It reported that for storms moving downstream across the watershed, an equilibrium discharge can be obtained even if the storm length is shorter than the watershed length.

Furthermore, several researchers derived analytical solutions for runoffs as functions of storm movements. They employed necessary simplified assumptions, yet provided deep understandings on the underlying mechanism of hydrologic transformation. For instance, Singh (1998) utilized kinematic wave equations to perform a methodical study to evaluate discharge from a one-dimensionally moving storm over a simplified 1-D watershed. Singh's results indicated how maximum peak flow occurs when the moving velocity of the storm is identical with that of the stream flow. Afterwards, de Lima and Singh (2002) further exercised finite-difference methods and derived analytical solutions to generate discharges resulting from a moving storm of spatially-distributed coverage over the catchment. Viglione et al. (2010) produced an analytical framework that explicitly quantified the significance of spatio-temporal variability of rain intensity as well as storm movement to the resulted peak discharges. Seo et al. (2012) established a conceptual model based on peak timing of rainfall intensity in hydrologic response to investigate the effects of storm movement on peak discharges and analytically interpret the causes for varied peak responses from a moving storm. Volpi et al., (2013) developed an analytical framework for analyzing hydrologic response from a catchment under moving rainstorms by assuming time-invariant of distribution of travel time within the basin.

For decades, studies in spatial pattern of storms utilized rain gauge networks as a major data source which inherently lacks the density necessary to capture the spatial variability (Michaud and Sorooshian, 1994). It was not until weather radar systems were widely applied in hydrology

in 1990s that a detailed understanding of rainfall events with high spatial resolution was acquired. Numerous studies took advantage of this radar rainfall technology and generated meaningful outcomes about spatial and temporal structures (e.g., Feral et al., 2003; von Hardenberg 2003; Wealands et al., 2005; Peleg and Morin, 2014).

Previously developed rainfall generators have incorporated storm movement and spatial variability of rain intensity (e.g., Cowpertwait et al., 2002; Peleg and Morin, 2012) with the simplified assumption of temporally constant rainfall. Despite its obvious deviation from the nature of rainfall, this simplification was made because temporal change of rain intensity was considered a lesser factor to hydrologic response compared with movement, size, or duration of the storm. However in the authors' opinion, this simplified assumption prevents investigating particular scenarios where the interaction between temporal change of rain intensity and other storm parameters (e.g., storm movement and size) can be impactful although it holds well considering on average how catchments react to storms. Therefore, the DMS generator here aims to provide generality in designing synthetic storms by allowing users to specify temporal change rate and the peak timing of rain intensity. In addition, most analytical frameworks used to study hydrologic responses to storm parameters assume time-invariant flow velocity across the drainage network and thus neglect flow attenuation from channel routing (e.g. Gupta et al., 1996; Mandapaka et al., 2009; Seo et al., 2012., Volpi et al., 2013). Moreover, numerical modeling approaches in some aforementioned studies (e.g. Stephenson, 1984; Singh, 1998; de Lima and Singh, 2002; Lee and Huang, 2007) adopt kinematic wave approximation which is incapable of modeling flow attenuation. The inadequate representation of flow attenuation effect in previous studies of moving storms compromises the validity of their findings in watersheds with flat terrain and mild slopes. Therefore, the authors think that it is imperative to re-evaluate the

hydrologic responses from the flat watersheds under moving storms using the most realistic hydrologic model available.

Given that Houston, Texas has undergone a rapid urban expansion and is prone to intense tropical storms with a history of major flood events like Tropical Storm Allison (2001) and Tropical Storm Harvey (2017), the Brays Bayou watershed (**Figure 1**) with 334 $km^2$ of area is selected for this study. The watershed is one of the most highly-developed areas in Houston and is characterized by very flat terrain, tight clay soils, and stream channel slopes of 0.2 to 0.4 m/km (Bedient et al., 2003). Brays Bayou is a flood-prone watershed due to over a 90% development percentage. In the late 1960s, a major concrete channelization project was completed to accommodate storms at the 100-year level. However, by the year of 1983, a 10-year storm would bring the channel to its full capacity due to the rapid development throughout the watershed. By the year of 1993 the watershed had reached over 90% development, leading to bankfull conditions at Main Street in response to only a 5 to 10-year storm (Bedient et al., 2018). After the devastation of Tropical Storm Allison in 2001, major efforts from the City of Houston, Federal Emergency Management Agency (FEMA), Harris County Flood Control District (HCFCD), and the Texas Medical Center (TMC) were invested to resolve the flooding issues (Fang et al., 2011; Bass et al., 2016). Given the considerable hydrologic experiences and practices in Brays Bayou, this well-studied watershed serves as an ideal study area for hydrologic analysis. By performing thorough sensitivity analyses between individual and joint factors from moving storms and the computed runoffs from hydrologic models, a better understanding of the dependency of hydrologic responses on spatiotemporal characteristics of rainfall can be achieved. This study is performed to achieve the following objectives:

1) To conceptualize a rainfall event using three fundamental components, i.e. spatial variability of rainfall intensity, temporal variability of rainfall intensity and storm movement.
2) To conduct global sensitivity analysis (GSA) on rainfall parameters and evaluate their individual contribution to the variation of corresponding hydrologic responses in a real catchment.
3) To explore the pattern in the hydrologic responses from the sensitive rainfall parameters (determined by GSA in Step 2) as well as their potential interdependences using pairwise sensitivity analysis (PSA).

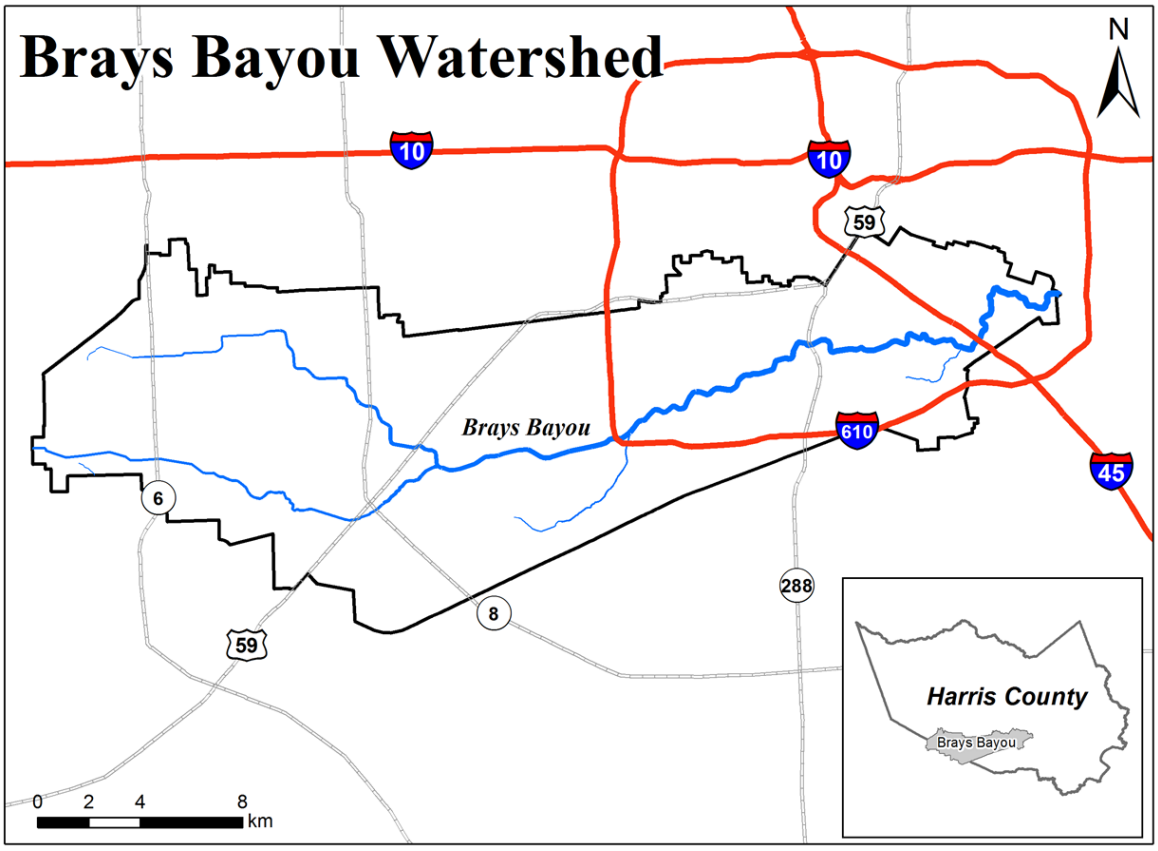



**Figure 1** The Brays Bayou watershed in Houston, Texas.

## 2. METHODOLOGY

### *2.1 Concept of Dynamic Moving Storm*

The fundamental framework of the Dynamic Moving Storm (DMS) generator consists of three basic modules, i.e. spatial variability of rain intensity, temporal variability of rain intensity, and storm movement, as shown in **Figure 2**. First, spatial variability refers to the spatial distribution of instantaneous rainfall intensity within the storm. In this study, storms are represented by idealized rain cells as isotropic elements with maximum rainfall intensity at the storm centroid. Second, temporal variability represents the changing pattern of rainfall intensity with respect to time. Finally, storm movement in this study refers to the temporal change of storm position on a 2-dimensional horizontal plain. For simplicity, storm movement is essentially represented by the movement of storm centroid disregarding circulation or minor dynamics within the storm.

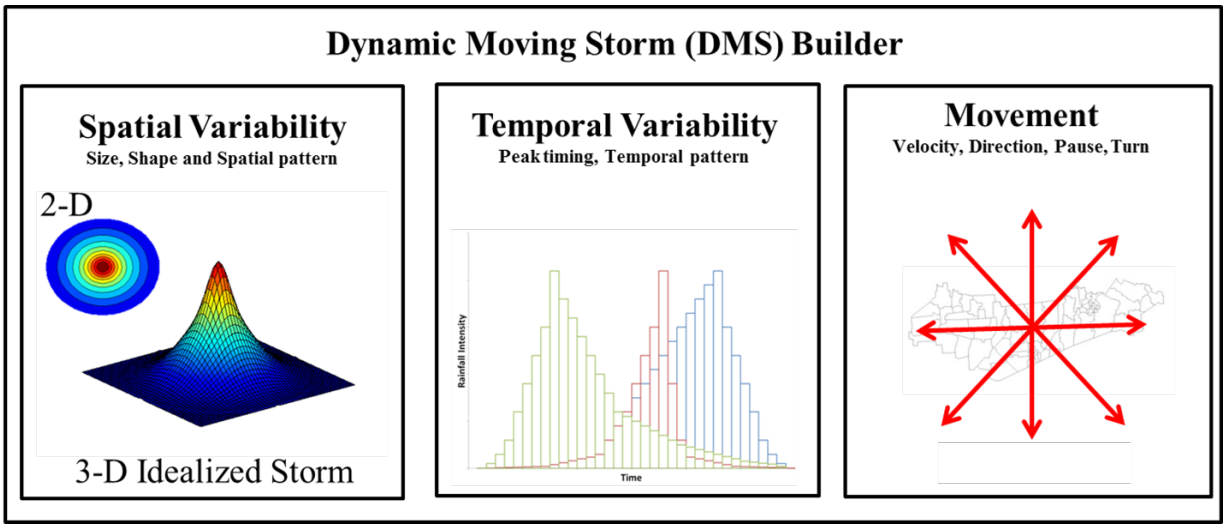



**Figure 2:** Modeling structure of the Dynamic Moving Storm (DMS) generator.

### *2.2 Simulation Scheme*

**Figure 3** illustrates the simulation scheme using the DMS generator over Brays Bayou. A synthetic storm is designed as a circular-shaped isotropic rain cell moving and evolving over time. For all tested scenarios, a single storm is initially positioned with its perimeter tangent to the watershed boundary. The storm moving direction ($\theta$) is defined as the counter-clockwise angle from the west-east direction. The storm with a constant diameter ($L_s$) moves in a straight line across the watershed's centroid with a constant velocity ($v$). The storm ceases once its coverage completely exits the watershed. Secondly, a temporal pattern controls the dynamic change of the rainfall intensity over time via a simplified linear relationship (**Equation 1**):

$$R_0 = \max\left(R_{max}\left(1 - \gamma\left|t - T \cdot t_p\right|\right), 0\right), \quad 0 \le t \le T \tag{1}$$

where $R_0$ is the rainfall intensity at the storm center; $t$ is elapsed time and $T$ is the length of time between the initial occurrence of precipitation in the watershed and the final occurrence; $R_{max}$ is the maximum rainfall intensity and is kept constant as 76mm/hr (3 in/hr) which corresponds to the hourly rainfall of a 10-year storm in Houston (National Weather Service, 2017); $\gamma$ is the changing rate of rainfall intensity with the unit of [time]$^{-1}$, with $\gamma = 0$ meaning a temporally uniform pattern of rainfall intensity; $t_p$ is the dimensionless timing (between 0 and 1) when the $R_{max}$ occurs, in percentage of $T$. For instance, $t_p$ = 0.5 means the rainfall intensity peaks when the storm travels halfway over the watershed. The operator max[] returns the maximum of the variables. For simplicity, the same $\gamma$ value is assumed for both the intensification and decay of rainfall intensity. In addition, $T$ and $\gamma$ jointly determine the duration of catchment rainfall ($D$): $D$ is constrained by $T$ when the storm moves fast and $T$ is small; whereas a rapid change rate of rainfall intensity (a large $\gamma$) dictates $D$. With $R_0$ generated for each time step, the spatial

distribution of rain intensity within the circular storm can be modeled using an areal decay relationship (**Equation 2**):

$$\boldsymbol{R = R_0 e^{-2\alpha r/L_s}, \quad 0 \le r \le L_s/2} \tag{2}$$

where $R$ is the rainfall intensity within the rain cell; $\alpha$ is a shape-controlling coefficient; and $r$ is the distance to storm centroid ranging from 0 to the storm radius. A greater $\alpha$ value indicates a steeper areal decay rate of rainfall intensity, while the value of $\alpha$ being zero means a spatially uniform distribution. **Equation 2** generates the peak intensity at the cell centroid and decreasing values toward the perimeter, as illustrated by a 3-D idealized storm in **Figure 3**. Similar spatial structure for storm cells was studied by many previous researchers (e.g. von Hardenberg et al, 2003; Feral et al, 2003; Morin et al, 2006; Curtis, 2007).

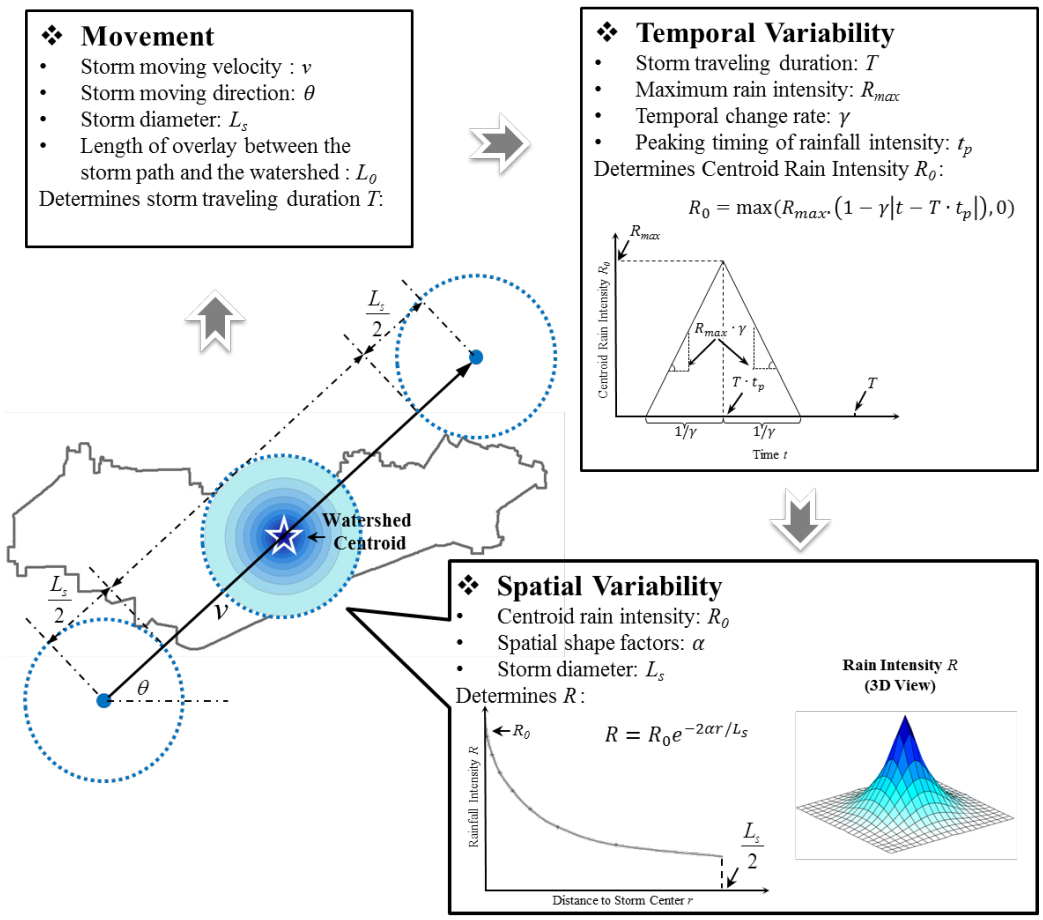



**Figure 3:** Simulation scheme with the DMS Generator over the Brays Bayou watershed.

The time series of gridded incremental rainfall from the DMS generator needs to be converted to sub-basin hyetographs via calculating the corresponding mean areal precipitation (MAP) values at each time step. A hydrologic (Hydrologic Engineering Center Hydrologic Modeling System, HEC-HMS) model is used to receive the preprocessed hyetographs and simulate hydrologic responses from Brays Bayou. The Brays Bayou HEC-HMS model uses Clark Unit Hydrograph to rout overland flow in each of the 76 sub-basins. Therefore, the overland routing in HEC-HMS model assumes time-invariant travel time distributions and time-invariant flow velocities. Regarding channel routing, the Modified Puls method is used to route flows through reaches along Brays Bayou (HEC, 2006). For each reach, the Modified Puls

routing method requires a storage-outflow curve to represent the geometry and storage/attenuation effects for that section of the channel (Fang et al., 2011). As opposed to overland routing, channel routing in the Brays Bayou HEC-HMS model accounts for time-variant flow velocities and the individual flow attenuation effect from each section of the channel.

The Brays Bayou HEC-HMS model was originally developed as part of the products from the Tropical Storm Allison Recovery Project (TSARP). Calibration was later conducted to improve the hydrologic simulation of this model in recent studies (Bass et al., 2016; Gao and Fang, 2018) and the simulation performance driven by radar rainfall was considered accurate compared to the observed hydrographs from stream gauges. The HEC-HMS model used in this study represents the watershed condition in 2016. In addition, the Brays Bayou HEC-HMS model is adjusted to 100% impervious to avoid influences on direct runoff volume caused by the spatial heterogeneity in the infiltration process. It should be noted that infiltration is considered as a minor factor in the hydrological process of the current Brays Bayou watershed due to the fact that the watershed is above 90% developed.

### *2.3 Metrics for Hydrologic Response*

This study focuses on investigating peak discharges ($Q_p$) generated by various combinations of DMS parameters. In order to dissect the intricacies in $Q_p$, two metrics of hydrographs are utilized: (1) runoff volume $V_{runoff}$, i.e. the integral of the hydrograph over time, (2) the standard deviation of timings of runoff, $\delta_T$ (i.e. the temporal dispersion of a hydrograph). These two metrics are independent from each other and can reveal individual aspect of hydrologic responses. In addition, the authors utilize the spatial statistics of catchment rainfall with respect to channel length to estimate the spatial location of rainfall in relative to the channel network.

Catchment rainfall is simply the rainfall received by a catchment, while channel length means the distance from the outlet of a sub-basin to the outlet of the watershed along the path of channels. Various measures of catchment rainfall were developed based on channel length in many previous studies. For instance, Smith et al., (2002, 2005a) utilized a series of scaled metrics called normalized flow distance and normalized dispersion to examine rainfall variability. The mathematical formula (**Equation 3**) is similar to those characterizing flow paths in previous research (Zoccatelli et al., 2011):

$$\Delta = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m} r_{ij} d_i a_i}{|A|^{-1}\sum_{i=1}^{n}\sum_{j=1}^{m} r_{ij} a_i \cdot \sum_{i=1}^{n} d_i a_i} \quad (3)$$

where $\Delta$ is the dimensionless spatial moments of catchment rainfall with respect to channel length; $r_{ij}$ is the rainfall at sub-basin *i* and time step *j*; $a_i$ is the area of sub-basin *i* and *A* is the total area of all the sub-basins; *n* is the total number of subbasins and *m* is the total number of time steps; $d_i$ is the channel length from sub-basin *i* to the watershed outlet. **Figure 4** shows the channel length values of the 76 sub-basins in Brays Bayou. *Δ* describes the location of centroid of catchment rainfall with respect to the centroid of channel network (i.e. the average value of channel length). Values of *Δ* being one (1) reflect a rainfall distribution concentrated at the watershed centroid, with values less than one (< 1) indicating that rainfall is concentrated near the watershed outlet, and values greater than one (> 1) indicating that rainfall is concentrated near the watershed headwaters (Zoccatelli et al., 2011).

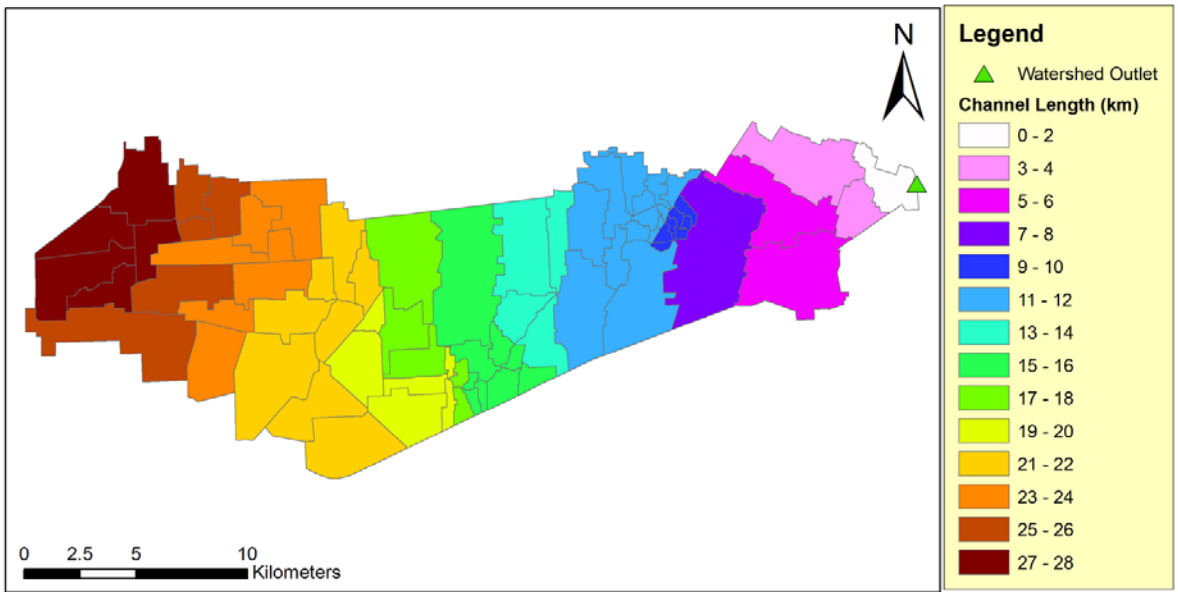



**Figure 4**: Channel length values of the sub-basins in Brays Bayou

### *2.4 Sobol' Sensitivity Analysis*

In this study, global sensitivity analysis (GSA) is conducted to investigate the effects of DMS parameters on the peak discharge at the watershed outlet generated by a single moving storm over Brays Bayou. Sobol's method (Sobol', 1993) is a GSA approach based on variance decomposition, in which any model can be represented in **Equation 4**:

$$Y = f(X) = f(X_1, \dots, X_p) \tag{4}$$

where $Y$ is the model output; $f(X)$ is the model and $X = X_1, \dots, X_n$ is the parameter set. In Sobol's method, the total variance of the model output ($V_t$) is decomposed into component variances from individual parameters and their interactions as shown in **Equation 5**:

$$V_t = \sum_i V_i + \sum_{i<j} V_{i,j} + \sum_{i<j<k} V_{i,j,k} + \cdots + V_{1,2,\dots,n} \tag{5}$$

where $V_i$ is the variance due to the $i^{th}$ parameter $X_i$, and $V_{i,j}$ is the variance due to the interaction between parameters $X_i$ and $X_j$. Sobol's sensitivity indices of different orders, i.e. the sensitivity of single parameter or parameter interaction, is defined as their percentage contribution to the total

variance. In this study, the first-order index $S_i$ is used to quantify the main effect from a single parameter as shown in **Equation 6**:

$$\boldsymbol{S_i = \frac{V_i}{V_t}} \quad (6)$$

The second order index $S_{ij}$ measures the effect of the interaction two parameters as shown in **Equation 7**:

$$S_{ij} = \frac{V_{i,j}}{V_t} \quad (7)$$

The total-order index $S_{Ti}$ measures the main effect of the parameter and its interaction with all the other parameters as shown in **Equation 8**:

$$S_{Ti} = 1 - \frac{V_{\sim i}}{V_t} \quad (8)$$

where $V_{\sim i}$ is the variance due to all the parameters except for $X_i$.

The GSA includes six parameters, i.e. storm moving velocity ($v$), storm moving direction ($\theta$), storm diameter ($L_s$), spatial shape factor of rainfall intensity ($\alpha$ in **Equation 2**), temporal change rate of rainfall intensity ($\gamma$ in **Equation 1**), and peak timing of rainfall intensity ($t_p$ in **Equation 1**). For sampling approach, the authors employ the quasi-Monte Carlo, which is characterized by an enhanced convergence rate compared to the crude Monte Carlo (Sobol', 2001). Without knowledge of the true parameter distributions, the quasi-Monte Carlo uses deterministic uniformly distributed points to effectively prevent sampling from clustering and hence improve convergence rate (Homma and Saltelli, 1995). A large sampling size of 20,000 is used for implementing Sobol's method.

Since the quasi-Monte Carlo requires only ranges and no default values of the parameters, an investigation was conducted to determine reasonable ranges as summarized in **Table 1**. The authors mainly focus on non-tropical storms in the investigation because the spatial structure of rainfall (**Equation 2**) used in the DMS generator is based on non-tropical storm cells. The historical storms are selected based on daily rainfall recorded by the Community Collaborative Rain, Hail and Snow Network (CoCoRaHS). After a threshold of 200-mm daily rainfall (equivalent to a 10-year storm for the study area) being applied in a filtering process, 20 non-tropical storms are identified with the corresponding rain gauges that received the most rainfall during the storms as shown in **Figure 5**. Storm movement is extracted from a Next Generation Radar (NEXRAD) level-III product (code: NST/58) providing locations of identified storm cells every 5 minutes. The NST/58 products are obtained from National Climate Data Center (NCDC) for the past twenty non-tropical storms that occurred in/near Harris County**.** Storm velocities and directions are calculated using 86,972 storm cells detected by the NEXRAD radar in Houston (ID: KHGX) during the 20 storms, as shown in **Figure 6**. It is found that the storm velocities range from 0 to 120 km/hr and that storm directions cover all possible angles (0° to 360°). For spatial parameters of storms, Olivera et al (2008) reported the largest storm area in Texas extracted from NEXRAD radar data was 800 $km^2$ which corresponds to the upper limit of storm diameter (30 km) in this study. Also based on a previous study on storm structure by Curtis (2007), lower limit of storm diameter is estimated to be 5 km according to storm cells extracted from 15-minute radar rainfall data. The upper limit of spatial shape factor $\alpha$ ($\alpha$ = 3) is determined also based on the study by Curtis (2007) where relationships between rain intensity and normalized storm area were estimated. The lower limit of $\alpha$ (= 0) is determined to incorporate simplest possible case of spatially uniform rain intensity. In terms of temporal parameters of

storms, the upper limit of $\gamma$ (= 24) corresponds to a storm duration of 5 minutes, representing the cases found in the NST/58 products where storm cells occur in only one time step. The lower limit of $\gamma$ (=0) represents the simplest possible temporal pattern: a storm of steady intensity moving across the watershed. The range of peak timing of rainfall intensity ($t_p$) is set to cover all possible cases i.e. from 0 to 1. This is because variation of $t_p$ can be viewed as a result of possible locations of storm sequences relative to the watershed which can be assumed to uniformly cover Brays Bayou.

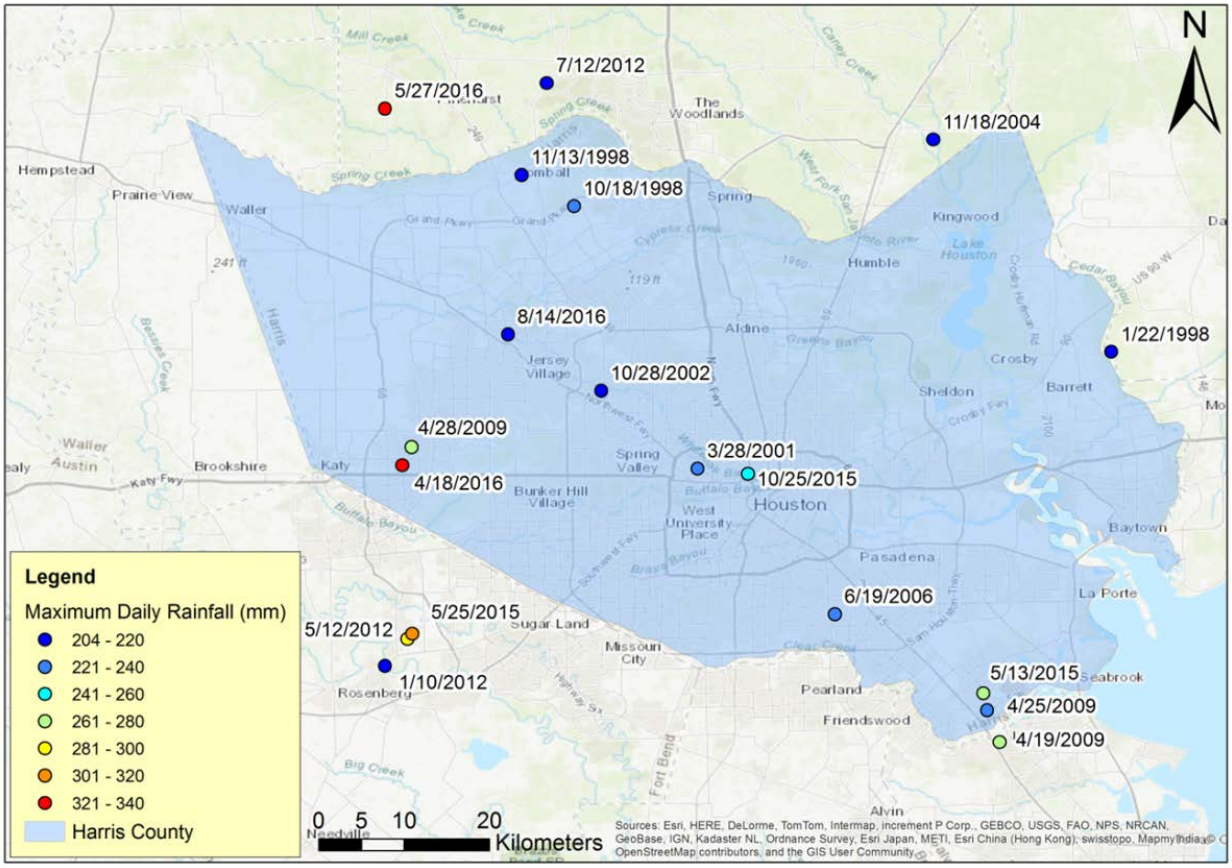



**Figure 5:** Maximum daily rainfall of 20 historical storms in/near Harris County, Texas.

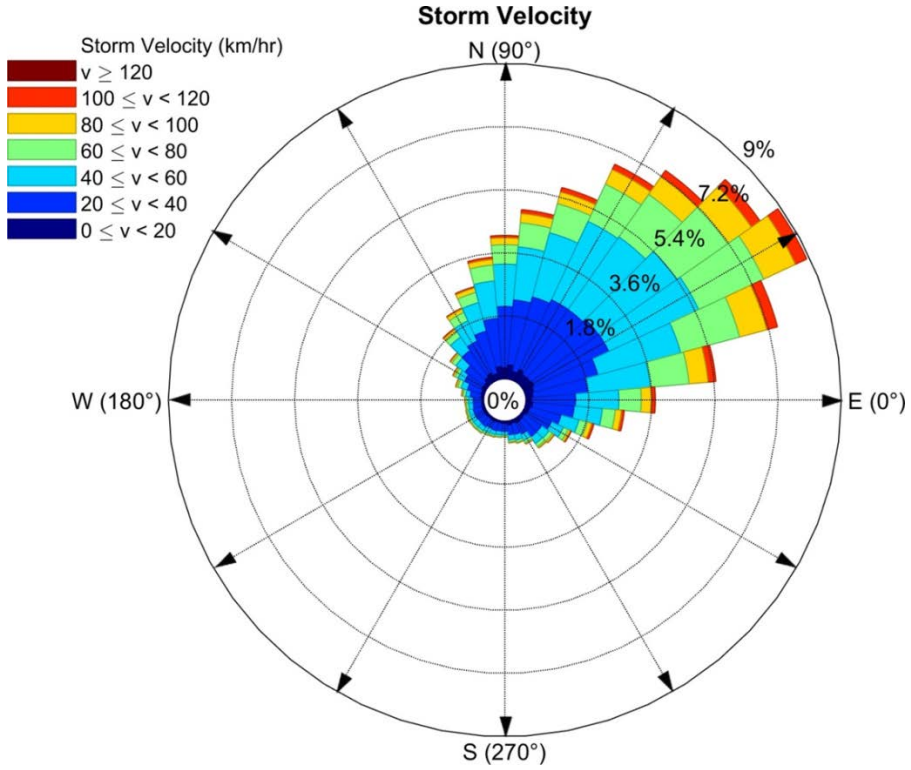



**Figure 6:** Velocities and Directions of Storm Movement Extracted from 20 storms captured by NEXRAD radar KHGX.

**Table 1:** Parameter ranges for Global Sensitivity Analysis (GSA)

| Parameters | Minimum | Maximum |
|---|---|---|
| **Moving Velocity $v$ (km/h)** | 0 | 120 |
| **Moving Direction $\theta$ (°)** | 0 | 360 |
| **Size/Diameter $L_s$ (km)** | 5 | 30 |
| **Spatial Shape Factor $\alpha$** | 0 | 3 |
| **Temporal Change Rate $\gamma$ ($h^{-1}$)** | 0 | 24 |
| **Peak Timing of Rainfall Intensity $t_p$** | 0 | 1 |

## 3. RESULTS AND DISCUSSION

### *3.1 Results from Global Sensitivity Analysis (GSA)*

The first- and total-order sensitivity indices of the six parameters are shown in **Figure 7A**. In **Figure 7A**, the first order indices are represented by the blue bars measuring individual parameter contributions to the variance of peak discharge at the watershed outlet. The total-order indices are signified by the total height of bars, which measure the individual and interactive parameter contributions. It should be noted that yellow bars measure the interactive contribution of one parameter with all the other parameters. It can be found that storm moving velocity ($v$) is the most sensitive parameter with both the first- and total- order indices being the highest among all the parameters. It should also be noted that parameter interactions are higher than the first order indices for all the parameters

Among the six parameters, $v$, $L_s$, $\alpha$, and $\gamma$ generate distinctly higher total-order indices than $\theta$ and $t_p$. This can be generally explained by the variation of excess rainfall as an indicator for peak discharge. According to **Equations 1** and **2**, variation of $L_s$, $\alpha$, and $\gamma$ affects the total amount of rainfall carried by the storm. In comparison, varying $\theta$ and $t_p$ values only changes the catchment rainfall as a proportion of the total amount carried by the storms, and thus are not as influential as $L_s$, $\alpha$, and $\gamma$. Regarding the moving velocity ($v$), it changes the duration of storm traveling across the watershed ($T$), and further the catchment rainfall. In addition, $v$ affects peak discharge due to more complex mechanisms, as discussed later in the section of pairwise sensitivity analysis (PSA).

**Figure 7B** shows the second-order sensitivity indices measuring the contributions of the interactions between any two of the six parameters to the variance of peak discharge at the watershed outlet ($Q_p$). It is found that interactions between $v$ and the other five parameters generate highest second-order sensitivity indices and hence are most influential to $Q_p$. Therefore,

the interactions between $v$ and the other five parameters are investigated individually via pairwise sensitivity analysis (PSA) as detailed in the following sections.

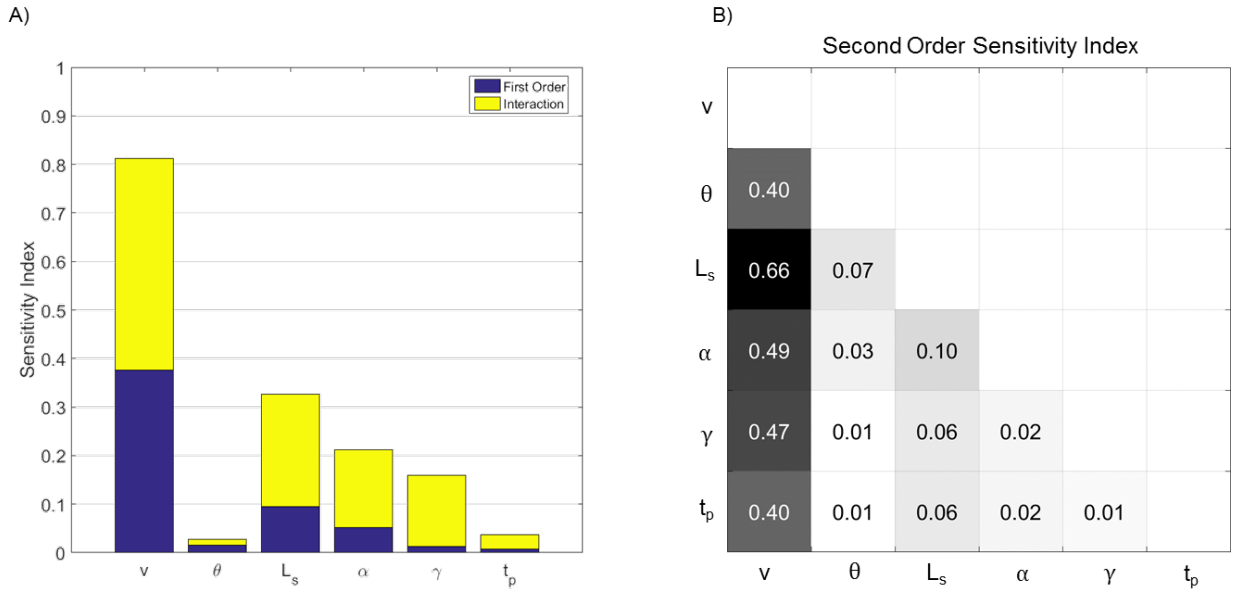
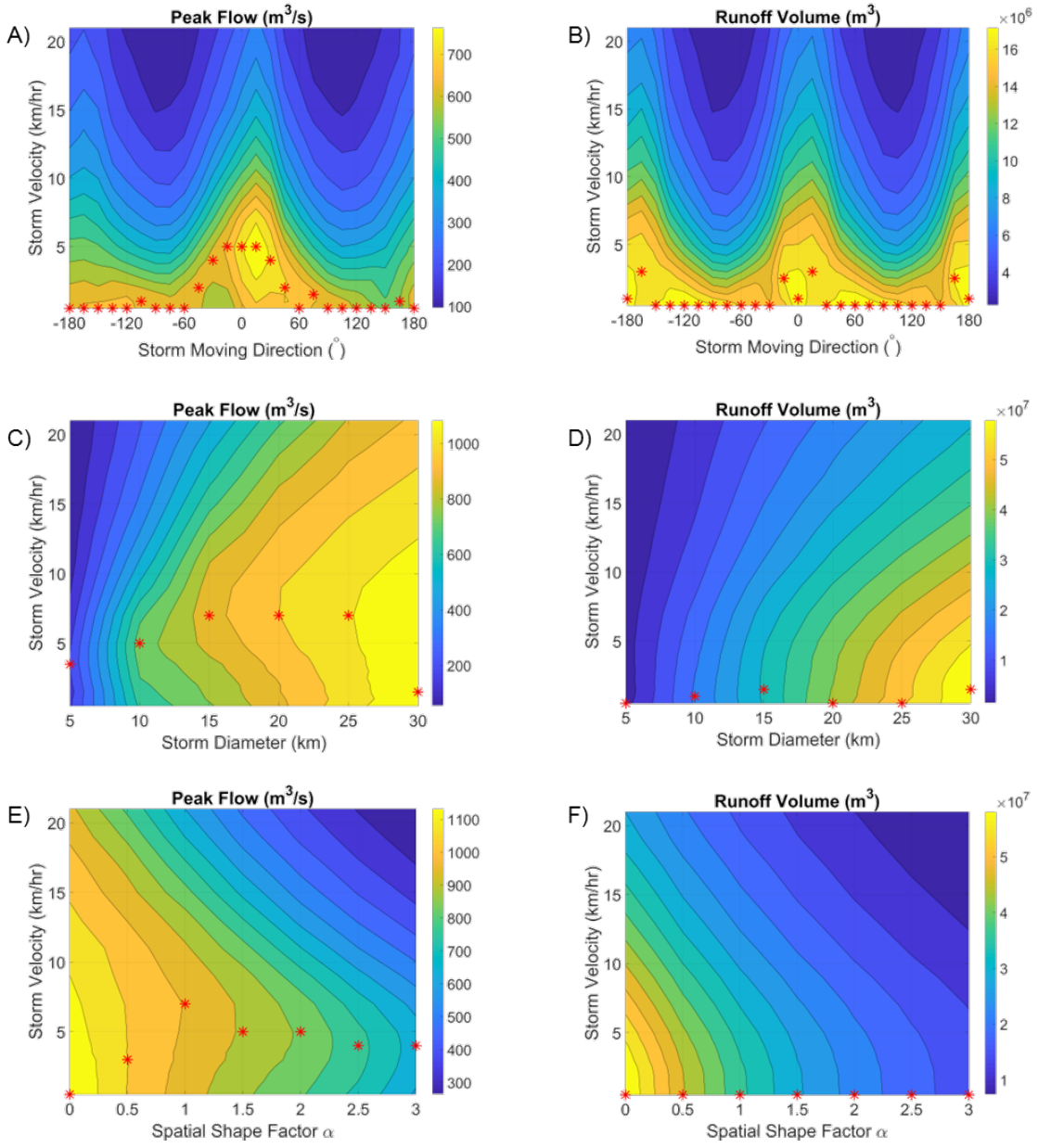



**Figure 7:** Sensitivity indices of the DMS parameters on peak discharge at the watershed outlet, A) first- and total- order sensitivity indices and B) second-order sensitivity indices.

### *3.2 Results from Pairwise Sensitivity Analysis (PSA)*

Pairwise sensitivity analysis (PSA) focuses on investigating the pattern of hydrologic responses from individual parameter as well as their potential interdependence. The PSA is conducted by varying two paired parameters within their own ranges while keeping all the other parameters constant to their default values (**Table 2**). It should be noted that the parameter ranges in PSA are in general smaller subsets of those in GSA, which aims to focus on the scenarios that generate meaningful variation in hydrologic responses and require illustrations. The settings of the default values are for the same purpose. The default value of $L_s$ is specially changed from 10 km to 30 km for the PSA between $v$ and $\alpha$, in order to generate comparable variations to those from the PSA between $v$ and $L_s$. **Figure 8** shows a panel of contour plots of

the peak flow rates ($Q_p$, 1st column – **Figures 8A**, **8C**, **8E**, **8G**, and **8I**) and runoff volume values ($V_{runoff}$, 2nd column - **Figures 8B**, **8D**, **8F**, **8H**, and **8J**) from varied storm moving velocity ($v$) and the other five parameters (five rows), i.e. storm moving direction ($\theta$), storm diameter ($L_s$), spatial shape factor ($\alpha$), temporal change rate ($\gamma$), and peak timing of rainfall intensity ($t_p$); in each panel, the Y-axis represents $v$, and X-axis represents one of the other five parameters; for each X value, the red asterisk marks the Y value ($v$) that generates the highest peak flow or runoff volume.



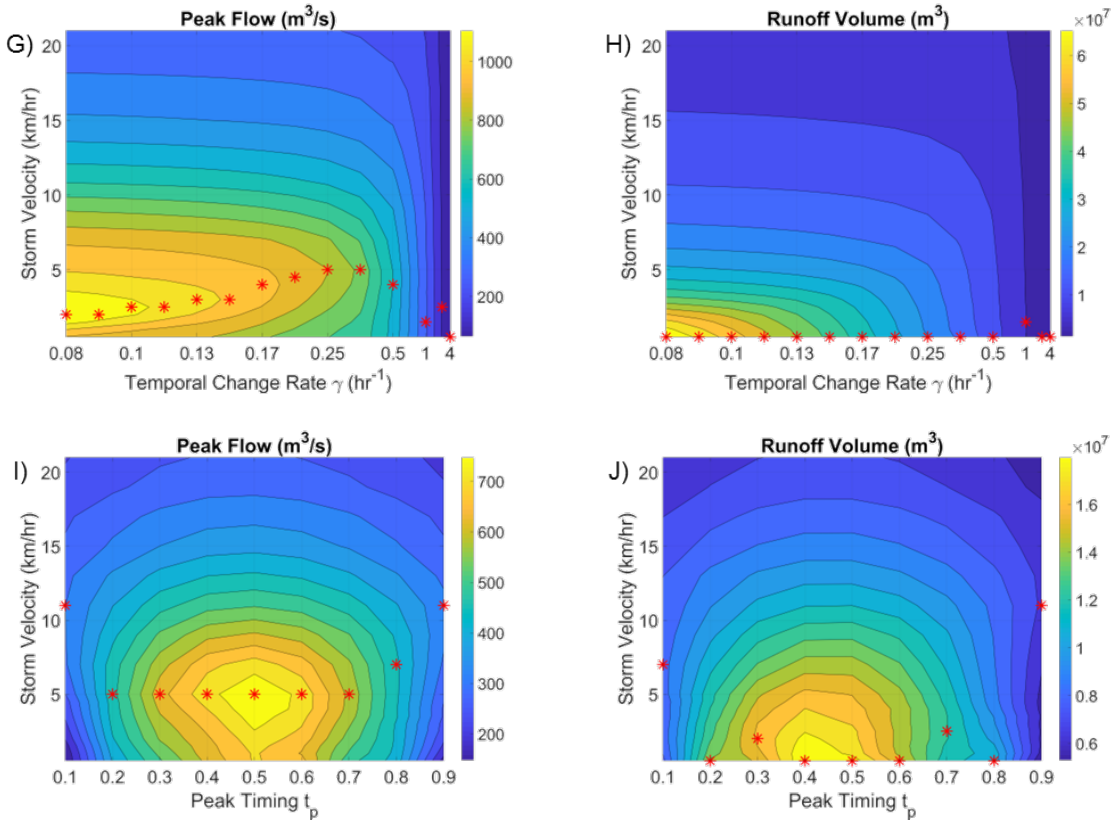



**Figure 8:** Peak flow rates ($Q_p$) and runoff volume ($V_{runoff}$) at watershed outlet from various storm moving velocities ($v$) paired with various storm moving directions (8A and 8B), storm diameters $L_s$ (8C and 8D), spatial shape factors $\alpha$ (8E and 8F), temporal change rate $\gamma$ (8G and 8H), and peak timing $t_p$ (8I and 8J) with the red asterisks marking the $v$ that generates the highest $Q_p$ or $V_{runoff}$.

It is interesting to note the similarities and differences between the PSA results of $Q_p$ and $V_{runoff}$. From any two panels from the same row in **Figure 8**, similarities in general can be traced between the patterns in $Q_p$ and $V_{runoff}$ values. Based on all the values in **Figure 7**, the correlation coefficient between $Q_p$ and $V_{runoff}$ is 0.91.

**Table 2:** Parameter ranges for Pairwise Sensitivity Analysis (PSA)

| Parameters | Minimum | Maximum | Default Value |
|---|---|---|---|
| **Moving Velocity $v$ (km/h)** | 0.5 | 20 | 5 |
| **Moving Direction $\theta$ (°)** | 0 | 360 | 5 |
| **Size/Diameter $L_s$ (km)** | 5 | 30 | 10 (30 *) |
| **Spatial Shape Factor $\alpha$** | 0 | 3 | 0 |
| **Temporal Change Rate $\gamma$ ($h^{-1}$)** | 0.083 | 4 | 0.33 |
| **Peak Timing of Rainfall Intensity $t_p$** | 0.1 | 0.9 | 0.5 |

* : Storm diameter is set to 30 for the PSA of storm moving velocity ($v$) and spatial shape factor ($\alpha$)

The varying patterns of $V_{runoff}$ in **Figure 8** are discussed in detail as follows. With varying storm moving velocity values, the traveling duration ($T$) of the moving storm changes. The shortened $T$ due to increasing $v$ further causes $V_{runoff}$ to reduce as shown in all the panels of **Figure 8**. Regarding storm moving directions, they cause $V_{runoff}$ to vary due to the overlay between the areal extents of the storm path and the Brays Bayou watershed. Because of the elongated shape of Brays Bayou, $V_{runoff}$ values become larger when the storm moves eastward ($\theta = 0^\circ$) or westward ($\theta = 180^\circ$) and smaller when storms moves towards the north ($\theta = 90^\circ$) or the south ($\theta = -90^\circ$), as shown in **Figure 8B**. Besides $\theta$, the storm diameter $L_s$ is another parameter affecting $V_{runoff}$ via varied coverages of the storm over the watershed — a larger $L_s$ value means a greater storm coverage thus a larger runoff volume (**Figure 8D**). Unlike $L_s$, the spatial shape factor $\alpha$ does not affect the storm coverage but the amount of rainfall carried by the storm: a smaller $\alpha$ indicates a milder areal decay rate of rainfall intensity and thus a greater volume of rainfall, as shown in **Figure 8F**. Similar to the storm moving velocity $v$, the temporal change rate $\gamma$ dictates the duration of catchment rainfall ($D$) : rapid rates of change, or greater $\gamma$ values, generate shorter $D$ and thus smaller $V_{runoff}$, as shown in **Figure 8H**. It should be

noted that X-axes of **Figures 8G** and **8H** are plotted in log-scale because $Q_p$ and $V_{runoff}$ values from various temporal change rates ($\gamma$) are positively skewed. Finally, $t_p$ along with storm movement determines spatially where the most intense portion of rainfall (the storm core) occurs relative to the watershed outlet: small $t_p$ values combined with the default westward storm movement ($\theta = 5°$) cause the storm core to be centered at headwaters while rainfall is concentrated near the watershed outlet with large $t_p$ values. Such effects from $t_p$ and the oval shape of Brays Bayou jointly determine that catchment rainfall or $V_{runoff}$ is maximized when the storm core occurs at midstream or $t_p$ is close to 0.5, as shown in **Figure 8J**.

Despite their similarities, $Q_p$ and $V_{runoff}$ show evident discrepancies in the patterns along the Y-axis from varying storm moving velocities. As indicated by the red asterisks in all the panels of **Figure 8**, the storm velocity of the highest $Q_p$ (1st column of the panels) is generally greater than that of the highest $V_{runoff}$ (2nd column of the panels) except for the scenarios where storms travel along other directions than along the streamflow ($\theta$ near 0) in **Figures 8A** and **8B**. According to previous studies, the peak discharge is attributed not only to rainfall volume received by catchments, but also spatial concentration of catchment rainfall (Nicotina et al., 2008; Volpi et al., 2013). A higher spatial concentration of catchment rainfall essentially enhances the temporal superposition of runoff contributing to the watershed outlet, which can be effectively indicated by a smaller standard deviation of runoff timings, $\delta_T$. This study benefits from employing $\delta_T$ as a hydrologic metric since it indicates peak flow magnification independently from $V_{runoff}$. In the following sections, the authors investigate the intricacy of $\delta_T$ to gain deeper insights on the hydrologic responses with respect to each of the storm parameters including storm moving direction ($\theta$), storm diameter ($L_s$), spatial shape factor ($\alpha$), temporal change rate ($\gamma$), and peak timing of rainfall intensity ($t_p$).

*3.2.1 Storm Movement – Storm Moving Direction*

**Figure 9A** shows the range, interquartile range, and median value of $\delta_T$ from storm moving directions ($\theta$) between -60° and 60° for various storm moving velocities ($v$). The selection of $\theta$ values aims to focus on the storms directions with positive downstream components. It is found that as $v$ increases, $\delta_T$ first decreases, gets minimized at $v \approx 5$ km/hr, and then increases. The authors attribute the change of $\delta_T$ with increasing $v$ to two counteracting mechanisms as discussed below. First, the temporal dispersion of hydrographs ($\delta_T$) reach minimum when storms travel downstream along the main channel at a certain speed. This phenomenon is the so-called "resonance condition" in literature (Surkan, 1974; Seo et al., 2012; Volpi et al., 2013). In essence, the "resonance condition" is caused by an exact superposition of the excess rainfall in time (Seo et al., 2012). Under resonance conditions, not only is the temporal dispersion of hydrographs reduced as shown in **Figure 9A,** but also the flood peak is magnified as indicated by the red asterisks in **Figure 8A**. The resonance condition dictates the initial decrease of $\delta_T$ with increasing $v$. Second, increasing $v$ reduces the spatial concentration of catchment rainfall and further increases the dispersion in arrival times of runoff (larger $\delta_T$) at the watershed outlet. Such effect becomes evident when $v$ exceeds the resonance value, as indicated by the increase of $\delta_T$ in **Figure 9A**.

In order to illustrate the directional behavior of hydrologic responses, the maximal directional bias introduced by Niemczynwicz (1984b) is utilized as shown by **Equation 9**.

$$\sigma_{max} = \frac{M_{DOWN} - M_{UP}}{M_{UP}} \quad (9)$$

where $\sigma_{max}$ is the maximal directional bias of a hydrologic metric ($\delta_T$ herein); $M_{DOWN}$ and $M_{UP}$ are the hydrologic metrics for storms moving downstream and upstream respectively. **Figure 9B**

shows two boxplots of $\sigma_{max}$ of $Q_p$ and $\delta_T$ respectively, for storm moving directions from -60° to 60°. First, it can be found that all $\sigma_{max}$ of $Q_p$ values are positive while all $\sigma_{max}$ of $\delta_T$ are negative, meaning that peak flow is higher and the temporal dispersion of hydrographs is lower when storms moving downstream than upstream. This explains the differences of $Q_p$ and $V_{runoff}$ values regarding storm moving directions (along X-axis) in **Figures 8A** and **8B** — the downstream-moving storms generate higher $Q_p$ but similar $V_{runoff}$ than the upstream-moving storms. Second, the varying patterns of $\sigma_{max}$ of $Q_p$ and $\sigma_{max}$ of $\delta_T$ in **Figure 9B** are almost symmetrical, which means $\delta_T$ is very indicative of $Q_p$ in this case. In addition, the peak $\sigma_{max}$ of $Q_p$ and the lowest $\sigma_{max}$ of $\delta_T$ coincide when storm moves downstream approximately along the stream flow direction of Brays Bayou ($\theta$ = -15° ), which confirms the finding by Niemczynwicz (1984b) the peak $\sigma_{max}$ of $Q_p$ is reached when a storm moves in the same direction as the channel flow.

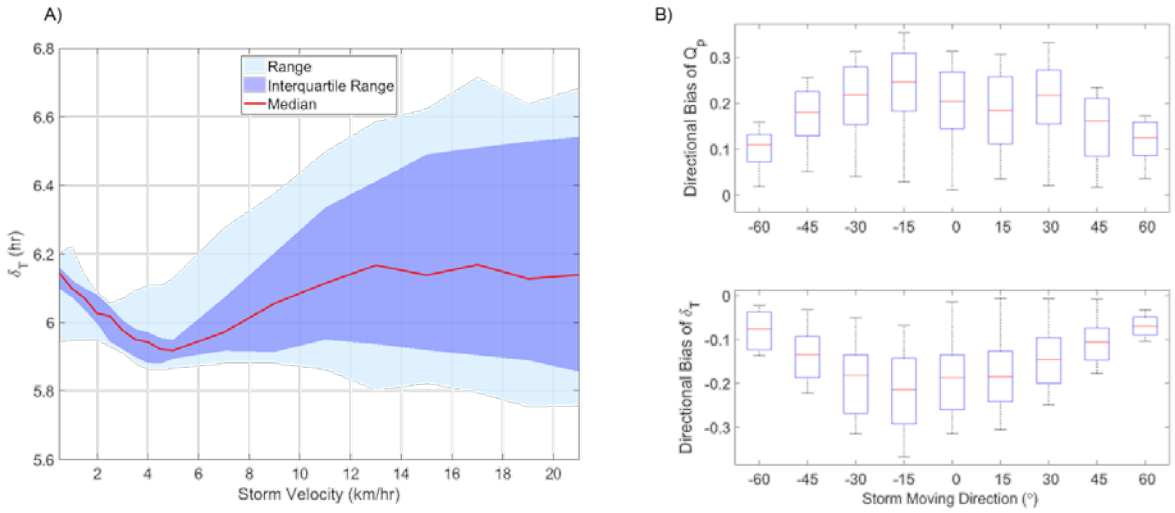



**Figure 9:** Standard deviation of runoff timings ($\delta_T$) illustrated by (A) the range, interquartile range, and median values from storm moving directions ($\theta$) between -60° and 60° for various

storm moving velocities ($v$) and (B) boxplots of $\sigma_{max}$ of $Q_p$ and $\delta_T$ from various storm moving velocities ($v$) for storm moving directions ($\theta$) between -60° and 60°.

*3.2.2 Spatial Variability – Storm Diameter $L_s$ and Spatial Shape Factor $\alpha$*

**Figures 10A** and **10B** show the contour plots of $\delta_T$ from various storm moving velocities ($v$) versus storm diameters ($L_s$) and spatial shape factors ($\alpha$), respectively; for each X value ($L_s$ or $\alpha$), the red circles marks the $v$ value generating the lowest $\delta_T$. First, it can be found that the patterns of $\delta_T$ in **Figures 10A** and **10B** are almost symmetrical along the X direction. This indicates that the increasing $L_s$ and decreasing $\alpha$ in their own varying ranges have similar effects on $\delta_T$. It can be found that either a smaller storm diameter or a steeper areal decay rate of rainfall intensity can reduce $\delta_T$ value. This is due to the increased spatial concentration of catchment rainfall further decreasing the temporal dispersion of runoff arriving at the watershed outlet. Second, $L_s$ and $\alpha$ interact with storm moving velocity $v$ in reaching the aforementioned resonance condition. As indicated by the red circles, the $v$ values associated with resonance conditions increase with greater $L_s$ or lower $\alpha$ values because a faster storm moving velocity is required to offset the effect from the reduced spatial concentration of catchment rainfall and reach the exact superposition of flow at the watershed outlet. Additionally, the $v$ values generating the lowest $\delta_T$ (the red circles in **Figures 10A** and **10B**) do not necessarily coincide with those generating the highest $Q_p$ (the red asterisks in **Figures 8C** and **8E**), especially for higher $L_s$ or lower $\alpha$ values. This is due to the increased $V_{runoff}$ (from the higher $L_s$ or lower $\alpha$ values) overwhelming the decreased spatial concentration of catchment rainfall in terms of affecting the peak flow.

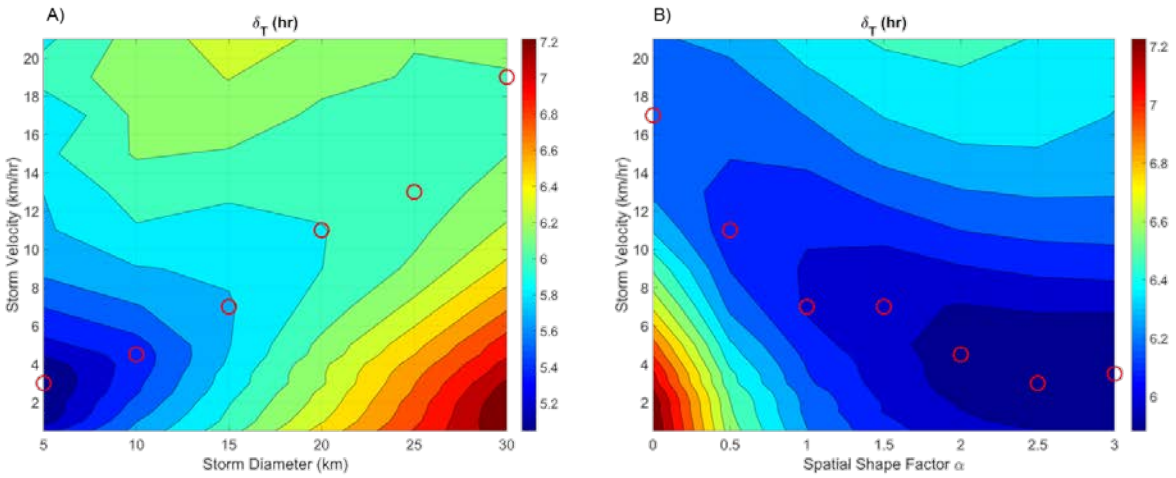



**Figure 10:** Standard deviations of runoff timings ($\delta_T$) from various storm moving velocities ($v$) versus various (A) storm diameters ($L_s$) and (B) spatial shape factors ($\alpha$) with red circles marking the $v$ generating the lowest $\delta_T$

*3.2.3 Temporal Variability –Temporal Change Rate $\gamma$*

**Figure 11A** shows the range, interquartile range, and median value of $\delta_T$ from various temporal change rates ( $\gamma$ ) for various storm moving velocities ($v$). It is shown that $\delta_T$ first decreases rapidly, reaches the minimum at $v \approx 5$ km/hr, and then slowly increases with greater $v$ values, which corresponds to the variation of peak flow as shown in **Figure 8G**. This is also due to the resonance condition mentioned before. It should also be noted that the interquartile range is very close to the top of the range because $\delta_T$ values from various temporal change rates ($\gamma$) are positively skewed, analogous to $Q_p$ and $V_{runoff}$ values in **Figures 8G and 8H**.

Similar to **Figure 11A**, **Figure 11B** shows the same statistics (range, interquartile range, and median) of $\delta_T$ but from various $v$ values for $\gamma$ values plotted in log scale. First, with greater $\gamma$, $\delta_T$ is found to decrease because a rapid temporal change of rainfall intensity along with the storm movement increases the spatial concentration of catchment rainfall and further reduces the

temporal dispersion of hydrographs. It is interesting to note that change of $\gamma$ causes two counteracting effects on peak discharge. On one hand, increasing $\gamma$ values lead to smaller $\delta_T$ as shown in **Figure 11B**, indicating a more spiky shape of the hydrograph. On the other hand, increasing $\gamma$ values shorten the rainfall duration in the watershed and further reduce $V_{runoff}$ as shown in **Figure 8H.** The second effect appears to be stronger than the first based on the pattern of $Q_p$ shown in **Figure 8G**.

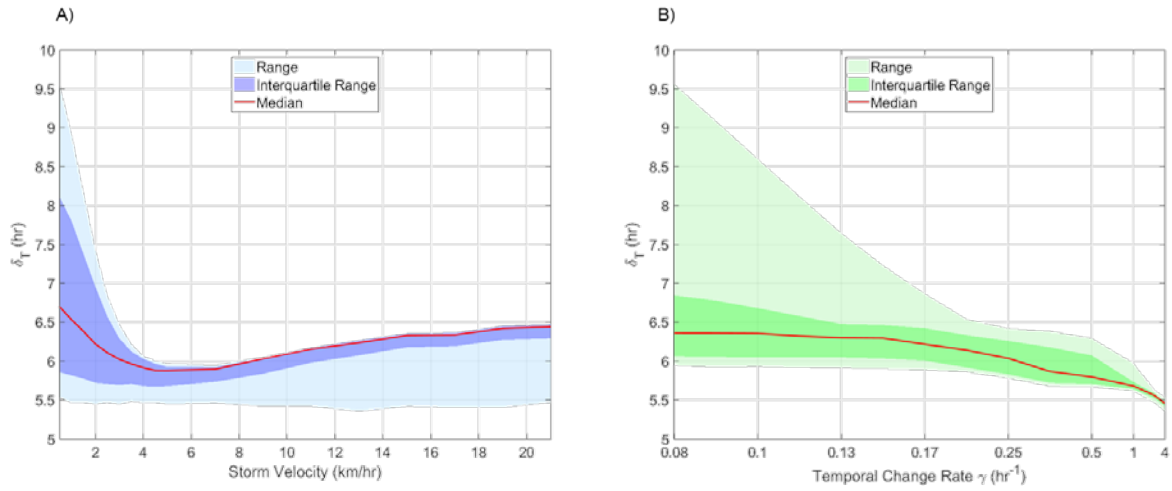



**Figure 11:** Standard deviations of runoff timings ($\delta_T$) illustrated by the range, interquartile range, and median values (A) from various temporal change rates ($\gamma$) for various storm moving velocities ($v$) and (B) from various storm moving velocities ($v$) for various temporal change rates ($\gamma$).

*3.2.4 Temporal Variability –Peak Timing of Rainfall Intensity $t_p$*

**Figure 12A** and **12B** show the range, interquartile range, and median value of $\delta_T$ from various peak timings of rainfall intensity ($t_p$) for various storm moving velocities ($v$) and from various $v$ for various $t_p$ values, respectively. The varying trend of $\delta_T$ with increasing $v$ in **Figure 12A** reaffirms the resonance conditions found in the previous PSA sets. It can also be found in

**Figure 12B** that $\delta_T$ deceases rapidly with increasing $t_p$ from 0.1 to 0.3, then slowly decreases till $t_p$ = 0.7, and then slightly increases till $t_p$ = 0.9. The authors attribute the varying pattern of $\delta_T$ with $t_p$ to two main factors as discussed below. First, varying $t_p$ values along with the storm movement change the location of storm core relative to the channel network. **Figure 12** shows the dimensionless channel lengths from the storm core to the watershed outlet ($\Delta$ in **Equation 3**) calculated for various combinations of $t_p$ and $v$. When $v$ is small, increasing $t_p$ values cause the storm core to shift from the headwaters (higher $\Delta$) to near the watershed outlet (lower $\Delta$). Consequently, the center mass of rainfall excess goes through shorter main channel lengths and thus receives less attenuation as $t_p$ increases. Such effect dictates the initial decrease of $\delta_T$ with increasing $t_p$ as shown **in Figure 12B**. It can also be found in **Figure 13** that $t_p$ creates less variation in $\Delta$ with greater $v$ (green zone) due to the increasingly uniform spatial distribution of catchment rainfall. Gao and Fang (2018) also utilized $\Delta$ as an indicator to demonstrate that the peak discharge is inversely correlated to $\Delta$ in Brays Bayou and such relationship is stronger with greater spatial concertation of catchment rainfall. Second, besides the channel attenuation effect, the change of $t_p$ shifts the balance between the overland and channel routing. As $t_p$ increases and channel routing lessens (indicated by $\Delta$), overland routing constitutes a growing portion in the overall rainfall-runoff process. As a result, the hydrograph at the watershed outlet resembles characteristics of the unit hydrographs used in the overland routing approach, which features slower flow velocity and larger $\delta_T$. As shown in **Figure 12B**, the second factor becomes more pronounced in the variation of $\delta_T$ when $t_p$ exceeds approximately 0.3.

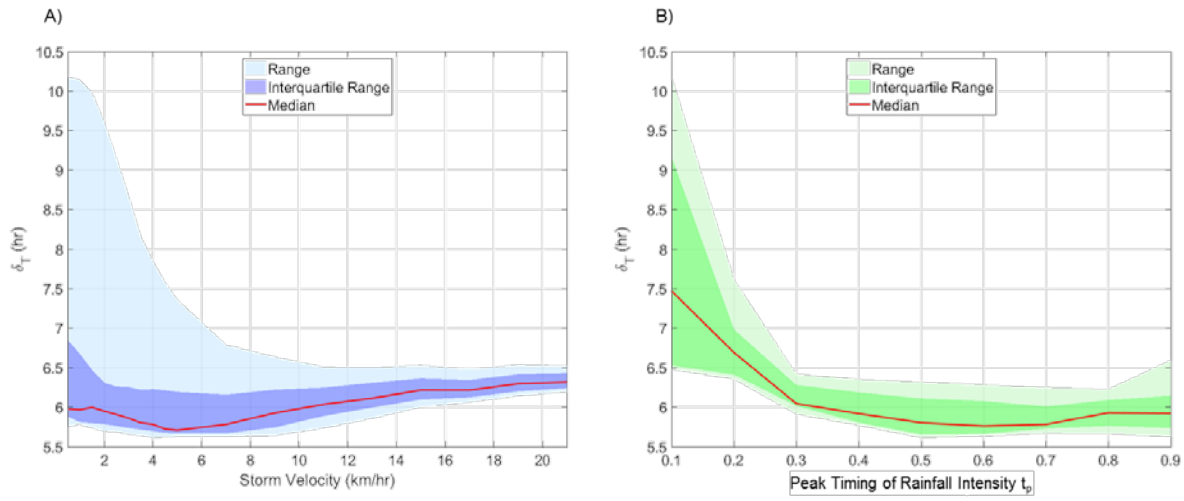



**Figure 12:** Standard deviations of runoff timings ($\delta_T$) illustrated by the range, interquartile range, and median values (A) from various peak timings of rainfall intensity ($t_p$) for various storm moving velocities ($v$) and (B) from various storm moving velocities ($v$) for various peak timings of rainfall intensity ($t_p$).

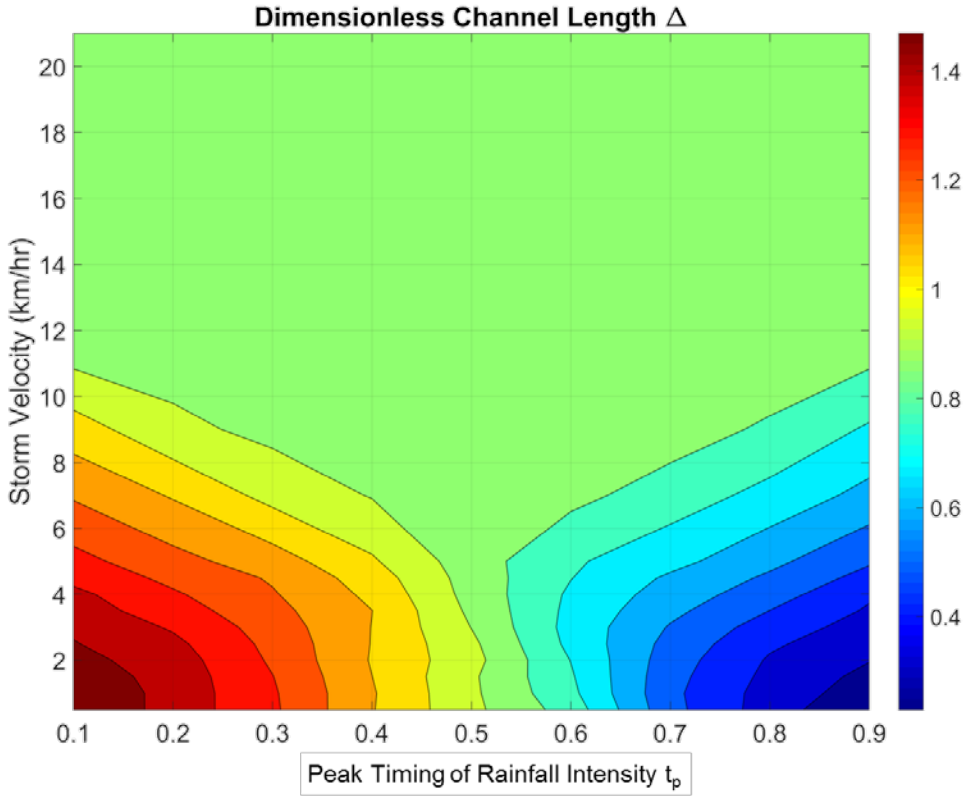


**Figure 13:** Dimensionless channel lengths from various storm moving velocities ($v$) versus various peak timings of rainfall intensity ($t_p$).

### *3.3 Discussion*

The results from GSA suggest that parameter interactions contribute more to the variation of peak discharge than individual parameters. In particular, contributions of $\gamma$ and $t_p$ coming almost entirely from parameter interactions suggest that the seemingly minor factors could lead to great hydrologic effects through interacting with other factors. Therefore, it is important for synthetic storms to sufficiently reflect and not oversimplify the temporal variability of rainfall for hydrologic simulations.

Results from PSA indicate that peak discharge is largely driven by runoff-volume or rainfall amount in the highly impervious Brays Bayou watershed. Many historical storm events occurring in the Houston region have exhibited this feature. For instance, the tremendous rainfall amount of Hurricane Harvey in 2017 caused record-breaking peak discharges in 9 out the 19 gauged rivers in Harris County (Blake and Zelinsky, 2018). Gao and Fang (2018) also reported a high correlation between $Q_p$ and $V_{runoff}$ in Brays Bayou based on numerous rainfall-runoff scenarios created by storm transposition. The volume-driven peak discharges are also due to the flat terrain of Brays Bayou which inherently has attenuation effects on the flood peaks. This finding is opposite to the case in a steep mountainous region where peak rainfall intensity is likely the prevailing factor. Reflecting the attenuation effects of various reach sections suggests space- and time- variant flow velocities in channel routing. Therefore, the simplified assumption of time-invariant flow velocities (made in previous studies) is not suitable for flat areas like Brays Bayou. Furthermore, it is worth noting the differences between the resonance conditions in this study and those in previous studies (Surkan, 1974; Seo et al., 2012; Volpi et al., 2013). The

temporal superposition reached under the assumption of time-invariant flow velocities is expected to be higher than the more realistic one reached in this study. In other words, resonance effect occurring in Brays Bayou would be overestimated if time-variant flow velocities were neglected.

Due to the 100% imperviousness assumed in this study, the storm core location does not affect any spatial distribution of infiltration processes to further result in varied hydrologic responses. Although this assumption is adequate for urban watersheds like Brays Bayou, it might be proved invalid elsewhere. For example, Syed et al. (2003) investigated the pronounced channel loss in arid watersheds focusing on its dependency on the relative location of storm core in the channel network. Therefore as a factor in runoff generation process, spatio-temporal variability of rainfall is more complex to natural watersheds than urban watersheds. If the findings from this study were to be extended to natural watersheds, the hydrologic responses would have to be associated with excess rainfall, which requires accurate estimation of losses due to soil moisture deficit created by evapotranspiration before each storm event.

## 4. CONCLUSIONS AND FUTURE WORK

This paper demonstrates a synthetic storm generator (DMS) that encompasses spatial and temporal variabilities of rain intensity as well as storm movement. Output from the DMS generator can be further connected to lumped or distributed hydrologic models to perform rainfall-runoff analysis. The hydrologic responses to the storm parameters are investigated by conducting global and pairwise sensitivity analyses on the DMS parameters. In particular, the DMS generator includes parameters controlling the temporal variability of rainfall intensity which were rarely available in previous storm generators. Moreover, the authors account for flow attenuation when analyzing the hydrologic responses to storm parameters, which was often

neglected in previous studies. Consequently, the insights gained from the analyses can better reveal the complex nature of rainfall-runoff process.

To decipher various hydrologic responses to storm parameters, runoff volume ($V_{runoff}$) and standard deviation of runoff timings ($\delta_T$) are utilized as independent metrics of hydrographs to reveal different aspects of the rainfall-runoff mechanism and dissect the intricacy in peak discharges. Dimensionless channel length ($\Delta$) is employed to characterize the spatial organization of catchment rainfall in the channel network. In addition, the DMS framework enables us to incorporate spatio-temporal variables into synthetic storms and reveal the corresponding hydrologic responses. The well-calibrated hydrologic model for Brays Bayou is used to accurately reveal the hydrologic responses to moving storms without compromising the flow attenuation effects from channel routing. The sensitivity analyses are conducted to investigate the effects on peak discharge from both the individual parameters and their interdependences. Understanding how storm characteristics synergize or counteract with each other to affect flood peaks is valuable for identifying the weakness in flood protection against certain type of storms. Such insights also help emergency responders for better preparedness against flood hazards. The findings obtained in this study about the hydrologic responses from Brays Bayou are summarized as follows:

1) According to the global sensitivity analysis, storm moving velocity significantly affects peak discharge not only individually but also through interacting with other storm characteristics.
2) Peak discharges are magnified when storms travel downstream along the main channel at the speed that corresponds to a temporal superposition of runoff (indicated by $\delta_T$) arriving at the watershed outlet.

3) Spatial and temporal variabilities of rainfall affect peak discharges via changing the spatial concentration of catchment rainfall and further the temporal dispersion of hydrographs at the watershed outlet (indicated by $\delta_T$). Specifically, a smaller storm diameter, a steeper areal decay rate of rainfall intensity, and a rapider temporal change rate of rainfall intensity all increase the spatial concentration of catchment rainfall and reduce the temporal dispersion of hydrographs.
4) The peak timing of rainfall intensity ($t_p$) along with storm movement alters spatially where the most intense portion of rainfall (the storm core) occurs in the channel network, which further changes the flow attenuation effects from channel routing. A storm core at headwaters leads to the most attenuation from channels, while rainfall centered near the watershed outlet generates runoff with nearly no attenuation from channels.

For future directions of this study, rainfall frequency analysis can be conducted if the DMS generator is embedded in a stochastic framework (see e.g., Wright et al 2013; Zhou et al., 2019). Additional stochastic processes are required to simulate how storm cells cluster in space and how individual storm cell starts and ends in time, which can be similar to other developed storm generators based on stochastic point process (see e.g., Cowpertwait et al., 2002; Burton et al., 2008; Leonard et al., 2008; Cowpertwait, 2010; Beuchat et al., 2011). A prerequisite for the stochastic DMS generator is to estimate the probability distribution of storm parameters. Ongoing effort is invested in retrieving parameters from radar rainfall data (see e.g., Peleg and Morin, 2012; Peleg and Morin, 2014). Once the stochastic DMS generator successfully reproduces key statistics from the historical rainfall, it will be further used in flood frequency analysis for ungauged watersheds by connecting to a hydrologic model for a long-term simulation (see e.g., Wright et al., 2014). The authors will determine appropriate hydrological

models for ungauged watersheds by taking into account the following factors: 1) the model, whether being distributed or lumped, should adequately reflect the effects of spatially distributed rainfall fields; 2) without stream gauges for calibration, uncertainty of model parameters should be reduced via regionalization (see e.g., Samaniego et al., 2010); and 3) the model should be able to capture the temporal variation of model states (e.g., antecedent soil moisture) via continuous simulations.

## 5. ACKNOWLEDGEMENTS

The authors would like to thank the West Gulf River Forecast Center (WGRFC) and the United States Army Corps of Engineers (USACE) for providing technical inputs. The authors would also like to thank Dr. Gianluca Botter and the four anonymous reviewers for their valuable comments and suggestions.

## 6. DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## 7. REFERENCES

Bass, B., Juan, A., Gori, A., Fang, Z. and Bedient, P. (2016). 2015 Memorial Day Flood Impacts for Changing Watershed Conditions in Houston. *Natural Hazards Review*, *18*(3), 05016007. DOI: 10.1061/(ASCE)NH.1527-6996.0000241

Bedient, P.B, Huber, W.C., and Vieux B.E. (2018). *Hydrology and Floodplain Analysis*, 6th Ed. Prentice-Hall, Upper Saddle River NJ: 763

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

Beuchat, X., Schaefli, B., Soutter, M., & Mermoud, A. (2011). Toward a robust method for subdaily rainfall downscaling from daily data. *Water Resources Research*, 47(9).

Blake E. S., and Zelinsky D. A. (2018). Hurricane Harvey - National Hurricane Center - NOAA. https://www.nhc.noaa.gov/data/tcr/AL092017_Harvey.pdf

Borga, M., Boscolo, P., Zanon, F., and Sangati, M. (2007) Hydrometeorological analysis of the 29 August 2003 flash flood in the Eastern Italian Alps. *J. of Hydromet.* 8(5), 1049-1067. DOI: https://doi.org/10.1175/JHM593.1

Burton, A., Kilsby, C. G., Fowler, H. J., Cowpertwait, P. S. P., & O'connell, P. E. (2008). RainSim: A spatial–temporal stochastic rainfall modelling system. *Environmental Modelling & Software*, 23(12), 1356-1369.

Cowpertwait, P. S. P., Kilsby, C. G., & O'Connell, P. E. (2002). A space‐time Neyman‐Scott model of rainfall: Empirical analysis of extremes. *Water Resources Research*, 38(8), 6-1.

Cowpertwait, P. S. (2010). A spatial‐temporal point process model with a continuous distribution of storm types. *Water Resources Research*, 46(12).

Curtis, D. (2007). Evaluation of the Spatial Structure of Storms and the Development of Design Storms. Proceedings of the ASCE/EWRI World Environmental and Water Resources

Congress 2007: Restoring Our Natural Habitat, May 15 2007. DOI: https://ascelibrary.org/doi/abs/10.1061/40927(243)286

De Lima, J.L.M.P., Singh, V.P. (2002). The influence of the pattern of moving rainstorms on overland flow. *Adv. Water Resour.* 25(7), 817–828. DOI: https://doi.org/10.1016/S0309-1708(02)00067-2

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

Fang, Z., Bedient, P., and Buzcu-Guven, B. (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. DOI: 10.1061/(ASCE)HE.1943-5584.0000374

Fang, Z., Shultz, M., Wienhold, K., Zhang, J., and Gao, S. (2019). Case Study: Comparative Analysis of Hydrologic Simulations with Areal-Averaging of Moving Rainfall. *Hydrology* 6, 12. DOI:10.3390/hydrology6010012

Feral, L., Sauvageot, H., Castanet, L., and Lemorton, J. (2003). HYCELL - A new hybrid model of the rain horizontal distribution for propagation studies: 1. Modeling of the rain cell. *Radio Sci.* 38(3), 18. DOI: https://doi.org/10.1029/2002RS002802

Furl, C., Sharif, H.O., El Hassan, A., Mazari, N., Burtch, D., and Mullendore, G.L. (2015) Hydrometeorological Analysis of Tropical Storm Hermine and Central Texas Flash Flooding, September 2010. *J. of Hydromet.* 16(6), 2311-2327. DOI: https://doi.org/10.1175/JHM-D-14-0146.1

Gao, S. and Fang, Z. (2018) Using Storm Transposition to Investigate the Relationships between Hydrologic Responses and Spatial Moments of Catchment Rainfall. *Natural Hazards Review*, 19(4), 04018015. DOI: 10.1061/(ASCE)NH.1527-6996.0000304

Gupta, V. K., Castro, S. L., & Over, T. M. (1996). On scaling exponents of spatial peak flows from rainfall and river network geometry. Journal of hydrology, 187(1-2), 81-104.

Homma, T., & Saltelli, A. (1995). Use of Sobol's quasirandom sequence generator for integration of modified uncertainty importance measure. Journal of Nuclear Science and Technology, 32(11), 1164-1173.

Hydrologic Engineering Center (HEC). (2006). HEC-HMS, Hydrologic Modeling System. US Army Corps of Engineers, Davis, Calif

Lee, K.T., Huang, J.K. (2007). Effect of moving storms on attainment of equilibrium discharge. *Hydrol. Process* 21(24), 3357–3366. DOI: https://doi.org/10.1002/hyp.6548

Leonard, M., Lambert, M. F., Metcalfe, A. V., & Cowpertwait, P. S. P. (2008). A space‐time Neyman–Scott rainfall model with defined storm extent. *Water Resources Research*, 44(9).

Liang, J. (2010). *Evaluation of Runoff Response to Moving Rainstorms*. PhD dissertation, Marquette Univ, Milwaukee, Wisc.

Mandapaka, P. V., Krajewski, W. F., Mantilla, R., & Gupta, V. K. (2009). Dissecting the effect of rainfall variability on the statistical structure of peak flows. Advances in Water Resources, 32(10), 1508-1525.

Marshall, R.J. (1980) The estimation and distribution of storm movement and storm structure, using a correlation-analysis technique and rain gauge data. *J. Hydrol.* 48(1–2), 19–39. DOI: https://doi.org/10.1016/0022-1694(80)90063-3

Morin, E., Goodrich, D.C., Maddox, R.A., Gao, X., Gupta, H.V., and Sorooshian, S. (2006). Spatial patterns in thunderstorm rainfall events and their coupling with watershed hydrological response. *Ad. in Water Resour.* 29(6), 843-860. DOI: https://doi.org/10.1016/j.advwatres.2005.07.014

Michaud, J.D., Sorooshian, S. (1994). Effect of rainfall-sampling errors on simulations of desert flash floods. *Water Resour. Res.* 30(10), 2765–2775. DOI: https://doi.org/10.1029/94WR01273

Niemczynowicz, J. (1984a). Investigation of the influence of rainfall movement on runoff hydrograph, Part I— Simulation on conceptual catchment. *Nord. Hydrol.* 15, 57–70. DOI: https://doi.org/10.2166/nh.1984.0005

Niemczynowicz, J. (1984b). Investigation of the influence of rainfall movement on runoff hydrograph, Part II—Simulation on real catchments in the City of Lund. *Nord. Hydrol.* 15, 71-84. DOI: https://doi.org/10.2166/nh.1984.0006

National Weather Service (2017, April 21). *Noaa Atlas 14 Point Precipitation Frequency Estimates: TX*. Retrieved from: https://hdsc.nws.noaa.gov/hdsc/pfds/pfds_map_cont.html?bkmrk=tx

Nicótina, L., Alessi Celegon, E., Rinaldo, A., and Marani, M. (2008). On the impact of rainfall patterns on the hydrologic response. *Water Resour. Res.* 44 W12401. DOI: http://dx.doi.org/10.1029/2007WR006654

Ogden, F.L., Richardson, J.R., and Julien, P.Y. (1995). Similarity in catchment response: 2. Moving rainstorms. *Water Resour. Res.* 31, 1543-1547. DOI: https://doi.org/10.1029/95WR00519

Olivera, F., Choi, J., Kim, D., & Li, M. H. (2008). Estimation of average rainfall areal reduction factors in Texas using NEXRAD data. *Journal of Hydrologic Engineering*, 13(6), 438-448.

Peleg, N., & E. Morin (2012), Convective rain cells: Radar-derived spatiotemporal characteristics and synoptic patterns over the eastern Mediterranean, *J. Geophys. Res.*, 117, D15116, doi:10.1029/2011JD017353.

Peleg, N., Morin, E. (2014). Stochastic convective rain-field simulation using a high-resolution synoptically conditioned weather generator (HiReS-WG). *Water Resour. Res.* 50, 2124–2139. DOI: https://doi.org/10.1002/2013WR014836

Richardson, J.R., Julien, P.Y. (1989). *One-dimensional Modeling of Moving Rainstorms, in Catchment Runoff and Rational Formula*, Yen, BC, ed, Water Resources Publications Littleton, CO, 155-167

Rozalis, S., Morin, E., Yair, Y., and Price, C. (2010). Flash flood prediction using an uncalibrated hydrological model and radar rainfall data in a Mediterranean watershed under changing hydrological conditions. *J. of Hydrol.* 394(1), 245-255. DOI: https://doi.org/10.1016/j.jhydrol.2010.03.021

Samaniego, L., Bárdossy, A., & Kumar, R. (2010). Streamflow prediction in ungauged catchments using copula‐based dissimilarity measures. *Water Resources Research*, *46*(2).

Seo, Y., Schmidt, A.R., and Sivapalan, M. (2012). Effect of storm movement on flood peaks: Analysis framework based on characteristic timescales. *Water Resour. Res.* 48. DOI: https://doi.org/10.1029/2011WR011761

Segond, M.L., Wheater, H.S., and Onof, C. (2007). The significance of spatial rainfall representation for flood runoff estimation: A numerical evaluation based on the Lee catchment, UK. *J. of Hydrol.* 347(1), 116-131. DOI: https://doi.org/10.1016/j.jhydrol.2007.09.040

Singh, V.P. (1998). Effect of the direction of storm movement on planar flow. *Hydrol. Proc.* 12: 147-170. DOI: https://doi.org/10.1002/(SICI)1099-1085(199801)12:1<147::AID-HYP568>3.0.CO;2-K

Smith, J. A., Baeck, M. L., Morrison, J. E., Sturdevant-Rees, P., Turner-Gillespie, D. F., & Bates, P. D. (2002). The regional hydrology of extreme floods in an urbanizing drainage basin. *Journal of Hydrometeorology*, 3(3), 267-282.

Smith, J. A., Baeck, M. L., Meierdiercks, K. L., Nelson, P. A., Miller, A. J., & Holland, E. J. (2005a). Field studies of the storm event hydrologic response in an urbanizing watershed. Water Resources Research, 41(10).

Smith, J.A., Miller, A.J., Baeck, M.L., Nelson, P.A., Fisher, G.T., and Meierdiercks, K.L. (2005b). Extraordinary flood response of a small urban watershed to short-duration convective rainfall. *J. of Hydromet.* 6(5), 599-617. DOI: https://doi.org/10.1175/JHM426.1

Sobol', I.M. (1993). Sensitivity estimates for nonlinear mathematical models. *Math Model Comput. Exp.* 1(4), 407 -417

Sobol', I. M. (2001). Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. Mathematics and computers in simulation, 55(1-3), 271-280.

Stephenson, D. (1984). Kinematic study of effects of storm dynamics on runoff hydrographs. *Water SA* 10, 189-196

Surkan, A.J. (1974). Simulation of storm velocity effects on flow from distributed channel networks. *Water Resour. Res.* 10(6), 1149–1160. DOI: https://doi.org/10.1029/WR010i006p01149

Syed, K.H., Goodrich, D.C., and Myers, D.E. (2003). Spatial characteristics of thunderstorm rainfall fields and their relation to runoff. *J. of Hydrol.* 271, 1–21. DOI: https://doi.org/10.1016/S0022-1694(02)00311-6

Tarolli, M., Borga, M., Zoccatelli, D., Bernhofer, C., Jatho, N., and Janabi, F.A. (2013). Rainfall space-time organization and orographic control on flash flood response: The Weisseritz Event of August 13, 2002. *J. of Hydrol. Eng.* 18(2), 183-193. DOI: 10.1061/(ASCE)HE.1943-5584.0000569.

Viglione, A., Chirico, G.B., Woods, R., and Blöschl, G. (2010). Generalised synthesis of space-time variability in flood response: An analytical framework. *J. Hydrol.* 394(1–2), 198–212. DOI: https://doi.org/10.1016/j.jhydrol.2010.05.047

Volpi, E., Di Lazzaro, M., & Fiori, A. (2013). Analytical modeling of the hydrologic response under moving rainstorms: Storm–catchment interaction and resonance. Journal of hydrology, 493, 132-139.

von Hardenberg, J., Ferraris, L,, and Provenzale, A. (2003). The shape of convective rain cells. *Geophys, Res. Lett.* 30(24). DOI: https://doi.org/10.1029/2003GL018539

Wealands, S.R., Grayson, R.B., Walker, J.P. (2005). Quantitative comparison of spatial fields for hydrological model assessment—some promising approaches. *Adv. Water Resour.* 28, 15–32. DOI: https://doi.org/10.1016/j.advwatres.2004.10.001

Wright, D. B., Smith, J. A., Villarini, G., & Baeck, M. L. (2013). Estimating the frequency of extreme rainfall using weather radar and stochastic storm transposition. *Journal of hydrology*, 488, 150-165.

Wright, D. B., Smith, J. A., & Baeck, M. L. (2014). Flood frequency analysis using radar rainfall fields and stochastic storm transposition. Water Resources Research, 50(2), 1592-1615.

Yen, B.C., Chow, V.T. (1969). A laboratory study of surface runoff due to moving rainstorms. *Water Resour. Res.* 5(5), 989–1006. DOI: https://doi.org/10.1029/WR005i005p00989

Zhou, Z., Long, Y., Liu, S., & Smith, J. (2018). The role of storm scale, position and movement in controlling urban flood response. *Hydrology and Earth System Sciences*, 22(1), 417.

Zhou, Z., Smith, J., Wright, D., Baeck, M., Yang, L., & Liu, S. (2019). Storm Catalog-Based Analysis of Rainfall Heterogeneity and Frequency in a Complex Terrain. *Water Resources Research*, 55(3), 1871 ‐ 1889. https://doi.org/10.1029/2018WR023567

Zoccatelli, D., Borga, M., Viglione, A., Chirico, G. B., & Blöschl, G. (2011). Spatial moments of catchment rainfall: rainfall spatial organization, basin morphology, and flood response. Hydrology and Earth System Sciences, 15(12), 3767-3783.

# Chapter 4:

# Stochastic Storm Simulation Using Optimal Estimation and Non-Parametric Generation of Ensemble Parameters

Shang Gao and Zheng Fang

## ABSTRACT

The historical record of rainfall observation rarely provides sufficient record length and resolution and therefore causes uncertainty in rainfall frequency analysis (RFA). The work described here serves as the first step in applying a long-term stochastic rainfall simulation in RFA. The framework of stochastic simulation adopts optimal estimation for spatio-temporal modeling of rain fields. A non-parametric approach featuring K-Nearest Neighbor Resampling (KNNR) plus the Genetic Algorithm (GA) mixing process is utilized for generating parameters in the long-term simulation. A case study is conducted in Dallas-Fort-Worth metroplex as the simulation domain. Ensemble parameters are generated using the KNNR+GA method from adjacent homogeneous areas and 10 years of radar rainfall observation. One hundred most rainy days in the 10 years are simulated at the resolutions of $4 \times 4$ km$^2$ and 1 hour for 50 ensemble members. The simulated rainfall is thoroughly evaluated against the observed radar rainfall with respects to statistical moments, spatio-temporal structure, and frequency distribution of rainfall at both near-point scale and domain scale. The results indicate that ensemble simulations successfully reproduce key statistical properties of the observed rainfall. In addition, the approach is also effective and flexible in capturing heavy rainfall values, which is important for RFA.

## KEY TERMS

Stochastic Rainfall Simulation, Optimal Estimation, K-Nearest Neighbor Resampling, Genetic Algorithm, Spatio-temporal Modeling of Rain Field, Non-parametric Approach, Rainfall Frequency Analysis

# 1. INTRODUCTION

The role of rainfall is essential for hydrology as it is the driving phenomenon of runoff mechanisms. Rainfall observation is the foundation of many hydrologic applications including flood risk management, design of urban drainage systems and dams (Cunnane, 1988). As the centerpiece of the wide range of applications, rainfall frequency analysis (RFA) reveals the relationship between rainfall magnitude and likelihood. Any RFA becomes more valid with longer record of rainfall observation, since the return period of infrequent severe storms are better captured. In the absence of a long data record, extrapolation is inevitable which adds uncertainty to RFA. Besides record length, another limitation in traditional RFA is caused by the spatially-sparse rain-gauge measurements. Due to the lack of spatial resolution, rain-gauge measurements can easily miss the most intense portion of storm, underestimating rainfall of large return period (e.g. 100 years or 500 years). In addition, a number of simplifying assumptions were made about the structure of extreme rainfall when high-resolution measurement of rainfall was not practical, including spatially and oftentimes temporally uniform ‘design storms’ (e.g. Koutsoyiannis, 1994), and area reduction factors (ARFs, e.g. Svensson and Jones, 2010). Some of these assumptions neglect the variety and complexity of hydro-meteorological processes (Wright et al., 2013b). Although the advent of weather radar estimation tremendously increases the spatiotemporal resolution of rainfall observation, its length of record hasn’t exceeded 30 years in the U.S. (NCEI, 2018). Therefore, lack of resolution and record length in rainfall observation has hindered reliable RFA.

Several stochastic modeling techniques have been developed to overcome the data limitations described above. Stochastic rainfall models are statistical models that can be used as random number generators whose output resembles the rainfall data to which they have been fit

(Woolhiser, 1992). The existing stochastic spatiotemporal modeling approaches can be divided into two major classes. The first approach is based on the theory of stochastic point processes. In this approach, the basic precipitation units called “rain cells” are two-dimensional pulses of positive intensity and duration. The occurrence of rain cells are simulated as independent Poisson processes overlapping in space and time. This approach has become one of the most widely used for rainfall simulation (e.g., Cowpertwait et al., 2002; Burton et al., 2008; Leonard et al., 2008; Cowpertwait, 2010; Beuchat et al., 2011), wherein the rain cells are conceptualized as discs of randomly assigned durations, intensities, and radius and move in constant velocity. The stochastic point processes models allow for analytical derivation of statistical properties across spatiotemporal scales, thus reproduce mean rainfall statistics reasonably well. However, these models often fail to capture complex spatial structure and kinematic properties of the rain cells. Moreover, representation of more realistic rainfall patterns requires a large number of parameters, which further leads to a difficult calibration.

The second approach is based on the discovery that rainfall exhibits a scale-invariant behavior in both space (Mandelbrot, 1983) and time (Marsan et al., 1996; Venugopal et al., 1999). Multiplicative random cascade models were used to simulate the fractal nature of rainfall and were found to successfully reproduce the scale-invariant structure (Lovejoy and Schertzer, 1985; Schertzer and Lovejoy, 1987). This approach is attractive because the characteristics of rainfall across spatiotemporal scales are explicitly described via parsimonious parameterization. For this reason, random cascade models are extensively used for stochastic rainfall simulation (e.g. Over and Gupta, 1996; Menabde et al., 1997; Deidda, 2000). The State-of-the-art model based on this approach is the Short-Term Ensemble Prediction System (STEPS) which simulates the Lagrangian temporal evolution of the random cascades (Seed et al., 1999; Seed et al., 2013;

Raut et al., 2018). STEPS functions as rainfall nowcasting tool and is used by the Australian Bureau of Meteorology and UK Met Office. However, the spatiotemporal cascades inevitably become complicated if the scaling behaviors in space and time differ from each other distinctly (Seed, 2004). In addition, complexity will increase if the anisotropy in space and time needs to be addressed (Marsan et al., 1996; Niemi et al., 2016).

A common issue with the two widely implemented theories introduced above is that simulated rainfall heavily relies on calibration and post-processing to maintain key statistics (e.g. mean, variance, and spatiotemporal correlation). This would inevitably complicate the procedures when a long-term simulation is conducted with more stochastic processes involved. Alternatively, rainfall simulation using kriging estimators, or optimal estimation, is attractive in the sense that the key rainfall statistics are prescribed as model parameters and are inherently preserved throughout the modeling framework (Creutin and Obled, 1982; Tabios and Salas, 1985). As a result, the simulation using optimal estimation is statistically stable. In a kriging system, the estimate is essentially a linear combination of observations assigned with weights that are calculated to minimize error variance. It should be noted that the ‘observations’ herein are equivalent of the already-simulated pixels in the case of stochastic simulation. The kriging system takes advantage of the fact that optimal weights can be derived to minimize error variance without actually knowing the true value of the estimate.

Although the family of kriging estimation was extensively used in geo-statistics (e.g. Journel and Huijbregts, 1978; Solow, 1986; Oliver and Webster, 1990; Stein, 2012), the application in rainfall estimation was challenged by 1) the fractional nature of rainfall coverage and 2) the conditional bias. The kriging estimate would be the best statistical estimate only if the rainfall probability distribution were normal (Schweppe, 1973). Under fractional coverage

conditions, the probability distribution of rainfall is generally skewed and causes the kriging estimate to be less optimal. Therefore, procedures for optimal rainfall estimation must account for not only within-storm variability but also variability due to fractional coverage (i.e. rainfall intermittency). The advancement in optimal rainfall estimation was first enabled by Barancourt et al. (1992) who incorporated rainfall intermittency into the modeling framework. Further development was later made in a series of works (Seo, 1998a and b) in which conditional expectations and variance of rainfall occurrence and rainfall amount were analytically derived. The base of such estimation procedures, termed Double Optimal Estimation (DOE), was that the simple kriging estimator can be rewritten as the product of the expectation of rainfall given that it actually rained and the probability of actual raining. As a result, both the inner variability and intermittency of rainfall are explicitly considered. Because this approach is very general, DOE has been applied widely in real-time rainfall estimation using rain gauges only (Seo, 1998a), using rain gauges and radar combined (Seo, 1998b), as well as probabilistic quantitative precipitation forecast (Seo et al., 2000).

In the simple-kriging or ordinary-kriging system, because the error variance is minimized considering all involved observations (i.e. in the unconditional sense), kriging-estimates are subject to conditional biases. For instance, when used in rainfall estimation, kriging tends to underestimate heavy rainfall, because light to moderate rainfall observations are given more weights due to their more frequent occurrences than the heavy rainfall. However, in many applications, such as frequency analysis and risk management, what matters the most is the ability to accurately estimate large rainfall amount. As an effort to offset the excessive weighting to light/moderate rainfall, conditional bias-penalized kriging (CBPK, Seo, 2013) was developed as an extension to simple kriging or ordinary kriging. Due to its compatibility with the kriging

system, CBPK has been plugged into existing frameworks, and was found to significantly improve the accuracy of estimating heavy rainfall (Seo et al., 2014; Kim et al., 2016). In this study, it is considered very valuable to sufficiently reflect the extreme rainfall amounts in the stochastic simulation for facilitating frequency analysis. In addition, as suggested in Seo et al., (2014), the DOE may provide a natural framework for applying CBPK. Therefore, CB-penalized formulation of DOE is conducted and applied in stochastic rainfall simulation.

The goal of this work is to investigate the capability of stochastic storm simulation to realistically reproduce the key statistical properties observed from the radar rainfall. Specifically, simulation results are evaluated with respect to statistical moments, spatio-temporal structure, and frequency distribution of rainfall at both near-point scale and domain scale. This study serves the first step for applying the approach in rainfall frequency analysis. The paper is organized as follows: Section 2 describes the study area and data used. Section 3 introduces the methodology. Section 4 includes the results and discussion. Section 5 provides conclusion and future directions.

## 2. STUDY AREA AND DATA

The simulation domain is a 75 × 75 $km^2$ area in the Dallas-Fort-Worth metroplex as shown in **Figure 1A**. In addition, eight areas with the same size as the simulation domain are selected for generating ensemble parameters (detailed in Section 3.4). In this study, the total nine areas as shown in **Figure 1B** are considered homogeneous because of their similar precipitation patterns for various durations and frequencies in NOAA Atlas 14 (National Weather Service, 2017). The climate of the areas is humidsubtropical with hot summers and relatively mild winters (Kim et al., 2016). The radar data used are Multisensor Precipitation Estimates (MPE) obtained from the West Gulf River Forecast Center, since correction using rain gauges and

quality control are applied in MPE. All analyses in this work are performed on the Hydrologic Precipitation Analysis Project (HRAP) grid (Greene and Hudlow, 1982), which is approximately 4 km a side for the homogeneity areas. The MPE data are on a 16 × 16 HRAP grid (**Figure 1A**) over each homogeneity area. In addition, nine HRAP cells (**Figure 1A**) are selected in the simulation domain for evaluating simulation results at near-point scale.

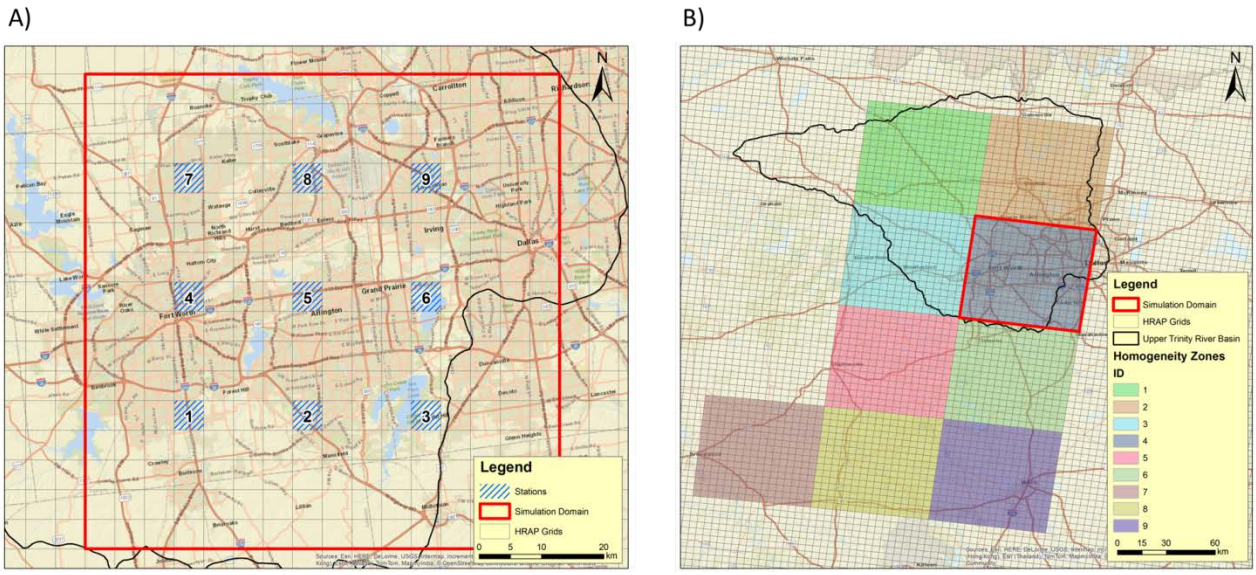



**Figure 1:** Study area – A) Dallas Fort Worth metroplex and B) nine homogeneity areas.

## 3. METHODOLOGY

### *3.1.Optimal Estimation*

In optimal estimation, the estimate is essentially a linear combination of observations. When optimal estimation is applied to simulate rain field, a pixel to be simulated is regarded as the estimate whereas those that have already been simulated are considered the observations. The simulation process goes through all the pixels in a random order to complete the whole rain field. For this study, Double Optimal Estimation (DOE, Seo, 1998a and b) is used for spatio-temporal simulating of rain fields. In a nutshell, DOE performs separate estimation of probability of

precipitation (POP) and conditional precipitation amount given that rainfall occurs, and then combines them for the final estimate. Specifically, simple kriging will be used to estimate three statistics of rainfall, i.e. POP, conditional mean and variance of rainfall amount given that rainfall occurs. The DOE can be viewed as a two-stage process, where POP is estimated first to determine the occurrence of rainfall and then the mean and variance of rainfall are estimated conditioned on rainfall occurrence. Since the POP is essentially the expected of value of rainfall indicator with "1" meaning rainfall and "0" meaning no rain, indicator kriging (e.g. Solow, 1986) is applied to estimate POP. The conditional mean and variance of rainfall are then estimated using simple kriging.

In addition, CB-penalized formulation is added into the second stage of DOE to evaluate its effect on the high rainfall values. This requires deriving a set of mathematical identities, which is documented with detail in the Appendix for the reader's reference. Note that by default the simulation in this study is conducted without CB-penalized formulation, which means $\beta$ in Equation A12 equals to 0. A comparison with scenarios with CB-penalized formulation is made and detailed in Section 4.4.

### *3.2. Spatio-temporal Simulation of Rain Field*

Spatio-temporal simulation is achieved by assuming that the rainfall amount in a pixel needs to be estimated using not only surrounding pixels at the current time step ($t$) but also those in the immediately preceding time step ($t$-1). In the context of kriging, this assumption requires a simple-cokriging estimator based on observations in both time steps. A challenge in the spatiotemporal simulation is that storm movement causes the spatial correlation function of rainfall between $t$ and $t$-1 to exhibit origin-asymmetry. Furthermore this leads to issues with generating the unique and optimal estimation. As a solution, the rainfall field at $t$-1 is transposed

along an advection vector and overlaid onto the rainfall field at *t*. Consequently, directional features (the origin-asymmetry) in the spatial correlation function between two time steps are effectively reduced.

The scheme for spatio-temporal simulation of rain field is illustrated in **Figure 2**. Initially, the rainfall field at *t*-1, termed $RF_{t-1}$, is transposed along an advection vector. A random pixel in the rain field at current time step ($RF_t$) is then selected for simulation, and its POP will be estimated both *n* nearby cells from $RF_t$ and *m* nearby cells from the transposed rainfall field at *t*-1 (termed $RF^*_{t-1}$) using **Equation A1** and **A2** in the Appendix. Next, a random number from the uniform distribution between 0 and 1 will be drawn and compared to the POP. If the random number is smaller than POP, then zero rainfall will be assigned to the pixel; otherwise the conditional mean $E\left[Z_{k0}\middle|z_{k-1}, z_k,\ Z_{k,0} > 0\right]$ and conditional variance $Var\left[Z_{k0}\middle|z_{k-1}, z_k,\ Z_{k,0} > 0\right]$ of positive rainfall is estimated also based on the nearby *n* pixels in $RF_t$ and *m* nearby pixels in $RF^*_{t-1}$ via **Equations A6**, **A7** and **A8** in the Appendix. $Z_{k0}$ is the rainfall to be estimated while $z_k$, and $z_{k-1}$ represent rainfall at nearby *n* pixels in $RF_t$ and *m* pixels in $RF^*_{t-1}$, respectively. Next, the rainfall value of the target pixel is randomly drawn from a probability distribution (e.g. log normal or Weibul) with the mean and variance being $E\left[Z_{k0}\middle|z_{k-1}, z_k,\ z_{k,0} > 0\right]$ (**Equation A7**) and $Var\left[Z_{k0}\middle|z_{k-1}, z_k,\ z_{k,0} > 0\right]$ (**Equation A8**), respectively. The simulation continues with the next random pixel till all pixels in $RF_t$ are visited.

With spatio-temporal simulation conducted using this scheme, there are underlying assumptions made as follows. First, both rainfall occurrence and rainfall amount are secondary homogeneous in the simulation domain, which means the mean and variance of rainfall indicator (0 and 1) and positive rainfall are identical everywhere in the simulation domain. Second, cross-correlograms between $RF_t$ and $RF^*_{t-1}$ ($\rho_{Ic}$ and $\rho_{rc}$) are assumed to be identical to the spatial

correlograms ($\rho_{Ik} = \rho_{Ic}$ and $\rho_{rk} = \rho_{rc}$) to ensure unique optimal solutions (see, e.g. Journel and Huijbregts, 1978). Third, the storm movement between *t* and *t*-1 can be characterized as a uniform advection vector over the whole simulation domain. It should also be noted that rainfall at *t* is not strongly correlated to rainfall at *t*-1 when rainfall quickly enters/exists the simulation domain or drastically intensifies/diminishes. Under such conditions, the simple-cokriging estimator is not able to generate a unique optimal solution based on both $RF_t$ and $RF^*_{t-1}$. Instead, a simple-kriging estimator is used considering only nearby pixels in $RF_t$. Fourth, because DOE only estimates the conditional mean and variance of rainfall instead of its full probability distribution, the simulation requires assuming a probability density function (pdf) for the positive rainfall amount to allow conditional simulation. A log-normal distribution is used for simulation by default and further compared to Weibul distribution (see Section 4.4) regarding the ability to capture high rainfall values.

It is worth noting that DOE cannot perfectly preserve probability of precipitation (POP), conditional mean ($E[Z_{k0}|\, z_{k,0} > 0]$) and conditional variance ($Var[Z_{k0}|z_{k,0} > 0]$) of positive rainfall with limited samples (herein 16 × 16 = 256 HRAP bins). According to Seo et al., 2000, a few hundred simulation traces may be needed to eliminate this uncertainty from sampling. For operational purposes, Seo et al, 2000 utilized post-processing to force the empirical cumulative distribution function (cdf) to match the cdf prescribed by POP, $E[Z_{k0}|\, z_{k,0} > 0]$, and $Var[Z_{k0}|z_{k,0} > 0]$. However, this post-processing strategy enviably distorts the spatio-temporal correlation structure of rainfall while preserving its univariate statistics (POP, $E[Z_{k0}|\, z_{k,0} > 0]$ and $Var[Z_{k0}|z_{k,0} > 0]$). In this study, an alternative strategy is employed: the one simulation that produces the best match of cdf with the prescribed is selected from a number of simulation traces without post-

processing. Due to many experiments, 30 simulation traces can generate a satisfactory candidate and thus are utilized in this study.

***Parameter Estimation***

The simulation requires specification of five parameters at each time step: $m_I$, the expected fraction of rain coverage; $m_r$, mean of positive rainfall; $\sigma_r$, standard deviation of positive rainfall; $v_E$, eastward velocity of advection; $v_N$, northward velocity of advection. To best reach possible realizations in ensemble simulations, these five parameters are estimated based on MPE data not only in the simulation domain but also the other eight homogeneity areas. Using 10 years of MPE data, $m_I$, $m_r$, and $\sigma_r$, can be directly estimated for each of the nine areas via **Equations** 1, 2 and 3, respectively.

$$m_I = n_{pr}/n_r \tag{1}$$

$$m_r = \left(\frac{1}{n_{pr}}\right)\sum z_{pr} \tag{2}$$

$$\sigma_r = \sqrt{\sum \frac{z_{pr}^2}{n_{pr}-1} - \frac{(\sum z_{pr})^2}{n_{pr}(n_{pr}-1)}} \tag{3}$$

where $n_{pr}$ is the number of radar pixels with positive rainfall; $n_r$ is the total number of radar pixels; $z_{pr}$ is positive rainfall in a radar pixel.

The eastward and northward velocities of advection ($v_E$ and $v_N$) are determined as follows: 1) the rainfall field at *t*-1 is shifted based on different potential velocities (from 3 m/s to 36 m/s with increments of 3 m/s) and directions (from 0° to 180° with increments of 3°); 2) correlation coefficients between the shifted rainfall field at *t*-1 and that at *t* are calculated; 3) the combination of velocity and direction that generates the highest correlation coefficient will be

converted to integer numbers of pixels in the eastward and northward directions as $v_E$ and $v_{N.}$,respectively

In addition, the simulation algorithms used in this study require modeling spatial correlograms of intermittency ($\rho_{Ik}$ in **Equation A13**) and inner variability ($\rho_{rk}$ in **Equation A9**) of precipitation (Seo, 1998a, b). For simplicity, the spatial correlograms are assumed to be time-invariant. Kim et al (2016) estimated climatological $\rho_{Ik}$ and $\rho_{rk}$ for North Central Texas region and for all months using isotropic exponential model based on radar data from 2002 and 2008, which is therefore adopted in this study. In addition,

### *3.3. Ensemble Parameter Generation*

In ensemble simulations, numerous time series of parameters ($m_I$, $m_r$, $\sigma_r$, $v_E$, and $v_N$) need to be generated. Key statistics from the observed rainfall, e.g., mean, variance, temporal autocorrelation, etc. need to be preserved in the ensemble traces. Moreover, the generated ensemble traces should exhibit diverse new patterns. In this study, the parameters estimated for all homogeneity areas in 10 years are reshuffled and organized into new patterns via a non-parametric approach (Lee and Ouarda, 2011; Lee et al., 2012; Lee and Jeong, 2014; Lee and Park, 2017). This approach was originally developed to temporally downscale time series, featuring k-nearest neighbor resampling (KNNR) and the Genetic Algorithm (GA) mixing process. With some modifications, this approach (termed KNNR+GA) is applied in this study for stochastically generating ensemble parameters.

In the KNNR+GA approach, a metric called ‘distance’ needs to be defined to quantify the difference between two time series of parameters. First for the time series of $m_I$, $m_r$ and $\sigma_r$, the distance needs to represent the differences in the three parameters collectively because they are correlated to each other. For instance if $m_I$ is zero (or positive), $m_r$ and.$\sigma_r$ must be zero (or

positive).Therefore, a composite distance for these three parameters is quantified to represent the difference in probability distribution of rainfall values, as $m_I$, $m_r$ and $\sigma_r$ can all be derived from the probability distribution function. Also the distance for $m_I$, $m_r$ and $\sigma_r$ does not consider the spatial distribution of rainfall, because $m_I$, $m_r$ and $\sigma_r$ are assumed to be identical everywhere in the simulation domain (second-order homogeneity). Therefore, the distance for $m_I$, $m_r$ and $\sigma_r$ is calculated in this study as follows: 1) rainfall values in the simulation domain (16 ×16 HRAP bins) are sorted in ascending order within each hour of the 24-hour duration and reshaped into an array with the length of 16×16×24; 2) root mean square error (RMSE) of any two compared arrays is calculated as the distance. As for $v_E$, and $v_N$, they don't have much dependence on rainfall and thus their ensemble traces are generated independently. The distance for $v_E$, or $v_N$ is simply defined as the RMSE of the two compared 24-hr time series of $v_E$, or $v_N$.

With the distances clearly defined, the procedures for generating ensemble traces for a 24-hr time series of parameters (target time series) using the KNNR+GA approach are as follows:

1) Estimate the distances between the target time series and all the other observed time series from the 9 homogeneity areas including the simulation domain. Each observed time series is obtained by sliding a 24-hour moving window forward by one hour, which means the obtained time series do not necessarily start from a fixed time of the day.

2) Rank the estimated distances in Step 1 in ascending order and select *h* time series that have the smallest distances and do not overlap with each other, from which then select the first *k* time series with $k \approx \sqrt{h}$ (Lee and Ouarda, 2011).

3) Randomly select one of the *k* stored time series, termed $R_p$, with the weighting probability $w_m$ given by **Equation 4**

$$w_m = \frac{1/m}{\sum_{j=1}^{k} 1/j}, \quad m = 1, \dots, k \tag{4}$$

4) Apply GA on the $R_p$ in Step 3 via the following three steps:

4a) Reproduction: Select another time series termed $Rp^*$ using Steps 1 to 3.

4b) Crossover: Replace each time step in $R_p$ with the corresponding time step in $R_p^*$ at the probability $P_c$ and under the conditions shown in **Equation 5**

$$R_{p,i} = R^*_{p,i} \;\; if \begin{cases} \varepsilon < P_c \\ \left(R_{p,i-1} - R^*_{p,i}\right)\left(R^*_{p,i} - R_{p,i+1}\right) > \left(R_{p,i-1} - R_{p,i}\right)\left(R_{p,i} - R_{p,i+1}\right) \end{cases}, \; i = 1, \dots, 24 \tag{5}$$

where $R_{p,i-1}$, $R_{p,i}$, and $R_{p,i+1}$ are *i-1*$^{th}$, *i*$^{th}$, and *i+1*$^{th}$ elements in and $R_p$, respectively; $R^*_{p,i}$ is *i*$^{th}$ elements in $R_p^*$; $\varepsilon$ is a uniformly distributed random number between 0 and 1. The second condition is to preserve gradual variation in the time series.

4c) Mutation: Replace each time step in $R_p$ with each corresponding time step in a time series randomly selected from the *h* stored time series, $R_h$, at the probability $P_m$ as shown in **Equation 6.**

$$R_{p,i} = R_{h,i} \;\; if \; \varepsilon < P_m, \; i = 1, \dots, 24 \tag{6}$$

where $R_{h,i}$ is *i*$^{th}$ elements in and $R_h$.

## 4. RESULTS AND DISCUSSION

In this study, 100 most rainy days are selected based on the daily areal average rainfall in the simulation domain and then simulated for 50 ensembles. The simulated results are validated

against the radar observations with respect to various aspects as discussed in the following sections.

### *4.1. Ensemble Parameters*

The KNNR+GA method is used to generate 50 ensemble time series of $m_I$, $m_r$, $\sigma_r$, $v_E$, and $v_N$ for the 100 rainy days. As an example, **Figure 2** shows the comparisons between the ensemble and observed parameters for four rainy days. It is found that the ensemble parameters well encompass the observed values, and the ensemble median values in general match the observed. Two tuning parameters, i.e. the crossover probability ($P_c$ in **Equation 5**) and the mutation probability ($P_m$ in **Equation 6**), are adjusted so that 1) the ensemble parameters can preserve key statistical properties of the observed parameters; and 2) the ensemble parameters can exhibit new and diverse patterns. As $P_c$ and $P_m$ decrease, the ensemble parameters contain less diverse new patterns but may be subject to smaller biases. $P_c$ = 0.3 and $P_m$ = 0.1 are initially tested and eventually tuned to $P_c$ = 0.1 and $P_m$ = 0.01. Under these two parameter settings, **Figure 3** shows key statistics, including mean, standard deviation, skewness, and maximum of the ensemble and observed parameters, respectively. It can be found that the final parameter setting ($P_c$ = 0.1 and $P_m$ = 0.01) causes the ensemble values to better match the observed, as indicated by ensemble median. In addition, although the final $P_c$ and $P_m$ are smaller than the initial, new patterns in the ensemble values are not significantly reduced as indicated by the interquartile range. Another important statistic is the temporal autocorrelation, as new patterns of ensemble parameters should be temporally correlated at the similar level as the observed parameters. **Figures 4** shows the temporal autocorrelations of the ensemble and observed parameters at one- to five-hour lags under the initial and final settings of tuning parameters, respectively. Note the final $P_c$ and $P_m$ values in general improve the match between the temporal

autocorrelations of ensemble and observed parameters. With $P_c$ = 0.1 and $P_m$ = 0.01, the temporal autocorrelations are preserved in the ensemble simulations for all the parameters at all lags.

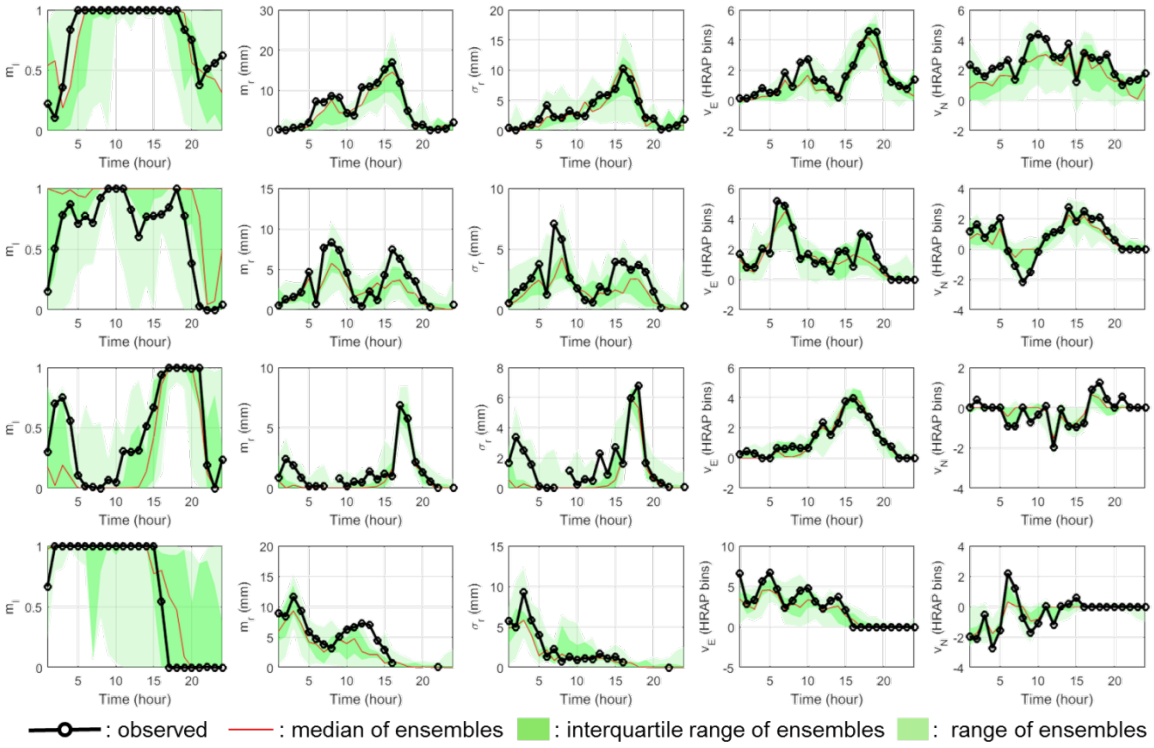



**Figure 2:** Comparison of ensemble and observed parameters ($m_I$, $m_r$, $\sigma_r$, $v_E$, and $v_N$) for four rainy days (four rows).

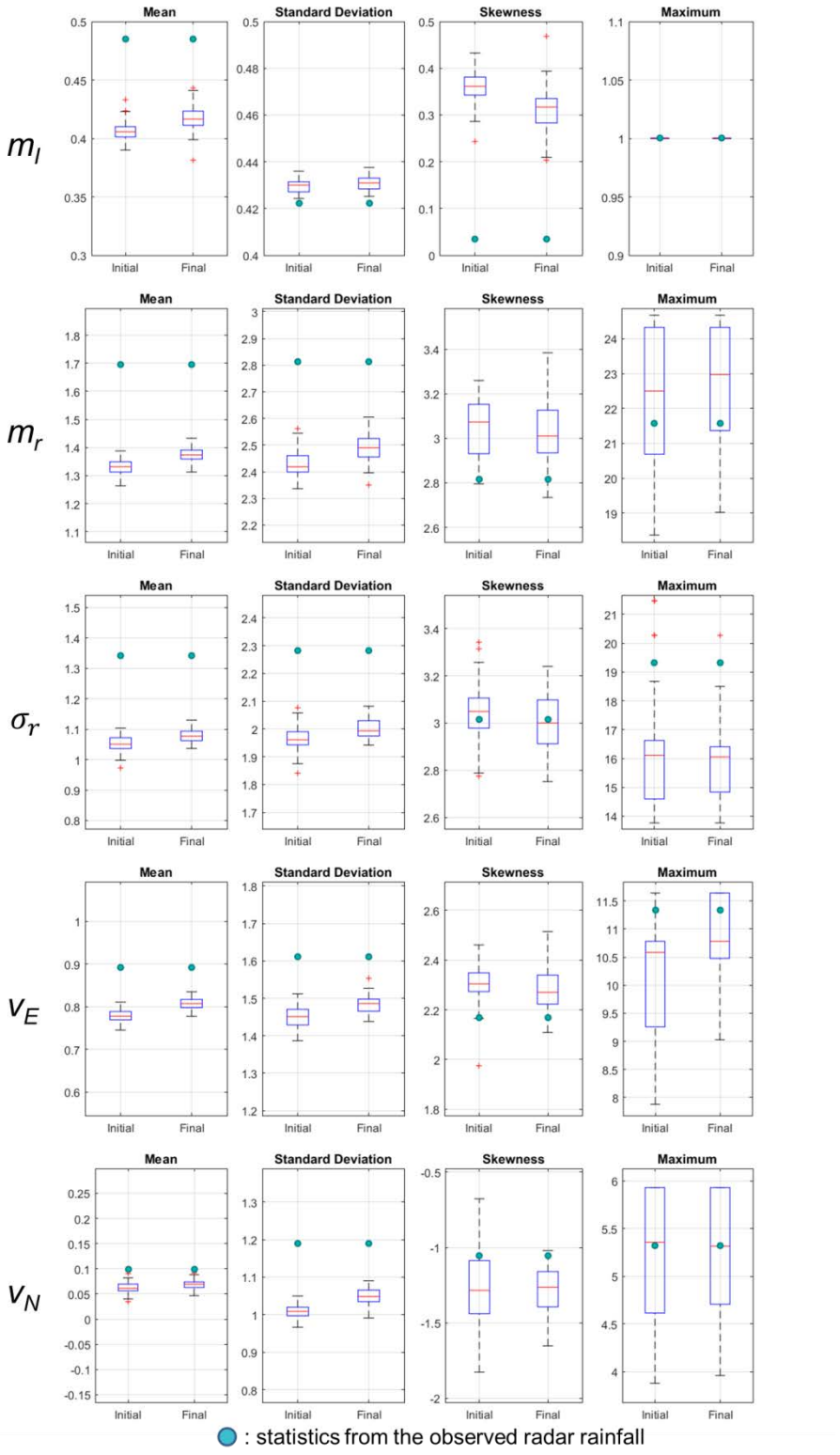



**Figure 3:** Key statistics (four columns) i.e. mean, standard deviation, skewness, and maximum of the ensemble and observed parameters (five rows).

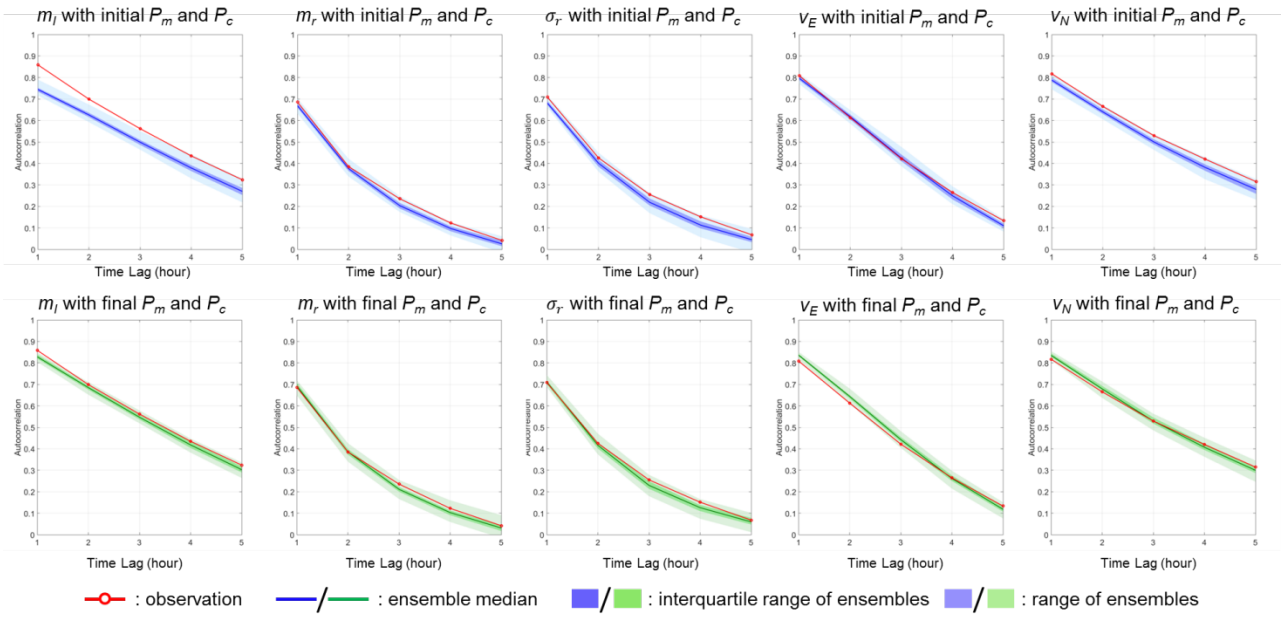



**Figure 4:** Temporal autocorrelations of the ensemble and observed parameters under the initial (upper row) and final (lower row) settings of tuning parameters

### *4.2. Spatial Structure of Rainfall*

#### *4.2.1. Daily Total*

The observed daily rainfall for the most rainy day in the 100 selected days is shown in the top-left panel in **Figure 5A**. The remaining 15 panels are the first 15 of the 50 ensemble simulations of the same event. In essence, each ensemble simulation is a realization of the same storm with similar but different evolutions of $m_I$, $m_r$, $\sigma_r$, $v_E$, and $v_N$. This is the main reason the accumulation patterns of the ensemble simulations appear different from each other. Furthermore, maxima (minima) of the daily rainfall in the ensemble simulations occur at different locations in the simulation domain because of the assumption of secondary homogeneity. **Figure 5B** shows the empirical cumulative distribution functions (cdfs) of the ensemble (same 15 members as in **Figure 5A**) and observed daily rainfall. Similarities can be

found between the ensemble and observed CDFs of daily rainfall with some ensemble simulations underestimating small to midrange values.

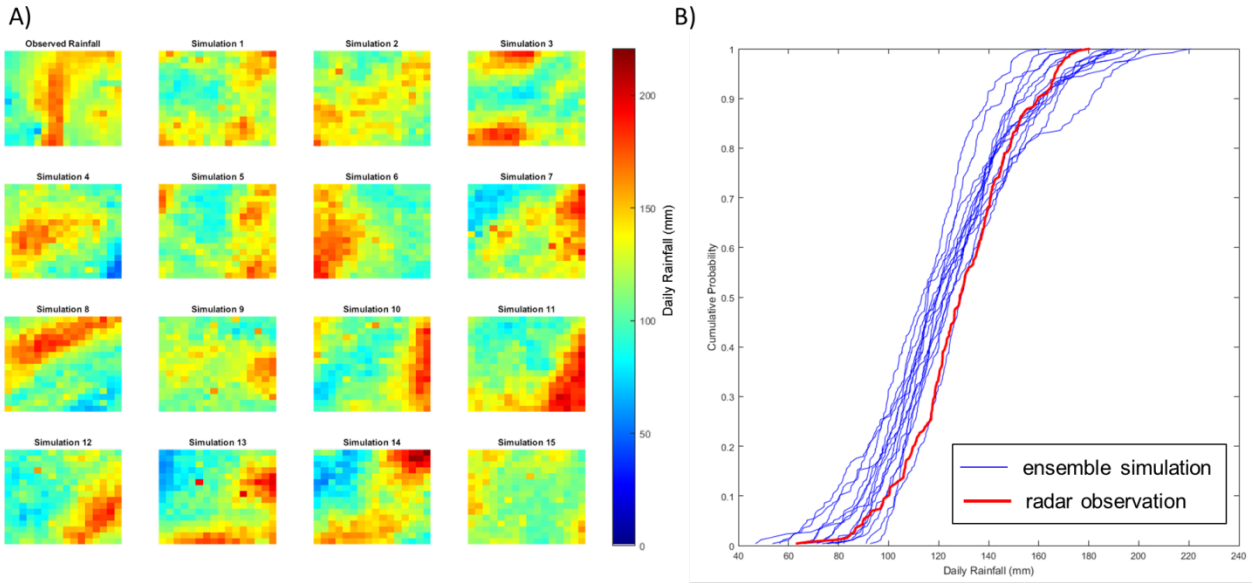



**Figure 5:** Comparison of A) daily total rainfall and B) the corresponding empirical cumulative distribution functions (cdfs) on from 15 ensemble simulations and the radar observation.

### *4.2.2. Hourly Rainfall*

The spatial structure of hourly rainfall is examined using spatial autocorrelations calculated as the Moran's I. In order to increase sample size for a better estimation of Moran's I, the samples of hourly rainfall are collected according to the months. **Figure 6** shows the comparisons of spatial autocorrelations of the observation and ensemble simulations calculated for six months, along with the corresponding climatological values. The remaining months are excluded because majority of the 100 rainy days occurred in the selected six months. Since all ensemble members are simulated using the climatological correlograms estimated by Kim et al., (2016), the spatial autocorrelations of the ensemble members align better with the climatological values than the observed. Overestimation is generally found for larger lag distances, which can

be due to the reduced sample size at larger lag distances. This issue could potentially be solved if more storm events were simulated.

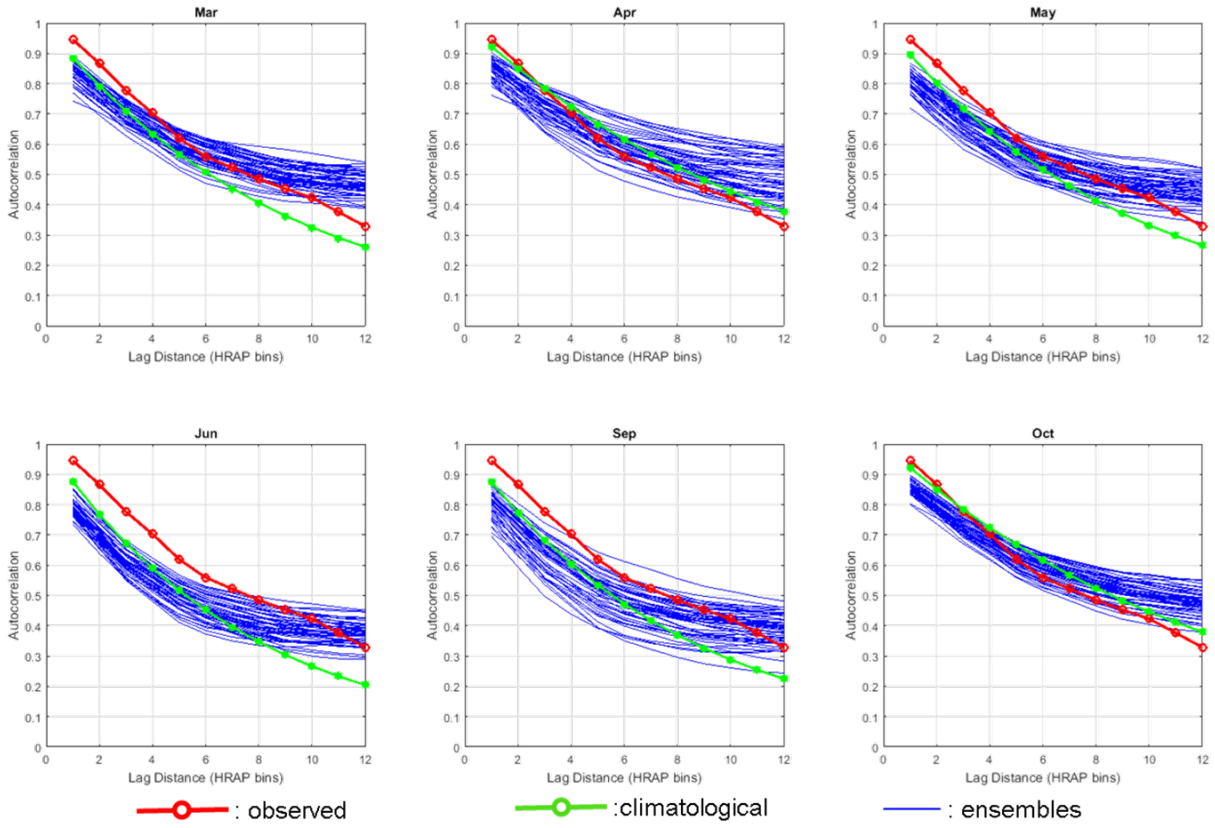



**Figure 6:** Spatial autocorrelations of hourly rainfall from the radar observation, ensemble simulations, and climatological values for six months.

### *4.3. Temporal Structure of Rainfall*

As learnt from the spatial structure of the simulated daily rainfall, the ensemble members exhibit a variety of patterns due to 1) the secondary homogeneity assumption and 2) the various patterns of the ensemble parameters. These variations mean that there are considerable differences at individual points in the simulation domain. In the following sections, the simulated rainfall time series at nine selected locations (**Figure 1A**) are discussed.

### *4.3.1. Autocorrelation Structure*

Although the autocorrelation structure of rainfall at the scale of simulation domain is preserved by the ensemble parameters ($m_I$ and $m_r$), temporal autocorrelations of rainfall at smaller scale (e.g. 1 HRAP bin) are more important and thus examined here. Correlograms of rainfall time series at the nine selected locations from observations and ensemble simulations are shown in **Figure 7**. In general, the ensemble simulations reproduce temporal correlation structure at this resolution reasonably well, with the observed autocorrelations lying within the envelope of the ensemble. Furthermore, the ensemble autocorrelations are generally similar from station to station, while differences can be found in the observed autocorrelations at the nine stations. Specifically, the ensemble mean of autocorrelations is very well simulated within the lag of 5 hours at Stations 4, 6, 7 and 9. However at the other stations, the simulations show either over- (Stations 1) or underestimation (Stations 2, 3, 5 and 8). This is probably due to the secondary homogeneity which is assumed in the ensemble simulations but is not completely true for the observed rainfall during the 100 rainy days.

### *4.3.2. Frequency Distribution*

Frequency distribution of hourly rainfall at the nine locations in the simulation domain are shown in **Figure 8**. The ensemble simulations are in good agreement with the observed frequency distribution at all nine locations. The ensemble variability well encompasses the observed frequency distribution, and its range increases with greater hourly rainfall. Note that simulations generally work well generating large hourly rainfall values, expect for slightly overestimating their frequencies at Station 4 and 5 as indicated by the longer tails.

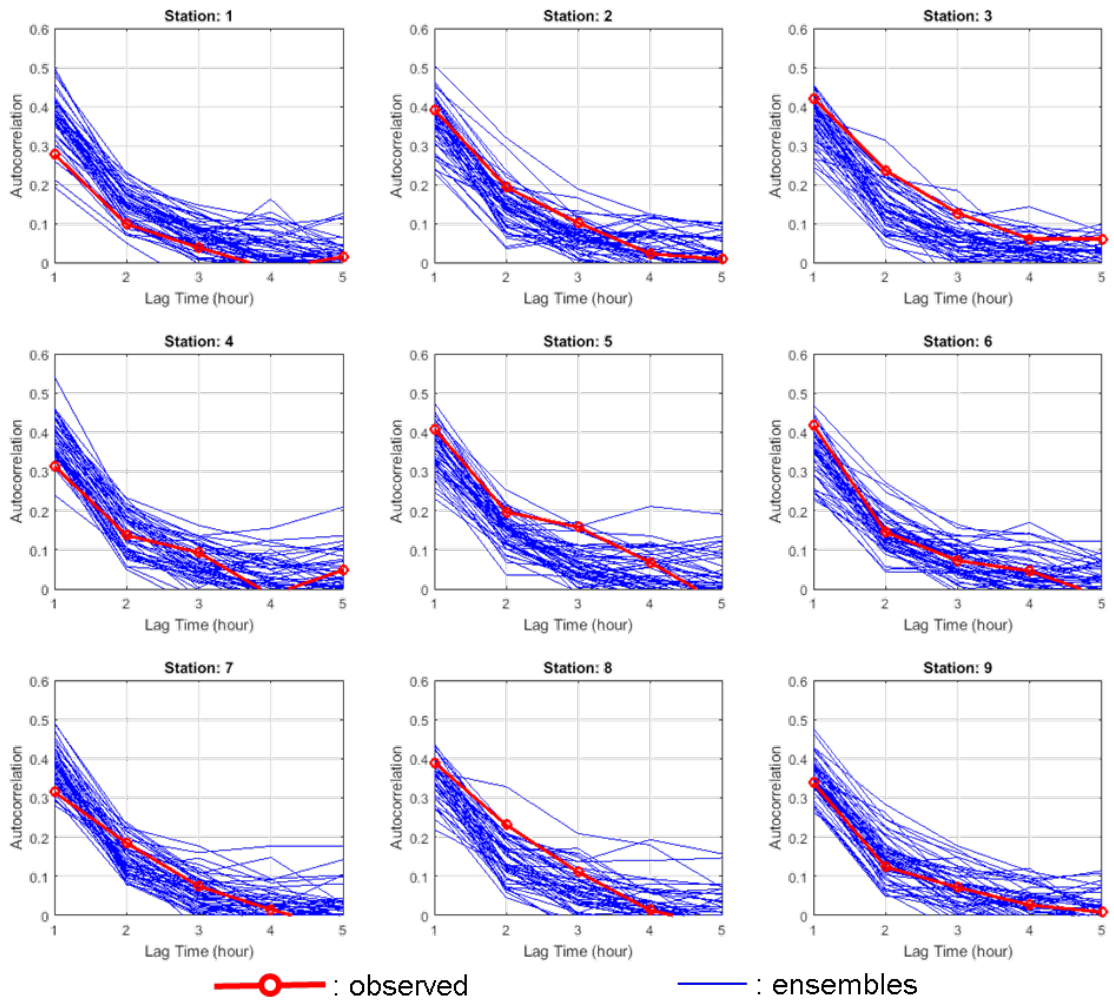



**Figure 7:** Temporal autocorrelations of hourly rainfall at nine selected stations from the radar observation and ensemble simulations.

### *4.1. High Rainfall Values*

Although the simulation in this study inherently preserves probability of precipitation (POP), conditional mean ($E[Z_{k0}|\, z_{k,0} > 0]$) and conditional variance ($Var[Z_{k0}|z_{k,0} > 0]$) of positive rainfall, what matters most in rainfall frequency analysis is the ability to estimate infrequent large rainfall amount as accurately as possible. In order to investigate the ability to capture high rainfall values, two approaches, i.e. changing pdf and changing DOE formulation, are tested. Specifically, log-normal distribution and Weibul distribution are tested with the latter

having a ‘fatter tail’ than the former; while DOE frameworks with and without CBP formulation are also tested.

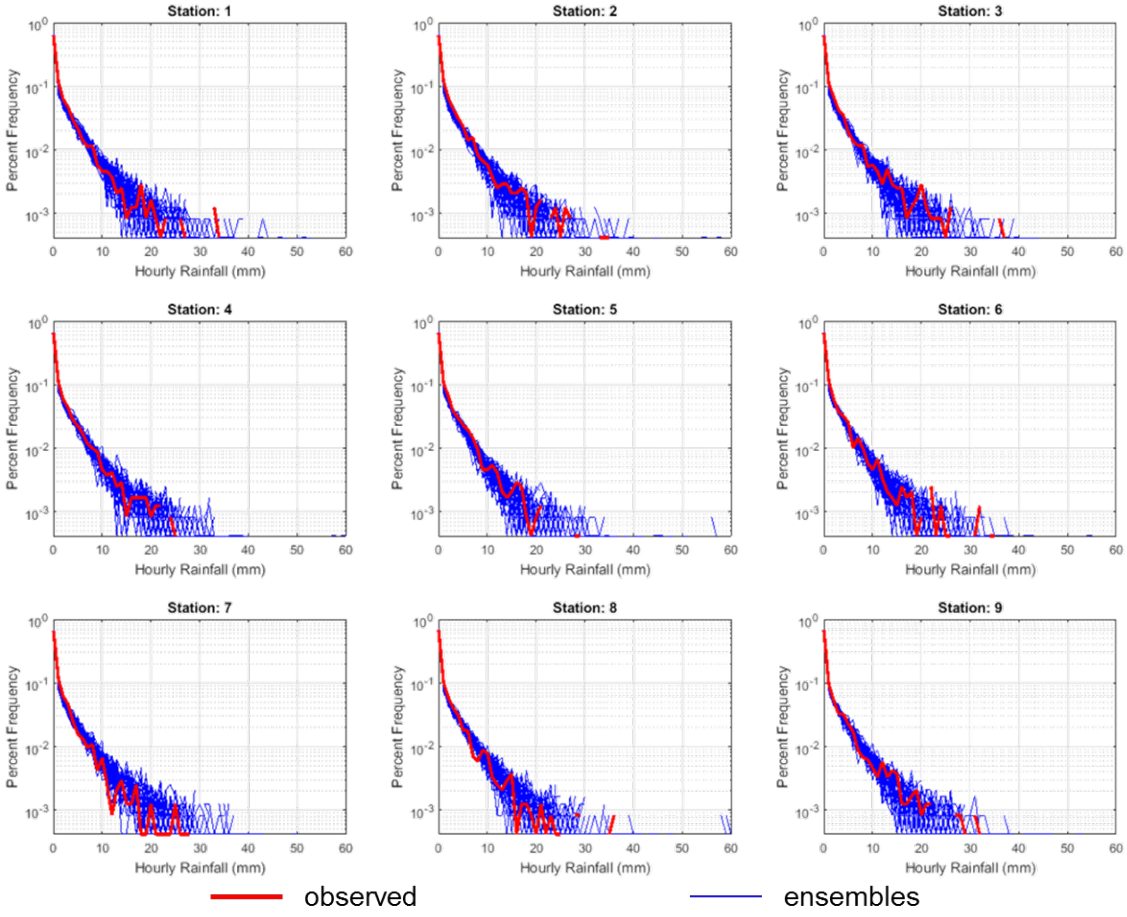



**Figure 8:** Frequency distributions of hourly rainfall at nine selected stations from the radar observation and ensemble simulations.

**Figure 9** shows frequency distributions of the hourly rainfall values at all pixels in the simulation domain and during all 100 rainy days in the 50 ensemble simulations under three combinations: 1) log-normal distribution and DOE, 2) Weibul distribution and DOE, and 3) log-normal distribution and CBP-DOE. It is found that change from log-normal to Weibul distribution does not generate any noticeable change in the frequency distribution of large hourly rainfall. In contrast, the CBP formulation causes the tail of frequency distribution to shift towards

higher hourly rainfall values. Moreover, tail behavior of the CBP-DOE simulation can further be adjusted by assigning more/less weight to the CB term via the parameter $\beta$ (Equation A12).

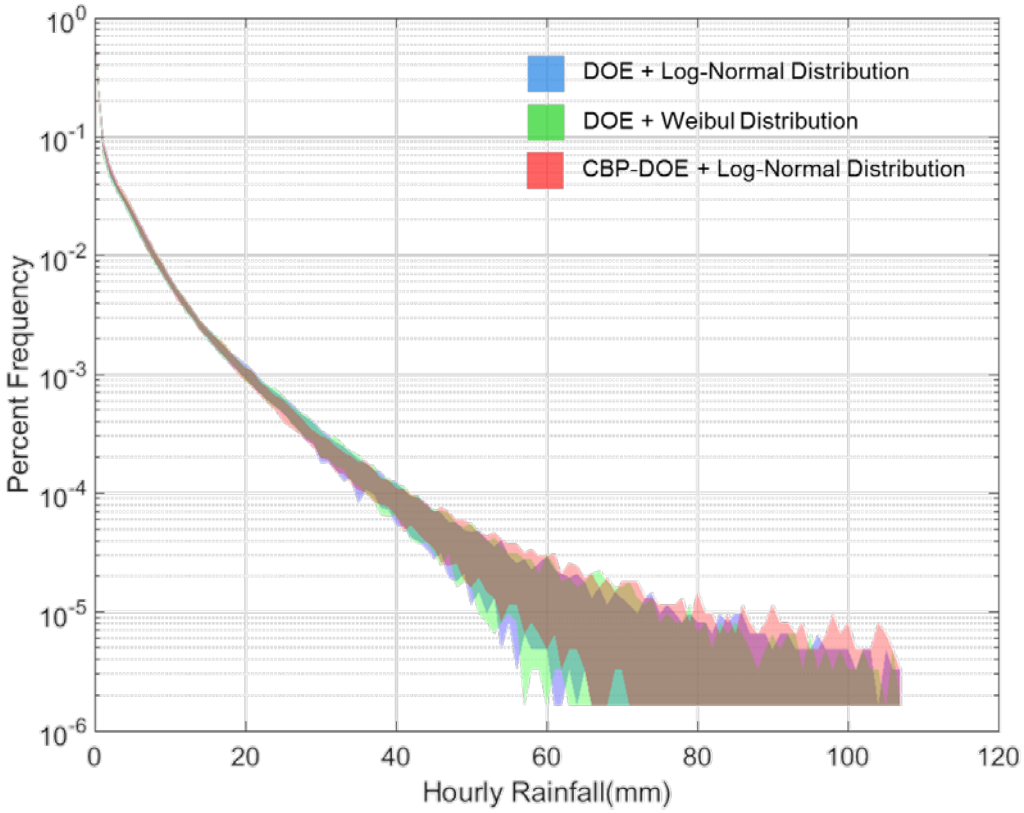



**Figure 9:** Frequency distributions of all hourly rainfall from ensemble simulations under three combinations of frequency distributions and formulations of optimal estimation.

### *4.2. Discussion*

This study is performed to demonstrate how realistically intense storms can be simulated without calibration or post-processing. However, prior to application in rainfall frequency analysis, simulation performance can potentially be improved by tackling the following issues. First, the ensemble generation of parameters is still subject to biases even after tuning $P_c$ and $P_m$ as indicated by the ensemble mean (**Figure 3**). One possible cause could be that the selected 100 days are too rainy to replicate using rainfall from other times in the 10 years and other eight

homogeneity zones. In other words, the results could potentially be improved if the targets were not the most rainy days but days better representing the average storm climatology. Second, post-processing as commonly used in previous studies (e.g. Seo et al., 2000 and Raut et al., 2018) could improve the results effectively. For this study, post-processing methods like quantile-quantile mapping could have been used to improve statistics of the ensemble parameters. A location-specific bias correction could also have been made to deal with any exceptions from the assumption of secondary homogeneity. Third, uncertainty exists when statistics are calculated based on limited samples. Essentially, the simulation is conducted at the resolutions of 4×4 $km^2$ and 1 hour. Given the size of simulation domain (16 HRAP × 16 HRAP bins) and the length of simulation duration (100 rainy days), estimation of spatial and temporal correlations lacks sample size especially at large lags. Improvements could be expected should the simulation be conducted for a longer period or at finer spatio-temporal resolution. The assumption of secondary homogeneity probably holds better with a longer simulation period. Moreover, the other two assumptions i.e. $\rho_{Ik} = \rho_{Ic}$, $\rho_{rk} = \rho_{rc}$ and uniform advection, would become more valid at finer temporal resolution. It is also worth noting that enlarging simulation domain as another solution to increase sample size might cause the assumption of secondary homogeneity to be invalid and thus requires extra caution.

## 5. CONCLUSIONS AND FUTURE WORK

In this study, the authors demonstrate an approach to stochastically generate numerous realizations of storm events based on radar rainfall data. To achieve spatio-temporal simulation of rainfall fields, the approach employs double optimal estimation (DOE) due to the following benefits: 1) the simulation is parsimonious requiring only five time-variant parameters i.e., $m_I$, $m_r$, $\sigma_r$, $v_E$, and $v_N$; 2) the simulation is statistically stable as $m_I$, $m_r$, $\sigma_r$ are inherently preserved

and thus no calibration is needed for maintaining these three statistics. In addition, the conditional biases commonly seen in kriging estimates can be effectively mitigated by the CBP-formulation, which also enable us to better capture high rainfall values. In order to realize diverse new patterns of the 5 parameters in ensemble simulations, the non-parametric KNNR+GA approach is used to borrow information from other locations and times. In essence, this approach trades space with time by assuming identical/similar climatology in the homogeneity areas. Fifty ensemble members of 100 simulated rainy days are evaluated against the observed from the radar with major findings summarized as below.

1) The KNNR+GA approach successfully generates ensemble parameters of new patterns while reproducing the key statistical properties of the observed parameters.
2) Realistic spatial patterns and temporal evolutions of rainfall fields are replicated by the spatio-temporal simulation using the DOE framework.
3) Diverse spatial structures of daily rainfall are due to the variability in ensemble parameters as well as the assumption of secondary homogeneity. Spatial structure of hourly rainfall aligns well with the prescribed climatological values.
4) Also at near-point scale, frequency distribution and temporal autocorrelations are reproduced at 9 selected stations.
5) As ways to capture high rainfall values, CBP formulation of DOE is more effective and flexible than altering probability density functions of hourly rainfall.

In terms of future directions, the authors will conduct continuous simulation from which rainfall frequency analysis will then be carried out. Comparison will then be made between the intensity-duration-frequency results from the simulated rainfall and those in the current NOAA

Atlas 14. Note that the continuous simulation will be based on not only radar rainfall data but also rain gauge data for generating ensemble parameters. Estimation of correlograms will be based on storm type and stage of maturity, instead of the climatological values used in this work. In addition, the authors see further applications of the stochastic storm generator in many areas: 1) enabling flood frequency analysis via hydrologic simulations at non-gauged locations or where streamflow records lack length or resolution, 2) downscaling outputs from global circulation model to investigate future changes due to climatic variability; 3) merging extrapolation-based forecasts and numerical weather prediction to improve accuracy and lead time of rainfall forecasts.

## ACKNOWLEDGMENT

The authors would like to thank Dr. Dong-jun Seo, Dr. Philip Bedient, Dr. Habib Ahmari, and Jerry Cotter for their valuable inputs to this work.

## APPENDIX

### *Equations Used in Spatiotemporal Simulation*

The conditional POP at the target cell ($E[I_{k0}|i_{k-1}, i_k]$) will be based on both *n* nearby cells from time *t* and *m* nearby cells from time *t*-1:

$$\left(\lambda_{k,1}, \ldots, \lambda_{k,n}, \lambda_{k-1,1}, \ldots, \lambda_{k-1,m}\right) = \left[P_{k,0} P_{k-1,0}\right] \begin{bmatrix} P_{k,k} & P_{k,k-1} \\ P_{k-1,k} & P_{k-1,k-1} \end{bmatrix}^{-1} \quad \text{(A1)}$$

$$E[I_{k0}|i_{k-1}, i_k] = E\left[I_{k,0}\right] + \sum_{i=1}^{n} \lambda_{k,i}\left(i_{k,i} - E\left[I_{k,i}\right]\right) + \sum_{j=1}^{m} \lambda_{k-1,j}\left(i_{k-1,j} - E\left[I_{k-1,}\right]\right) \quad \text{(A2)}$$

where subscripts *k* and *k*-1 denote time *t* and *t*-1, respectively; $i_{k,i}$ is the rainfall indicator at nearby pixel *i* (*i* = 1,2, …, *n*) in $RF_t$; $i_{k-1,j}$ is the rainfall indicator at nearby pixel *j* (*j* = 1,2, …, *m*) in $RF^*_{t-1}$; $\lambda_{k,i}$ and $\lambda_{k-1,j}$ are optimal weights associated with *n* and *m* nearby pixels in $RF_t$ and,

respectively; $P_{k,0}$ are 1 by $n$ vector whose $i^{th}$ entry is given by $Cov(I_{k,0}, I_{k,i})$; $P_{k-1,0}$ are 1 by $m$ vectors whose $i^{th}$ entry is given by $Cov(I_{k,0}, I_{k-1,j})$; $P_{k,k}$, $P_{k,k-1}(= P_{k-1,k}^T)$ and $P_{k-1,k-1}$ are $n$ by $n$, $n$ by $m$, and $m$ by $m$ matrices whose $ij^{th}$ entries are given by $Cov(I_{k,i}, I_{k,j})$, $Cov\left(I_{k,i}, I_{k-1,j}\right)$, and $Cov\left(I_{k-1,i}, I_{k-1,j}\right)$, respectively. Assuming the precipitation occurrence at time t and time t-1 is jointly wide-sense second-order homogeneous in the simulation domain, the indicator covariance terms can be calculated via Equations A3, A4, and A5.

$$Cov(I_{k,i}, I_{k,j}) = \rho_{Ik}(u_i - u_j)\sigma_{Ik}^2 \quad \text{(A3)}$$

$$Cov(I_{k-1,i}, I_{k-1,j}) = \rho_{Ik-1}(u_i - u_j)\sigma_{Ik-1}^2 \quad \text{(A4)}$$

$$Cov(I_{k,i}, I_{k-1,j}) = \rho_{Ic}(u_i - u_j)\sigma_{Ik}\sigma_{Ik-1} \quad \text{(A5)}$$

where $\rho_{Ik-1}()$ and $\rho_{Ic}()$ is the spatial correlation function of rainfall indicator at time $t$-1 and between $t$ and $t$-1, respectively; $\sigma_{Ik-1}$ is the standard deviation of rainfall indicator at time $t$-1. The conditional mean ($E\left[Z_{k0}\middle|z_{k-1}, z_k,\ z_{k,0} > 0\right]$) and conditional variance ($Var\left[Z_{k0}\middle|z_{k-1}, z_k,\ z_{k,0} > 0\right]$) of positive rainfall will be estimated via **Equations A6**, **A7** and **A8**:

$$\left(\Gamma_{k,1}, \dots, \Gamma_{k,n}, \Gamma_{k-1,1}, \dots, \Gamma_{k-1,m}\right) = (1+\beta)\left[Q_{k,0}Q_{k-1,0}\right]\left(\begin{bmatrix} Q_{k,k} & Q_{k,k-1} \\ Q_{k-1,k} & Q_{k-1,k-1} \end{bmatrix} + \beta \begin{bmatrix} M_{k,k} & M_{k,k-1} \\ M_{k-1,k} & M_{k-1,k-1} \end{bmatrix}\right)^{-1} \quad \text{(A6)}$$

$$E\left[Z_{k0}\middle|z_{k-1}, z_k,\ z_{k,0} > 0\right] = E\left[Z_{k,0}|z_{k,0} > 0\right] + \sum_{i=1}^{n} \Gamma_{k,i}\left(z_{k,i} - E\left[Z_{k,i}|z_{k,0} > 0\right]\right) + \sum_{j=1}^{m} \Gamma_{k-1,j}\left(z_{k-1,j} - E\left[Z_{k-1,j}|z_{k,0} > 0\right]\right) \quad \text{(A7)}$$

$$Var\left[Z_{k0}\middle|z_{k-1}, z_k,\ z_{k,0} > 0\right] = Var\left[Z_{k0}\middle|\ z_{k,0} > 0\right] - \left(\Gamma_{k,1}, \dots, \Gamma_{k,n}, \Gamma_{k-1,1}, \dots, \Gamma_{k-1,m}\right)\left[Q_{k,0}Q_{k-1,0}\right]^T \quad \text{(A8)}$$

where $Z_{k0}$ is the rainfall to be estimated; $z_{k,i}$ is the rainfall at nearby pixel $i$ ($i$ = 1,2, …, $n$); $z_{k-1,j}$ is the rainfall at nearby pixel $j$ ($j$ = 1,2, …, $m$) at time $t$-1; $\Gamma_{k,i}$ and $\Gamma_{k-1,j}$ are optimal weights obtained by minimizing error variance, $E[\left(Z_{k,0} - Z_{k,0}^{*}\right)^2 | z_{k,0} > 0]$ and conditional bias (CB), $E[\left(Z_{k,0} - E\left(Z_{k,0}^{*} | Z_{k,0}\right)\right)^2 | z_{k,0} > 0]$; $Q_{k,0}$ are 1 by $n$ vector whose $i^{th}$ entry is given by $Cov(Z_{k,0}, Z_{k,i} | z_{k,0} > 0)$; $Q_{k-1,0}$ are 1 by $m$ vectors whose $i^{th}$ entry is given by $Cov(Z_{k-1,0}, Z_{k,j} | z_{k,0} > 0)$; $Q_{k,k}$, $Q_{k,k-1} (= Q_{k-1,k}^{T})$ and $Q_{k-1,k-1}$ are $n$ by $n$, $n$ by $m$, and $m$ by $m$ matrices whose $ij^{th}$ entries are given by $Cov(Z_{k,i}, Z_{k,j} | z_{k,0} > 0)$, $Cov(Z_{k,i}, Z_{k-1,j} | z_{k,0} > 0)$, and $Cov(Z_{k-1,i}, Z_{k-1,j} | z_{k,0} > 0)$, respectively. $M_{k,k}$, $M_{k,k-1} (= M_{k-1,k}^{T})$ and $M_{k-1,k-1}$ are $n$ by $n$, $n$ by $m$, and $m$ by $m$ matrices whose $ij^{th}$ entries are given by $Cov\left(Z_{k,0}, Z_{k,i} | z_{k,0} > 0\right) Cov\left(Z_{k,0}, Z_{k,j} | z_{k,0} > 0\right) / \sigma_{rk}^2$, $Cov\left(Z_{k,0}, Z_{k,i} | z_{k,0} > 0\right) Cov\left(Z_{k,0}, Z_{k-1,j} | z_{k,0} > 0\right) / (\sigma_{rk} \sigma_{rk-1})$, and $Cov\left(Z_{k,0}, Z_{k-1,i} | z_{k,0} > 0\right) Cov\left(Z_{k,0}, Z_{k-1,j} | z_{k,0} > 0\right) / \sigma_{rk-1}^2$, respectively. $\sigma_{rk}$ and $\sigma_{rk-1}$ are the standard deviation of positive rainfall at $t$ and $t$-1, respectively. $\beta$ is a scalar weight given to the CB. When $\beta = 0$, all equations in the Appendix become the original ones in DOE (Seo 1998a and b). A discussion on $\beta$ is given later.

The derivation of $Cov(Z_{k,0}, Z_{k,i} | z_{k,0} > 0)$, $Cov\left(Z_{k,i}, Z_{k,j} | z_{k,0} > 0\right)$, and $Cov(Z_{k,i}, Z_{k-1,j} | z_{k,0} > 0)$, and $Cov\left(Z_{k-1,i}, Z_{k-1,j} | z_{k,0} > 0\right)$ is detailed in Seo (1998a, b), which involves using the covariance of positive rainfall calculated via **Equations A9, A10**, and **A11** under the assumption of secondary homogeneity:

$$Cov(Z_{k,i}, Z_{k,j} | z_{k,i} > 0, z_{k,j} > 0) = \rho_{rk}(u_i - u_j)\sigma_{rk}^2 \quad \text{(A9)}$$

$$Cov(Z_{k-1,i}, Z_{k-1,j} | z_{k-1,i} > 0, z_{k-1,j} > 0) = \rho_{rk-1}(u_i - u_j)\sigma_{rk-1}^2 \quad \text{(A10)}$$

$$Cov(Z_{k,i}, Z_{k-1,j}|z_{k,i} > 0, z_{k-1,j} > 0) = \rho_{rc}(u_i - u_j)\sigma_{rk}\sigma_{rk-1} \quad \text{(A11)}$$

where $\rho_{rk}()$ is the spatial correlation function of positive and $(u_i - u_j)$is the vector pointing from location $u_j$ to $u_i$ $\rho_{rk-1}()$ and $\rho_{rc}()$ is the spatial correlation function of positive rainfall at time *t*-1 and between *t* and *t*-1, respectively.

***Specification of β***

The weight given to CB term $\beta$ is specified as follows. First, the simple kriging solution without the CB term (i.e. $\beta$ = 0) is calculated as a reference. It is assumed that when the simple kriging solution is closer to $E[Z_{k,0}|z_{k,0} > 0]$, it becomes more skillful hence should be given more weight (Seo 2013; Seo et al., 2014; Kim et al., 2016). Thus $\beta$ varies with deviation of the conditional mean from the unconditional mean as shown in **Equation A12**:

$$\beta = C{Z_{SK}}^2 \quad \text{(A12)}$$

where $Z_{SK}$ is standardized solution from simple kriging based on $E[Z_{k,0}|z_{k,0} > 0]$ and $\sigma_{rk}$; and *C* is an empirical coefficient. As a result, in the midranges of the transformed normal distribution, $Z_{SK}$ is close to zero (i.e., $\beta$ is very small) and hence the final estimate is closer to simple kriging solution, whereas in the tails $Z_{SK}$ is large (i.e., $\beta$ is large) and hence the final estimate is closer to the CBP estimate.

## REFERENCES

Berne, A., Delrieu, G., Creutin, J. D., & Obled, C. (2004). Temporal and spatial resolution of rainfall measurements required for urban hydrology. Journal of Hydrology, 299(3-4), 166-179.

Beuchat, X., Schaefli, B., Soutter, M., & Mermoud, A. (2011). Toward a robust method for subdaily rainfall downscaling from daily data. Water Resources Research, 47(9).

Bowler, N. E., Pierce, C. E., & Seed, A. (2004). Development of a precipitation nowcasting algorithm based upon optical flow techniques. Journal of Hydrology, 288(1-2), 74-91.

Burton, A., Kilsby, C. G., Fowler, H. J., Cowpertwait, P. S. P., & O'connell, P. E. (2008). RainSim: A spatial–temporal stochastic rainfall modelling system. Environmental Modelling & Software, 23(12), 1356-1369.

Brown, J. D., & Seo, D. J. (2013). Evaluation of a nonparametric post‐processor for bias correction and uncertainty estimation of hydrologic predictions. Hydrological Processes, 27(1), 83-105.

Cunnane, C. (1988). Methods and merits of regional flood frequency analysis. Journal of Hydrology, 100(1-3), 269-290.

Cowpertwait, P. S. P., Kilsby, C. G., & O'Connell, P. E. (2002). A space‐time Neyman‐Scott model of rainfall: Empirical analysis of extremes. Water Resources Research, 38(8), 6-1.

Cowpertwait, P. S. (2010). A spatial‐temporal point process model with a continuous distribution of storm types. Water Resources Research, 46(12).

Creutin, J. D., & Obled, C. (1982). Objective analyses and mapping techniques for rainfall fields: an objective comparison. Water resources research, 18(2), 413-431.

Deidda, R. (2000). Rainfall downscaling in a space‐time multifractal framework. Water Resources Research, 36(7), 1779-1794.

Greene, D. R., & Hudlow, M. D. (1982). HRAP project. *National Weather Service, Hydrologic Research Laboratory, Silver Spring, MD, internal publication*.

Journel, A. G., & Huijbregts, C. J. (1978). Mining geostatistics (Vol. 600). London: Academic press.

Kim, B., Seo, D. J., Noh, S. J., Prat, O. P., & Nelson, B. R. (2016). Improving multisensor estimation of heavy-to-extreme precipitation via conditional bias-penalized optimal estimation. Journal of Hydrology.

Karlin, S. (2014). A first course in stochastic processes. Academic press. New York.

Koutsoyiannis, D. (1994). A stochastic disaggregation method for design storm and flood synthesis. Journal of Hydrology, 156(1-4), 193-225.

Lee, T., & Jeong, C. (2014). Nonparametric statistical temporal downscaling of daily precipitation to hourly precipitation and implications for climate change scenarios. Journal of hydrology, 510, 182-196.

Lee, T., & Ouarda, T. B. (2011). Identification of model order and number of neighbors for k-nearest neighbor resampling. Journal of hydrology, 404(3-4), 136-145.

Lee, T., Ouarda, T. B., & Jeong, C. (2012). Nonparametric multivariate weather generator and an extreme value theory for bandwidth selection. Journal of hydrology, 452, 161-171.

Lee, T., & Park, T. (2017). Nonparametric temporal downscaling with event-based population generating algorithm for RCM daily precipitation to hourly: Model development and performance evaluation. Journal of hydrology, 547, 498-516.

Leonard, M., Lambert, M. F., Metcalfe, A. V., & Cowpertwait, P. S. P. (2008). A space‐time Neyman–Scott rainfall model with defined storm extent. Water Resources Research, 44(9).

Lovejoy, S., & Schertzer, D. (1985). Generalized scale invariance in the atmosphere and fractal models of rain. Water Resources Research, 21(8), 1233-1250.

Mandelbrot, B. B. (1983). The fractal geometry of nature/Revised and enlarged edition. New York, WH Freeman and Co., 1983, 495 p.

Marsan, D., Schertzer, D., & Lovejoy, S. (1996). Causal space‐time multifractal processes: Predictability and forecasting of rain fields. Journal of Geophysical Research: Atmospheres, 101(D21), 26333-26346.

Menabde, M., Seed, A., Harris, D., & Austin, G. (1997). Self‐similar random fields and rainfall simulation. Journal of Geophysical Research: Atmospheres, 102(D12), 13509-13515.

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

Nielsen, J. E., Thorndahl, S., & Rasmussen, M. R. (2014). A numerical method to generate high temporal resolution precipitation time series by combining weather radar measurements with a nowcast model. Atmospheric research, 138, 1-12.

Niemczynowicz, J. (1999). Urban hydrology and water management–present and future challenges. Urban water, 1(1), 1-14.

Niemi, T. J., Guillaume, J. H., Kokkonen, T., Hoang, T. M., & Seed, A. W. (2016). Role of spatial anisotropy in design storm generation: Experiment and interpretation. Water Resources Research, 52(1), 69-89.

Ogden, F. L., Sharif, H. O., Senarath, S. U. S., Smith, J. A., Baeck, M. L., & Richardson, J. R. (2000). Hydrologic analysis of the Fort Collins, Colorado, flash flood of 1997. Journal of Hydrology, 228(1-2), 82-100.

Oliver, M. A., & Webster, R. (1990). Kriging: a method of interpolation for geographical information systems. International Journal of Geographical Information System, 4(3), 313-332.

Over, T. M., & Gupta, V. K. (1996). A space‐time theory of mesoscale rainfall using random cascades. Journal of Geophysical Research: Atmospheres, 101(D21), 26319-26331.

Raut, B. A., Seed, A. W., Reeder, M. J., & Jakob, C. (2018). A Multiplicative Cascade Model for High‐Resolution Space‐Time Downscaling of Rainfall. Journal of Geophysical Research: Atmospheres, 123(4), 2050-2067.

Richardson, C. W. (1981). Stochastic simulation of daily precipitation, temperature, and solar radiation. Water resources research, 17(1), 182-190.

Schertzer, D., & Lovejoy, S. (1987). Physical modeling and analysis of rain and clouds by anisotropic scaling multiplicative processes. Journal of Geophysical Research: Atmospheres, 92(D8), 9693-9714.

Schweppe, F. C. (1973). Uncertain dynamic systems. Prentice Hall, Englewood Cliffs, NJ.

Seed, A. (2004). Modelling and forecasting rainfall in space and time. IAHS Publications-Series of Proceedings and Reports, 287, 137-152.

Seed, A. W., Pierce, C. E., & Norman, K. (2013). Formulation and evaluation of a scale decomposition ‐ based stochastic precipitation nowcast scheme. Water Resources Research, 49(10), 6624-6641.

Seed, A. W., Srikanthan, R., & Menabde, M. (1999). A space and time model for design storm rainfall. Journal of Geophysical Research: Atmospheres, 104(D24), 31623-31630.

Seo, D. J. (1998a). Real-time estimation of rainfall fields using rain gage data under fractional coverage conditions. Journal of Hydrology, 208(1-2), 25-36.

Seo, D. J. (1998b). Real-time estimation of rainfall fields using radar rainfall and rain gage data. Journal of Hydrology, 208(1-2), 37-52.

Seo, D. J., Perica, S., Welles, E., & Schaake, J. C. (2000). Simulation of precipitation fields from probabilistic quantitative precipitation forecast. Journal of Hydrology, 239(1-4), 203-229.

Seo, D. J. (2013). Conditional bias-penalized kriging (CBPK). Stochastic environmental research and risk assessment, 27(1), 43-58.

Seo, D. J., Siddique, R., Zhang, Y., & Kim, D. (2014). Improving real-time estimation of heavy-to-extreme precipitation using rain gauge data via conditional bias-penalized optimal estimation. Journal of hydrology, 519, 1824-1835.

Solow, A. R. (1986). Mapping by simple indicator kriging. Mathematical geology, 18(3), 335-352.

Stein, M. L. (2012). Interpolation of spatial data: some theory for kriging. Springer Science & Business Media.

Svensson, C., & Jones, D. A. (2010). Review of methods for deriving areal reduction factors. Journal of Flood Risk Management, 3(3), 232-245.

Tabios III, G. Q., & Salas, J. D. (1985). A comparative analysis of techniques for spatial interpolation of precipitation. JAWRA Journal of the American Water Resources Association, 21(3), 365-380.

Venugopal, V., Foufoula‐Georgiou, E., & Sapozhnikov, V. (1999). Evidence of dynamic scaling in space‐time rainfall. Journal of Geophysical Research: Atmospheres, 104(D24), 31599-31610.

Wang, L. P., Ochoa-Rodríguez, S., Van Assel, J., Pina, R. D., Pessemier, M., Kroll, S., ... & Onof, C. (2015). Enhancement of radar rainfall estimates for urban hydrology through optical flow temporal interpolation and Bayesian gauge-based adjustment. Journal of Hydrology, 531, 408-426.

Wilks, D. S. (1992). Adapting stochastic weather generation algorithms for climate change studies. Climatic change, 22(1), 67-84.

Wilks, D. S. (1999). Interannual variability and extreme-value characteristics of several stochastic daily precipitation models. Agricultural and Forest Meteorology, 93(3), 153-169.

Woolhiser, D. A. (1992). Modeling daily precipitation—progress and problems. Statistics in the environmental and Earth sciences, 5, 71-89.

Wright, D. B., Smith, J. A., & Baeck, M. L. (2013a). Critical examination of area reduction factors. Journal of Hydrologic Engineering, 19(4), 769-776.

Wright, D. B., Smith, J. A., Villarini, G., & Baeck, M. L. (2013b). Estimating the frequency of extreme rainfall using weather radar and stochastic storm transposition. Journal of hydrology, 488, 150-165

# Chapter 5: Conclusions and Future Research

The following sections summarize conclusions from previous chapters and outlines forthcoming research that is to be pursued in the near future.

## 8.1 CONCLUSIONS

In the U.S., historical record of rainfall observation hasn't exceeded 150 years and yet has driven many hydrologic/hydraulic (H/H) practices. With advancement in sensing technology, rainfall observation has been reaching higher levels of spatial granularity and scale. However, uncertainty still exists in rainfall observation so far, because H/H studies rely on heavy rainfall values that are rare in nature. What's more, one might argue the rainfall observation would never be sufficient if non-stationarity was assumed. In this regard, simulated rainfall might provide the needed solution to these issues if key statistical properties in the observed rainfall are well preserved in the simulations. This doctoral work explores rainfall simulation in the following two applications: 1) generating possible rainfall-runoff scenarios for investigating hydrologic response to storm characteristics, and 2) stochastically generating storm events for rainfall frequency analysis.

To this end, Chapter 2 presented an approach to transpose historical storms from their original locations to other possible areas, which in turn creates numerous rainfall-runoff scenarios. Hydrologic responses to storm characteristics were then analyzed using a flood-prone urban watershed in Houston, Brays Bayou. The study highlighted the role of Brays Bayou's flat terrain in affecting flood peaks based on two findings as follows: (1) flood peaks in Brays Bayou are highly correlated to/driven by total rainfall volume; (2) channel routing effect is pronounced in Brays Bayou in reducing the flood peaks. Also, several lessons were learned from this

exploratory work. First, deterministic storm transposition was limited by transposition distance and the number of severe storms. Second, storm location is the only factor allowed by this approach for investigation; other impactful storm characteristics, for instance storm movement and spatiotemporal variabilities of rainfall intensity cannot be modified and thus studied. These limitations motivated the development of a deterministic storm generator as introduced in Chapter 3.

The Dynamic Moving Storm (DMS) generator as detailed in Chapter 3 conceptualizes a storm event as a single-core storm cell moving and evolving over time. There are three basic components in DMS generator's framework, i.e. storm movement, spatial and temporal variabilities of rainfall intensity. By parameterizing each component, hydrologic responses can be associated either to individual storm parameters or to parameter interaction. A large variety of synthetic storm events allowed a global sensitivity analysis (GSA) on storm parameters, which was unavailable in Chapter 2. The GSA revealed the significant impact of storm velocity on flood peaks and the higher influence from parameter interaction on flood peaks than from individual parameters. A pairwise sensitivity analysis (PSA) was further conducted to examine the interactive patterns between storm velocity and the other parameters. Using DMS generator, I confirmed and augmented the finding in Chapter 2 that flood peaks in Brays Bayou are driven by total rainfall volume not peak rainfall intensity. Also this work confirmed the resonance effect of storm movement on flood peaks as discovered in previous studies.

Finally, the flat terrain's influence on the flood peaks was highlighted throughout the chapter. In essence, the DMS generator in Chapter 3 revealed a bigger picture than the storm transposition in Chapter 2 on hydrologic responses to storm characteristics, and deepened the understanding on the urban hydrometeorology of Brays Bayou. In order to eventually facilitate

rainfall frequency analysis (RFA), the future development of DMS generator as outlined at end of Chapter 3 requires addition of processes embedded in a stochastic framework. These additional processes in turn entail more parameters and a more difficult calibration, which has been seen in the previous studies as a common issue for storm-cell-based generators. After many trials and considerations, I switched to seek for a more parsimonious and statistically more stable method for stochastic storm generator as introduced in Chapter 4.

Chapter 4 introduced a stochastic storm generator which adopts optimal estimation for spatio-temporal modeling of rain fields and a non-parametric approach for generating model parameters. A case study in north central Texas was conducted as the first step towards applying a long-term stochastic rainfall simulation in RFA. This approach was proved to offer many promises as follows. First, the simulation is parsimonious requiring only five parameters. Second, the simulation is statistically stable as the first and second statistical moments of rainfall are inherently preserved. Third, the non-parametric approach for generating model parameters saves the effort to select and parameterize stochastic processes by borrowing information from other locations and times. In the case study, 50 ensembles of 100 rainy days were successfully simulated with the spatio-temporal patterns of observed rainfall well represented. In particular, the simulation framework shows the capability and flexibility in capturing heavy rainfall values or high-order extreme statistics, which is quite valuable for RFA.

## 8.2 FUTURE RESEARCH

This doctoral research includes three approaches to synthesize storm events as introduced in Chapters 2, 3, and 4, respectively. For the purpose of investigating hydrologic responses to storm characteristics, the DMS generator in Chapter 3 is a very direct approach as storm characteristics are represented by DMS parameters. However due to many simplified

assumptions, the synthetic storms from the DMS generator are less realistic than those from the stochastic storm generator in Chapter 4. Therefore, the stochastic storm generator can and will also be used for investigating hydrologic responses in the future research, from which the results are expected to be more meaningful the study area.

Chapters 2 and 3 produced some empirical relationships whose hydrologic/hydraulic implications were thoroughly discussed. Deeper insights could be expected should the hydrologic model be replaced by an analytical framework incorporating all important processes, e.g. overland routing and flow attenuation in channel routing. Future effort will be invested in utilizing a proper analytical framework (e.g. Menabde et al., 2001) for better explaining the empirical relationships.

The stochastic storm generator has many further applications as outlined at the end of Chapter 4. Among these many possibilities, the most attainable one would be a comparative analysis on the intensity-duration-frequency (IDF) relationships derived from the following two rainfall simulations: Simulation A based on only historical observed rainfall and Simulation B based on both historical observed rainfall and downscaled outputs from global circulation models (GCMs). Note that the two simulations will have the same duration (say till 2100) despite that Simulation B has longer base period than Simulation A. The comparison is expected to pinpoint the effects of climate change on rainfall at desired spatiotemporal scale and furthermore the implication of stationarity as long assumed in traditional RFA.

## 8.3 REFERENCES

Menabde, M., Veitzer, S., Gupta, V., & Sivapalan, M. (2001). Tests of peak flow scaling in simulated self-similar river networks. Advances in Water Resources, 24(9-10), 991-999.