UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

DEVELOPMENT AND EVALUATION OF A HYDROLOGICAL AND HYDRAULIC COUPLED FLOOD PREDICTION SYSTEM ENABLED BY REMOTE SENSING, NUMERICAL WEATHER PREDICTION, AND DEEP LEARNING TECHNOLOGIES

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

MENGYE CHEN

Norman, Oklahoma

2021

DEVELOPMENT AND EVALUATION OF A HYDROLOGICAL AND HYDRAULIC COUPLED FLOOD PREDICTION SYSTEM ENABLED BY REMOTE SENSING, NUMERICAL WEATHER PREDICTION, AND DEEP LEARNING TECHNOLOGIES

A DISSERTATION APPROVED FOR THE
SCHOOL OF CIVIL ENGINEERING AND ENVIRONMENTAL SCIENCE

BY THE COMMITTEE CONSISTING OF

Dr. Yang Hong, Chair

Dr. Ming Xue

Dr. Jonathan J. Gourley

Dr. Kanthasamy K. Muraleetharan

Dr. Randall L. Kolar

This dissertation is dedicated to my family.

To my late father, Xiaoke Wang, for building me a tough mind. And to my mother, Juan Chen, for unlimited support.

To my wife and my son, Chan and Arthur Shao, for motivation, love, and affection.

Acknowledgement

After a few years of entrepreneurship, I came to realize that it was almost impossible to achieve anything alone. Without the support and encouragement, intentionally or unintentionally, a personal accomplishment from me, can be very limited. I am fortunate to say that I have been surrounded by luxury support and help, and I am very thankful for all the giants in the scientific community that allow me to climb over their shoulders to take a peek.

It is not possible to overstate my gratitude towards my Ph.D. advisor, Prof. Yang Hong. The words are too plain. Without his active guidance, support, I would not have finished this program. I am very grateful for being involved in multiple research projects, such as NASA-SIRVIER, NSF-PIRE, etc. I am thankful to my members of committee, Dr. Randall L. Kolar, Dr. Ming Xue, Dr. Johnathan J. Gourley, and Dr. Kanthasamy K. Muraleetharan, for their support and guidance on my publications and researches.

I would like to thank all my colleagues in the HyDROS lab for the active work environment, especially, doctor-to-be Zhi Li, being the best work partner that I have worked with in my life thus far. I would also like to thank the HWS program for the teaching opportunities. And finally, I would like to thank Mr. John B. Davis and Mr. Nick Wu for showing me the very beginning of this exciting life journey.

I will forever be indebt to all the kindness and encouragement around me throughout the journey and I am prepared to do the same to others.

Table of Contents:

List of Figures

List of tables

Abstract

Floods triggered by extreme precipitation are the most frequently occurring and disastrous natural hazards in the world. However, it is still challenging to provide accurate and flood mapping, flood damage estimation, and flood forecast. The purpose of this dissertation is to develop a hydrological and hydraulic coupled flood prediction system, inundation MApping and Prediction (iMAP), which can provide comprehensive flood simulation and prediction including channel flow rate, flood return period, flood extent, surface flow speed, and direction, as well as inundation depth and soil moisture. Up until now, the Coupled Routing and Excess STorage (CREST) model family has been well documented and established both in research and in real-world operation. As a new member of the CREST family, the work in this dissertation carries on the features of CREST model, as being robust, efficient, automated, and globally applicable. Moreover, the study also evaluates multiple remote sensing and precipitation prediction technologies during the historical event Hurricane Harvey. The results of the studies demonstrate that the CREST-iMAP system has the ability to provide comparable Harvey flood simulation as multiple real-time and operational flood monitoring systems in the world, and the best result comes from using Multi-Radar Multi Sensor (MRMS) Quantitative Precipitation Estimates (QPE), which the combination considers as the best practice in the Contiguous United States (USA). The results also indicate that the uncalibrated precipitation estimates perform better during extreme events like Hurricane Harvey, and precipitation forecasts still need more improvement to provide more information on flood prediction. However, the Numerical Weather Prediction (NWP) product can provide a preliminary forecast of the maximum flood extent, while the deep learning method could potentially improve the displacement issues from NWP forecasts.

Chapter 1. Introduction

1.1 Problem statement

In the United States, severe storm-triggered floods cause $3.7 billion in property losses and about 110 fatalities per year (Ashley and Ashley, 2008). In recent years, the U.S. was not short on severe storm-related hazards, such as Hurricane Harvey, Irma, Maria, Michael and Florence causing the catastrophic flooding events in Houston TX, Key West FL, Puerto Rico etc. in 2017 and 2018, as well as the historic 2016 Louisiana Flood which was not triggered by a tropical cyclone. According to Barthel and Neumayer (2012), although there were increases in observations and early warning systems, more information about the infrastructure vulnerability and better construction and planning technologies, the impact of the storm-related natural disasters was still rising. This trend can be led by the increase of population, developing intrusion into the potentially hazardous area, transportation infrastructure expansion (Spiker and Gori, 2000), the increase of wealth and changing climate (Changnon, 2011). **As the trend of more frequent flooding is almost certain, it is important to estimate and predict where and when the flooded water would present, how deep the water would be and how fast the water would flow to fully evaluate the impact of a future flood to building structures, soil erosions, outdoor safety factors, etc. Therefore, the public needs a well-informed forecasting and warning system, which can assess, quantify, and present the life cycle of a flood comprehensively.**

Since the 1970s, the scientific community has made a great improvement of the capacity of flood modeling by combining with climate models, weathers models, hydrological models, and hydrodynamic models (Teng et al., 2017). The applications of flood modeling vary from flood risk assessment and mapping (Dutta et al., 2006), flood damage assessment (Merz et al., 2013),

global real-time flood forecasting (Flamig et al., 2020; Gourley et al., 2017; Sampson et al., 2015; Wing et al., 2017), engineering for flood prevention (Gallegos et al., 2009), post-flood river system hydrology (Dutta et al., 2013), soil and river bank erosion (Hardy et al., 2000), catchment hydrology (Abbott et al., 1986), and floodplain ecology (Zonta et al., 2005). Generally, the application requires considerations of the predictive output, spatial-temporal scale, and resolution while balancing the accuracy requirement and computing efficiency. For real-time simulation and forecast, faster run time and data assimilation are required to provide reliable results (Chen et al., 2013). However, despite decades of efforts from the hydrological and hydraulic research community, it is a great challenge to provide accurate flood modeling at a high spatial-temporal resolution even in regional scales (Teng et al., 2017).

The "grand challenge for hydrology" was raised by Wood et al. (2011) to provide the hyperresolution hydrological prediction capacities to the public, as the society critically demand the high spatial-temporal resolution forecasting for the floods and droughts. New remote sensing technology provides accurate, spatially and temporally fine observation over the globe, which helps to advance physics-based models for atmosphere, hydrologic and hydraulic processes. As well as the computing technology allows the high volume of the data computation, which further allows the combination of different modeling systems into an integrated real-time forecasting chain that could forecast all the cascading storm-related disastrous events while considering their related uncertainties. All these advancement in technology would all support the real-time decision making for risk management. Over the last decade, fundamental tools were gradually prepared to meet the "grand challenge" and provide new information to society. Yet, much effort is still needed to meet the challenge and to provide reliable information for flood monitoring and prediction.

1.2 Literature review

1.2.1 Hydrological modeling

Hydrological models have been implemented for water management since the first watershed hydrological model was developed in 1966 (Crawford and Linsley, 1966). Hydrological models have been evolving over the years with complexity from lumped-process models (Sugawara et al., 1984; Williams and Hann, 1978) to semi-distributed models (Beven and Kirkby, 1979) and fully distributed models (Abbott et al., 1986; Wang et al., 2011; Wigmosta et al., 1995).

In recent years, remote-sensing technologies provided new insights about global precipitation and runoff responses with real-time availability, which has improved the monitoring of the precipitation and modeling its cascading effect, such as the flood, landslides, debris flows, drought, etc. (Brakenridge et al., 2007; Hong et al., 2007b). To understand the environmental change and the terrestrial water cycle, the field of hydrology evolved to develop the global scale models and using remote sensing data as the observation sources (Hong et al., 2007a; Wang et al., 2011), with the addition of advanced weather and climate models, which further motivated the development of land surface hydrological models (National Research Concil, 1991). In the new generation of satellite missions, more data sources are enabling the development of better hydrological models, such as the Soil Moisture Active-Passive (SMAP) mission which provides global soil moisture at 1-10 km resolution (Entekhabi et al., 2010), the Surface Water and Ocean Topography (SWOT) mission which provides the surface water storage (Durand et al., 2010), and CoReH2O mission which provides the snow extent along with the equivalent water (Heliere et al., 2009). Advanced global and continental hydrologic models have shown very promising results thus far, which provide critical information over large spatial domains regarding the behavior of surface runoff (Hong et al., 2007a), streamflow, soil moisture, soil infiltration, and

evapotranspiration, with applications of coupled land surface model of the Coupled Routing and Excess Storage model- Ensemble Framework for Flash Flood Forecasting (CREST- EF5) (Clark et al., 2017; Flamig et al., 2020; Gourley et al., 2017; Wang et al., 2011), and Weather Research and Forecasting model – Hydrological Model (WRF-Hydro) (Arnault et al., 2016; Cohen et al., 2018; Gochis et al., 2017; Lin et al., 2018; Zhang et al., 2018).

Notwithstanding the remarkable achievements of the emerging data resources and advancing models in the field of hydrology, it is believed that the current models are not able to meet the societal demand of water management (Wood et al., 2011). For example, the prediction of water movement and storage alongside the accessible water with adequate water quality for target population was challenging to meet the demand of the developing counties (Sachs and McArthur, 2005). One specific societal demand from the hydrological model is flood simulation and forecasting, which require higher spatial-temporal resolutions than what is currently available. The high-resolution modeling at continental scales would better represent the reality when considering the spatial heterogeneity in topography, soils, and vegetations, which will impact the hydrology dynamics. Therefore, the simulation would consider in better details of the effects on solar radiation at slopes, snowmelt, soil moisture and evapotranspiration, as well as the channel routing, inundation extent, flood water depth, impact population and potential infrastructure damage. However, the available land surface models are constrained by the resolution of the global weather and climate models and the hydrological models' low spatial resolution (Wood et al., 2011).

1.2.2 Observation and conceptual modeling approaches to flood inundation mapping

Flood inundation mapping has been studied for over a century, and two traditional groups of research effort gained most of the attention of the research community: observations and

hydraulic models. The observations include ground measurements, survey, remote sensing technology (Shen et al., 2019b, 2019a) and statistical models (Kundzewicz et al., 2013). As the remote sensing technology could identify the flood extent over a vast area, while other methods can only provide data at single points, the flood mapping using remote sensing data has gained popularity in recent years. However, due to the limited representation of reality and the forecastability, the observation results are more often served as an input or benchmark data to validate and calibrate the hydrological and hydraulic models (Teng et al., 2017).

Instead of hydraulic modeling, a new branch has emerged in recent years for its simplicity and computing efficiency, which was labeled as the simplified conceptual models. This type of models is non-physics-based but are built on simplified hydrological concepts. The Rapid Flood Spreading Method (RFSM) model is one of the signature models in this group (Samuels and European Conference on Flood Risk Management Research Into Practice, 2009), which has a pre-processing stage to divide the flood plan into elementary depressing areas, and then uses a filling/spilling process to simulate the flood extent. Another popular model in this group is named Height Above the Next Drainage (HAND), which has been operating at a continental scale (Johnson et al., 2019; Nobre et al., 2011; Zhang et al., 2018). This model normalizes the topography based on the relative height of each cell in the drainage network, then simulate the flood extent using the "bathtub method".

The studies indicated that the conceptual model could approximate the final inundation distribution when there was no more excessive water input in the basin, and the computing time can be 1000 times faster comparing to hydraulic models (Néelz and Pender, 2013). This group of models is suitable for the applications with less demand of accuracy of flow dynamics but is a good approximation to the shallow water equation 2D model in the final inundation extent,

overbank water volume, and water depth. However, the conceptual models are not suitable for the cases of complex topographies, or modeling tsunamis, flash floods, or dam breaks types of flooding (Teng et al., 2017).

1.2.3 Hydraulic modeling approaches to high temporal-spatial resolution flood inundation

As another traditional approach for flood inundation mapping, the hydraulic models include one-dimensional, two-dimensional and three-dimensional methods that use the physical equations and laws to describe the fluid motion, where the degree of complexity varies. One-dimensional hydraulic model is considered the simplest representation of floodplain flow and it could simulate the open surface water flow with the assumption that the water flows in the same direction and the flow velocity is the average over the channel cross-section (Brunner, 2016). Generally, the one-dimensional model is to solve the St. Venant equation.

The two-dimensional models are believed to be the most widely implemented technologies for the flood extent simulation and risk estimation studies (Teng et al., 2017). The 2D models assume the water as a shallow ditch where no flow occurs vertically and solve the shallow water equation from the depth-averaging Navier-Stokes equations (Roberts et al., 2015).

In some special cases which need detailed information for engineering solutions, such as dam breaks, tsunamis or embankment failures, the 3D hydraulic models are implemented. However, for most of the floodplain analysis and simulation, the 2D shallow water approximation is considered adequate after proper model construction and validation (Alcrudo, 2004).

It is challenging to provide rapid and accurate simulation at high spatial-temporal resolution, due to the complex nature and the uncertainties of flooding (Rougier et al., 2013). Many efforts were made to overcome the challenge that the hydraulic models are enduring, to provide an

efficient floodplain simulation. One of the options is to utilize the advanced High-Performance Computation (HPC) tools and techniques, and another option is to utilize certain assumption to simplify the model, but in most cases, both methods were used. The studies showed that implementing the Graphics Processing Unit (GPU) could improve the computation speed by 100 times compared to the single core Central Processing Unit (CPU) 2D land surface runoff simulation using finite volume method, fully solved shallow water equation (Lacasta et al., 2014; Vacondio et al., 2014). Studies stated that the Message Passing Interface (MPI) parallel computing technique was mostly useful for large domain 2D land surface hydraulic modeling (Neal et al., 2009; Sampson et al., 2015), while the computation time approximately decreases linearly with the addition of the computational nodes (Roberts et al., 2015).

The fully solved 2D St. Venant equation for shallow water equation using the finite volume and the finite element methods were considered having the higher complexity among the 2D hydraulic modeling methods (Bates and De Roo, 2000). The storage cell or the cellular automata approach by solving Manning equation with finite difference methods was suggested to be a good approximation to the physical based model and the computation time was reduced by 30 times (Bates et al., 2010; de Almeida and Bates, 2013; Ghimire et al., 2013). It was further tested to prove the method was as efficient as other classes of models implementing HPC techniques (Bates et al., 2010; Sampson et al., 2015; Wing et al., 2017).

1.2.4 Coupling hydrologic and hydraulic modeling

Integrating hydrologic and hydraulic models has the benefit of utilizing the computation efficiency to model the hydrologic conditions over a long period and the flow dynamic representations when extreme hydrometeorological events occur (Anselmo et al., 1996). A recent study (Tanaka et al., 2018) has integrated a distributed hydrological model, Geomorphology-

Based Hydrological Model (GBHM), a 1D hydraulic model Mike11 and a 2D hydraulic model, Local Inertial Equation (LIE) model. The study found the integrated framework yields good agreement with the observation data of the stream discharge and the lake water level over four years span and was able to simulate a significant flooding event in 2000 over the study area. The author also indicated that the framework could simulate sediment movement downstream in the future research plan. The EF5 framework integrated the distributed hydrological model, CREST, and 1D hydraulic model, Kinematic Wave model, which successfully simulated multiple extreme precipitation events that caused flash flooding events in Oklahoma City and Houston at a continental scale implementation (Clark et al., 2017; Flamig et al., 2020; Gourley et al., 2017).

The National Water Center (NWC) took an effort on integrating WRF-Hydro and HAND at the national scale with 10 meters spatial resolution as the new National Water Model (NWM) (Cohen et al., 2018; Gochis et al., 2017; Johnson et al., 2019), as the agency was convinced that the HAND model performed good simulation of a 2016 Texas flooding event with good agreement with remote sensing observations and less computation cost (Zhang et al., 2018). However, based on the service assessment done by the National Water Service for 2017 Hurricane Harvey, the NWM simulation was not applied in the field operation due to the significant run-to-run variability and the consequential lack of confidence and distractions among the field operatives and agencies. The report suggested further improvement and communications (Murphy, 2018). The operational hydrologic models are currently still considered as academic projects but not an information vehicle that supports the public and government decision making.

1.3 Research objective

The overarching goal of this dissertation is to extend the capability of the existing multi-scale flood early warning system (e.g. CREST-EF5) and hydrological prediction of the terrestrial water by integrating two-dimensional hydrodynamic modeling and advanced numerical weather forecast products in the regional scale at the 10-m resolution.

1.4 Hypotheses

Based on the overarching objective, the following hypotheses are made in this dissertation:

A. The existing global flood early warning framework can successfully simulate the flooding events using both spaceborne and ground radar precipitation products in the one-dimensional fashion.

B. The 2D hydrologic & hydraulic coupled model can extend the capability of the existing flood early warning framework to simulate both riverine and terrestrial inundation extent and depth.

C. The flood predicting system can be enabled by the precipitation forecasts from Numerical Weather Prediction (NWP) products, and the uncertainties will be considerably neutralized through the model simulation.

1.5 Outline of the dissertation

This dissertation consists of five Chapters: the first Chapter is the introductive Chapter which describes the problem and raises the hypotheses, Chapters 2 to 4 are the three main Chapters followed by Chapter 5 which is an overall summary of this dissertation. Each chapter are relatively independent, however there are some repetitions in the content.

For the hydrological modeling and the water balance portion of the promoted hydrological and hydraulic coupled modeling is the CREST-EF5 (Flamig et al., 2020; Wang et al., 2011). The study area and all experiments are established in Harris County, TX and the Spring Basin locating at the northwestern region of Harris County. In Chapter 2, the MRMS products, the Global Precipitation Mission (GPM) Integrated Multi-satellitE Retrievals IMERG products, and the National Center of Environmental Prediction (NCEP) 4 km gridded precipitation field are evaluated by using CREST-EF5. In Chapter 3, the CREST-EF5 is coupled with a finite volume hydraulic model, Australia National University and Geophysics of Australia (ANUGA), which is added to the CREST modeling family and named CREST-iMAP. The CREST-iMAP simulated flood extent and flood depth are compared with other real-time operational flood monitoring maps in the world, such as the Radar Produced Inundation Diary (RAPID), National Water Model (NWM), and Fathom US, as well as the USGS survey being the benchmark. In Chapter 4, two NWP products, Rapid Refresh (RAP) and High Resolution Rapid Refresh (HRRR), as well as Artificial Intelligent (AI) based deep learning nowcast are used to simulate the flood flowrate, extent, and depth to examine the predictability of CREST-iMAP.

1.6 List of publications from the dissertation

Journal Publications:

Chapter 2

Chen, M.; Nabih, S.; Brauer, N.S.; Gao, S.; Gourley, J.J.; Hong, Z.; Kolar, R.L.; Hong, Y. Can Remote Sensing Technologies Capture the Extreme Precipitation Event and Its Cascading Hydrological Response? A Case Study of Hurricane Harvey Using EF5 Modeling Framework. *Remote Sensing*. 2020, 12(3), 445. Doi: 10.3390/rs12030445

Chapter 3

Chen, M.; Li, Z.; Luo, X.; Wing, O. L.; Gourley, J.J.; Kolar, R. L.; Hong, Y. A comprehensive flood inundation mapping for Hurricane Harvey using an integrated hydrological and hydraulic model. *Journal of Hydrometeorology*. 2021 Accepted for publication.

Chapter 4

Chen, M.; Gao, S.; Gourley, J. J.; Xue, M.; Kolar, R. L.; Hong, Y. A flood predictability study for Hurricane Harvey with the CREST-iMAP model using high resolution Quantitative Precipitation Forecasts and U-Net deep learning precipitation nowcasts. *Journal of Hydrology*. 2021 Accepted for publication.

Conference Presentations:

Chapter 2

Chen, M.; Luo, X.; Hong, Y. An efficient high-resolution hydrologic model for forecasting flood inundation at local to regional scales (poster). Oklahoma Transportation Research Day, Oklahoma City, Oklahoma, October **2018** 23rd

Chen, M.; Luo, X.; Hong, Y. An efficient high-resolution hydrologic model for forecasting flood inundation at local to regional scales (poster). 39th Oklahoma Governor's Water Conference and Research Symposium, Midwest City, Oklahoma, December **2018** 5- 6th

Chapter 3

Chen, M.; Luo, X.; Hong, Y. An Integrated Approach for a Real-time Forecasting and Risk Assessment of the Cascading Storm Triggered Flood Inundation (presentation). Beijing University 3rd Youth Remote Sensing and GIS Forum, Beijing, China, May **2019** 15-18th

Chen, M.; Luo, X.; Hong, Y. A Hydrologic and Hydraulic Modeling Approach for the Storm Triggered Cascading Flood Inundation (poster). CUAHSI Hydroinformatics Conference, Provo, Utah, July **2019** 29- 31st

Chen, M.; Luo, X.; Hong, Y. A Hydrologic and Hydraulic Modeling Approach for the Storm Triggered Cascading Flood Inundation (poster). OU International WaTER Conference, Norman, Oklahoma, September **2019** 16- 17th.

Chen, M.; Nabih, S.; Gao, S.; Gourley, J. J.; Kolar, R. L.; Hong, Y. A 2D hydrologic and hydraulic modeling approach for the storm triggered cascading flood inundation- a case study of Hurricane Harvey (Presentation). AGU Fall Meeting, San Francisco, California, December **2019** 9- 13th.

Chen, M.; Nabih, S.; Li, Z.; Luo, X.; Gao, S.; Gourley, J. J.; Wing, O.; Bates, P.; Shen, X.; Anagnostou, E.; Kolar, R. L.; Hong, Y. An Integrated Approach for a Real-Time Forecasting and Risk Assessment of the Cascading Extreme Storm Triggered Flood Inundation (Presentation). American Meteorological Society's 34th Conference on Hydrology, AMS 100th Annual Meeting, Boston, Massachusetts, January **2020** 12- 16th.

Chapter 4

Chen, M.; Li, Z.; Gao, S.; Xue, M.; Gourley, J.J.; Kolar, R.L.; Hong, Y. A Flood Inundation Prediction for Hurricane Harvey Using a Hyperresolution Hydrologic & Hydraulic Model Driven by the Remote Sensing Observations and Precipitation Forecasting Products (Interactive Poster). American Geophysics Union Fall Meeting, Virtual, December **2020** 1-17th.

Chen, M.; Li, Z.; Gao, S.; Xue, M.; Gourley, J.J.; Kolar, R.L.; Hong, Y. A case study of a hydrological and hydraulic coupled flood prediction for Hurricane Harvey using CREST-iMAP system (Presentation). American Meteorological Society's 35th Conference on Hydrology, AMS 101st Annual Meeting, Virtual, January **2021** 7 – 15th.

Reference

Abbott, M.B., Bathurst, J.C., Cunge, J.A., O'Connell, P.E., Rasmussen, J.V., 1986. An introduction to the European Hydrological System — Systeme Hydrologique European, "SHE", 1: History and philosophy of a physically-based, distributed modeling system. Journal of Hydrology 87, 45–59.

Alcrudo, F., 2004. A state of the art review on mathematical modelling of flood propagation.

Anselmo, V., Galeati, G., Palmieri, S., Rossi, U., Todini, E., 1996. Flood risk assessment using an integrated hydrological and hydraulic modelling approach: a case study. Journal of Hydrology 175, 533–554. https://doi.org/10.1016/S0022-1694(96)80023-0

Arnault, J., Wagner, S., Rummler, T., Fersch, B., Bliefernicht, J., Andresen, S., Kunstmann, H., 2016. Role of Runoff–Infiltration Partitioning and Resolved Overland Flow on Land–Atmosphere Feedbacks: A Case Study with the WRF-Hydro Coupled Modeling System for West Africa. Journal of Hydrometeorology 17, 1489–1516. https://doi.org/10.1175/JHM-D-15-0089.1

Ashley, S.T., Ashley, W.S., 2008. Flood fatalities in the United States. Journal of Applied Meteorology and Climatology 47, 805–817. https://doi.org/10.1175/2007JAMC1611.1.

Barthel, F., Neumayer, E., 2012. A trend analysis of normalized insured damage from natural disasters. Climatic Change 113, 215–237. https://doi.org/10.1007/s10584-011-0331-2

Bates, P.D., De Roo, A.P.J., 2000. A simple raster-based model for flood inundation simulation. Journal of Hydrology 236, 54–77. https://doi.org/10.1016/S0022-1694(00)00278-X

Bates, P.D., Horritt, M.S., Fewtrell, T.J., 2010. A simple inertial formulation of the shallow water equations for efficient two-dimensional flood inundation modelling. Journal of Hydrology 387, 33–45. https://doi.org/10.1016/j.jhydrol.2010.03.027

Beven, K.J., Kirkby, M.J., 1979. A physically based, variable contributing area model of basin hydrology / Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. Hydrological Sciences Bulletin 24, 43–69. https://doi.org/10.1080/02626667909491834

Brakenridge, G.R., Nghiem, S.V., Anderson, E., Mic, R., 2007. Orbital microwave measurement of river discharge and ice status: MICROWAVE MEASUREMENT OF RIVER DISCHARGE. Water Resour. Res. 43. https://doi.org/10.1029/2006WR005238

Brunner, G.W., 2016. HEC-RAS, River Analysis System Hydraulic Reference Manual (User Manual). US Army Corps of Engineers, Hydrologic Engineering Center, Davis, California.

Changnon, S.A., 2011. Temporal distribution of weather catastrophes in the USA. Climatic Change 106, 129–140. https://doi.org/10.1007/s10584-010-9927-1

Chen, H., Yang, D., Hong, Y., Gourley, J.J., Zhang, Y., 2013. Hydrological data assimilation with the Ensemble Square-Root-Filter: use of streamflow observations to update model states for real-time flash flood forecasting. Advances in Water Resources 59, 209–220.

Clark, R.A., Flamig, Z.L., Vergara, H., Hong, Y., Gourley, J.J., Mandl, D.J., Frye, S., Handy, M., Patterson, M., 2017. Hydrological Modeling and Capacity Building in the Republic of Namibia. Bulletin of the American Meteorological Society 98, 1697–1715. https://doi.org/10.1175/BAMS-D-15-00130.1

Cohen, S., Praskievicz, S., Maidment, D.R., 2018. Featured Collection Introduction: National Water Model. J Am Water Resour Assoc 54, 767–769. https://doi.org/10.1111/1752-1688.12664

Crawford, N.H., Linsley, R.K., 1966. Digital simulation in hydrology: Stanford watershed model IV (Technical report No. No. 39), Stanford University. Dept. of Civil Engineering. Stanford University, Stanford, California.

de Almeida, G.A.M., Bates, P., 2013. Applicability of the local inertial approximation of the shallow water equations to flood modeling: APPLICABILITY LOCAL INERTIAL. Water Resour. Res. 49, 4833–4844. https://doi.org/10.1002/wrcr.20366

Durand, M., Fu, L.-L., Lettenmaier, D.P., Alsdorf, D.E., Rodriguez, E., Esteban-Fernandez, D., 2010. The Surface Water and Ocean Topography Mission: Observing Terrestrial Surface Water and Oceanic Submesoscale Eddies. Proc. IEEE 98, 766–779. https://doi.org/10.1109/JPROC.2010.2043031

Dutta, D., Herath, S., Musiake, K., 2006. An application of a flood risk analysis system for impact analysis of a flood control plan in a river basin. Hydrological Processes 20, 1365–1384.

Dutta, D., Teng, J., Vaze, J., Lerat, J., Hughes, J., Marvanek, S., 2013. Storage-based approaches to build floodplain inundation modeling capability in river system models for water resources planning and accounting. Journal of Hydrology 504, 12–28.

Entekhabi, D., Njoku, E.G., O'Neill, P.E., Kellogg, K.H., Crow, W.T., Edelstein, W.N., Entin, J.K., Goodman, S.D., Jackson, T.J., Johnson, J., Kimball, J., Piepmeier, J.R., Koster, R.D., Martin, N., McDonald, K.C., Moghaddam, M., Moran, S., Reichle, R., Shi, J.C., Spencer, M.W., Thurman, S.W., Tsang, L., Van Zyl, J., 2010. The Soil Moisture Active Passive (SMAP) Mission. Proc. IEEE 98, 704–716. https://doi.org/10.1109/JPROC.2010.2043918

Flamig, Z.L., Vergara, H., Gourley, J.J., 2020. The Ensemble Framework For Flash Flood Forecasting (EF5) v1.2: Description and Case Study (preprint). Hydrology. https://doi.org/10.5194/gmd-2020-46

Gallegos, H.A., Schubert, J.E., Sanders, B.F., 2009. Two-dimensional, high-resolution modeling of urban dam-break flooding: A case study of Baldwin Hills, California. Advanced in Water Resources 32, 1323–1335.

Ghimire, B., Chen, A.S., Guidolin, M., Keedwell, E.C., Djordjević, S., Savić, D.A., 2013. Formulation of a fast 2D urban pluvial flood model using a cellular automata approach. Journal of Hydroinformatics 15, 676–686. https://doi.org/10.2166/hydro.2012.245

Gochis, D.J., Dugger, A.L., Yu, W., Yates, D.N., Sampson, K., Barlage, M., Pan, L., Zhang, Y., McCreight, J.L., RafieeiNasab, A., Karsten, L., Read, L., Gaydos, A., McAllister, M., Mills, J., Towler, E., Grim, J., FitzGerald, K., 2017. The NOAA National Water Model: Research to Operations to Research (Presentation). CUAHSI, Boulder, CO.

Gourley, J.J., Flamig, Z.L., Vergara, H., Kirstetter, P.-E., Clark, R.A., Argyle, E., Arthur, A., Martinaitis, S., Terti, G., Erlingis, J.M., Hong, Y., Howard, K.W., 2017. The FLASH Project: Improving the Tools for Flash Flood Monitoring and Prediction across the United States. Bulletin of the American Meteorological Society 98, 361–372. https://doi.org/10.1175/BAMS-D-15-00247.1

Hardy, R.J., Bates, P.D., Anderson, M.G., 2000. Modelling suspended sediment deposition on a fluvial floodplain using a two-dimensional dynamic finite element model. Journal of Hydrology 229, 202–218.

Heliere, F., Lin, C.C., Fois, F., Kern, M., Thompson, A., Bensi, P., 2009. Cold region hydrology high-resolution observatory (CoReH2O): A new microwave earth explorer core mission candidate, in: 2009 IEEE Radar Conference. Presented at the 2009 IEEE Radar Conference, IEEE, Pasadena, CA, USA, pp. 1–6. https://doi.org/10.1109/RADAR.2009.4977083

Hong, Y., Adler, R.F., Hossain, F., Curtis, S., Huffman, G.J., 2007a. A first approach to global runoff simulation using satellite rainfall estimation. Water Resources Research 43. https://doi.org/10.1029/2006WR005739.

Hong, Y., Gochis, D., Cheng, J.-T., Hsu, K., Sorooshian, S., 2007b. Evaluation of PERSIANN-CCS rainfall measurement using the NAME event rain gauge network. Journal of Hydrometeorology 8, 469–482. https://doi.org/10.1175/JHM574.1.

Johnson, J.M., Munasinghe, D., Eyelade, D., Cohen, S., 2019. An integrated evaluation of the National Water Model (NWM)–Height Above Nearest Drainage (HAND) flood mapping methodology. Nat. Hazards Earth Syst. Sci. 19, 2405–2420. https://doi.org/10.5194/nhess-19-2405-2019

Kundzewicz, Z.W., Pińskwar, I., Brakenridge, G.R., 2013. Large floods in Europe, 1985–2009. Hydrological Sciences Journal 58, 1–7. https://doi.org/10.1080/02626667.2012.745082

Lacasta, A., Morales-Hernández, M., Murillo, J., García-Navarro, P., 2014. An optimized GPU implementation of a 2D free surface simulation model on unstructured meshes. Advances in Engineering Software 78, 1–15. https://doi.org/10.1016/j.advengsoft.2014.08.007

Lin, P., Rajib, M.A., Yang, Z.-L., Somos-Valenzuela, M., Merwade, V., Maidment, D.R., Wang, Y., Chen, L., 2018. Spatiotemporal Evaluation of Simulated Evapotranspiration and Streamflow over Texas Using the WRF-Hydro-RAPID Modeling Framework. J Am Water Resour Assoc 54, 40–54. https://doi.org/10.1111/1752-1688.12585

Merz, B., Kreibich, H., Lall, U., 2013. Multi-variate flood damage assessment: a tree-based data-mining approach. Natural Hazards and Earth System Sciences 13, 53–64.

Murphy, J.D., 2018. Service assessment August–September 2017 Hurricane Harvey (US DOC). NOAA National Weather Service, Silver Spring, Maryland.

National Research Concil, 1991. Opportunities in the Hydrologic Sciences. National Academies Press, Washington, D.C. https://doi.org/10.17226/1543

Neal, J., Fewtrell, T., Trigg, M., 2009. Parallelisation of storage cell flood models using OpenMP. Environmental Modelling & Software 24, 872–877. https://doi.org/10.1016/j.envsoft.2008.12.004

Néelz, S., Pender, G., 2013. Benchmarking the latest generation of 2D hydraulic flood modelling packages (Techical report No. SC120002). Environment Agency, Bristol, UK.

Nobre, A.D., Cuartas, L.A., Hodnett, M., Rennó, C.D., Rodrigues, G., Silveira, A., Waterloo, M., Saleska, S., 2011. Height Above the Nearest Drainage – a hydrologically relevant new terrain model. Journal of Hydrology 404, 13–29. https://doi.org/10.1016/j.jhydrol.2011.03.051

Roberts, S.G., Nielson, O.M., Gray, D., Sexton, J., Davis, G., 2015. ANUGA User Manual 2.0. (Manual). Commonwealth of Australia (Geoscience Australia) and the Australian National University.

Rougier, J., Sparks, S., Hill, L.J., 2013. Risk and Uncertainty Assessment for Natural Hazards. Cambridge University Press, New York, New York.

Sachs, J., McArthur, J., 2005. The Millennium Project: a plan for meeting the Millennium Development Goals. The Lancet 365, 347–353. https://doi.org/10.1016/S0140-6736(05)17791-5

Sampson, C.C., Smith, A.M., Bates, P.D., Neal, J.C., Alfieri, L., Freer, J.E., 2015. A high-resolution global flood hazard model: A HIGH-RESOLUTION GLOBAL FLOOD HAZARD MODEL. Water Resour. Res. 51, 7358–7381. https://doi.org/10.1002/2015WR016954

Samuels, P., European Conference on Flood Risk Management Research Into Practice (Eds.), 2009. Flood risk management: research and practice ; proceedings of the European Conference on Flood Risk Management Research into Practice (FLOODrisk 2008), Oxford, UK, 30 September - 2 October 2008. Presented at the European Conference on Flood Risk Management Research into Practice, CRC Press, Boca Raton, Fla.

Shen, X., Anagnostou, E.N., Allen, G.H., Robert Brakenridge, G., Kettner, A.J., 2019a. Near-real-time non-obstructed flood inundation mapping using synthetic aperture radar. Remote Sensing of Environment 221, 302–315. https://doi.org/10.1016/j.rse.2018.11.008

Shen, X., Wang, D., Mao, K., Anagnostou, E., Hong, Y., 2019b. Inundation Extent Mapping by Synthetic Aperture Radar: A Review. Remote Sensing 11, 879. https://doi.org/10.3390/rs11070879

Spiker, E.C., Gori, P.L., 2000. National Landslide Hazards Mitigation Strategy A Framework for Loss Reduction (Open-File Report No. 00–450). U.S. Department of Interior, Reston, Virginia.

Sugawara, M., Watanabe, I., Ozaki, E., Katsugama, Y., 1984. Tank Model with Snow Component (Techical report No. 65). National Research Center for Disaster Prevention of Japan, Ibaraki-Ken, Japan.

Tanaka, T., Yoshioka, H., Siev, S., Fujii, H., Fujihara, Y., Hoshikawa, K., Ly, S., Yoshimura, C., 2018. An Integrated Hydrological-Hydraulic Model for Simulating Surface Water Flows of a Shallow Lake Surrounded by Large Floodplains. Water 10, 1213. https://doi.org/10.3390/w10091213

Teng, J., Jakeman, A.J., Vaze, J., Croke, B.F.W., Dutta, D., Kim, S.S.H., 2017. Flood inundation modeling: A review of methods, recent advances and uncertainty analysis. Environmental Modelling & Software 90, 201–216.

Vacondio, R., Dal Palù, A., Mignosa, P., 2014. GPU-enhanced Finite Volume Shallow Water solver for fast flood simulations. Environmental Modelling & Software 57, 60–75. https://doi.org/10.1016/j.envsoft.2014.02.003

Wang, J., Hong, Y., Li, L., Gourley, J.J., Khan, S.I., Yilmaz, K.K., Adler, R.F., Policelli, F.S., Habib, S., Irwn, D., Limaye, A.S., Korme, T., Okello, L., 2011. The coupled routing and excess storage (CREST) distributed hydrological model. Hydrological Sciences Journal 56, 84–98. https://doi.org/10.1080/02626667.2010.543087

Wigmosta, M.S., Nijssen, B., Storck, P., 1995. The Distributed Hydrology Soil Vegetation Model (Techical report). Pacific Northwest National Laboratory, Richland, Washingtong USA.

Williams, J.R., Hann, R.W., 1978. Optimal operation of large agricultural watersheds with water quality restraints (Techical report No. TR-96). Texas Water Resources Institute, Temple, Texas.

Wing, O.E.J., Bates, P.D., Sampson, C.C., Smith, A.M., Johnson, K.A., Erickson, T.A., 2017. Validation of a 30 m resolution flood hazard model of the conterminous United States: 30 m RESOLUTION FLOOD MODEL OF CONUS. Water Resour. Res. 53, 7968–7986. https://doi.org/10.1002/2017WR020917

Wood, E.F., Roundy, J.K., Troy, T.J., van Beek, L.P.H., Bierkens, M.F.P., Blyth, E., de Roo, A., Döll, P., Ek, M., Famiglietti, J., Gochis, D., van de Giesen, N., Houser, P., Jaffé, P.R., Kollet, S., Lehner, B., Lettenmaier, D.P., Peters-Lidard, C., Sivapalan, M., Sheffield, J., Wade, A., Whitehead, P., 2011. Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water: OPINION. Water Resour. Res. 47. https://doi.org/10.1029/2010WR010090

Zhang, J., Huang, Y.-F., Munasinghe, D., Fang, Z., Tsang, Y.-P., Cohen, S., 2018. Comparative Analysis of Inundation Mapping Approaches for the 2016 Flood in the Brazos River, Texas. J Am Water Resour Assoc 54, 820–833. https://doi.org/10.1111/1752-1688.12623

Zonta, R., Collavini, F., Zaggia, L., Zuliani, A., 2005. The effect of floods on the transport of suspended sediments and contaminants: A case study from the estuary of the Dese River (Venice Lagoon, Italy). Environment International 31, 948–958.

Chapter 2. Can Remote Sensing Technologies Capture the Extreme Precipitation Event and Its Cascading Hydrological Response? A Case Study of Hurricane Harvey Using EF5 Modeling Framework

Abstract

A new generation of precipitation measurement products has emerged, and their performances have gained much attention from the scientific community, such as the Multi-Radar Multi-Sensor system (MRMS) from the National Severe Storm Laboratory (NSSL) and the Global Precipitation Measurement Mission (GPM) from the National Aeronautics and Space Administration (NASA). This study statistically evaluated the MRMS and GPM products and investigated their cascading hydrological response in August of 2017, when Hurricane Harvey brought historical and record-breaking precipitation to the Gulf Coast (>1500 mm), causing 107 fatalities along with about USD 125 billion worth of damage. Rain-gauge observations from Harris County Flood Control District (HCFCD) and stream-gauge measurements by the United States Geological Survey (USGS) were used as ground truths to evaluate MRMS, GPM and National Centers for Environmental Prediction (NCEP) gauge-only data by using statistical metrics and hydrological simulations using the Ensemble Framework for Flash Flooding Forecast (EF5) model. The results indicate that remote sensing technologies can accurately detect and estimate the unprecedented precipitation event with their near-real-time products, and all precipitation products produced good hydrological simulations, where the Nash–Sutcliff model efficiency coefficients (NSCE) were close to 0.9 for both the MRMS and GPM products. With the timeliness and seamless coverage of MRMS and GPM, the study also demonstrated the capability and efficiency of the EF5 framework for flash flood modeling over the United States and potentially additional international domains.

2.1 Introduction

Floods are believed to be among the most hazardous and frequent natural disasters to human society (Ashley and Ashley, 2008; Barredo, 2007; Benito et al., 2004; Smith and Ward, 1998). Flooding can generally damage infrastructure, cost lives, and even cause further water contamination as well as waterborne diseases(Zhang et al., 2015). In particular, floods over urbanized area are more likely to cause fatalities and severe economic damage because of the population density and developed infrastructure, which leads to the intensification of the meteorological extremes (Nigussie and Altunkaynak, 2019; Zhang et al., 2018) and increased surface runoff peaks (Zhang et al., 2015). Globally, the Gulf Coast of North America is one of many places that is heavily affected by tropical storms and their cascading floods in an urbanized area (Adhikari et al., 2010). On August 25th, 2017, Hurricane Harvey made its first landfall at the northern end of San Jose Island, TX. Since then, Harvey stalled over the greater Houston area and produced over 1500 mm of rain in 4 days, which set the US record of total precipitation since the 1880s, when the reliable rainfall records started (Eric and Zelinsky, 2018). During this event, southeast Texas received 20 to 30 trillion tons of water with a return period exceeding 9000 years at some locations (van Oldenborgh et al., 2018), interconnected the Colorado River and San Bernard River overland, and caused unprecedented flooding. Hurricane Harvey was estimated to cause about USD 125 billion worth of damage and 107 fatalities, and 127 flash flood warnings were issued during the event (Murphy, 2018). As much as technology has advanced, society is still searching for tools to improve prediction and mitigate the damage from floods.

Over the past few decades, the scientific community has made great improvements in the capacity of flood modeling by combining climate models, weather models, hydrological models,

river models, and hydrodynamic models (Teng et al., 2017). The applications of flood modeling vary from flood risk assessment and mapping (Dutta et al., 2006) to flood damage assessment (Merz et al., 2013), real-time flood forecasting (Cohen et al., 2018; Flamig et al., 2020; Gourley et al., 2017; Sampson et al., 2015), engineering for flood prevention (Gallegos et al., 2009), post-flood river system hydrology (Dutta et al., 2013), soil and riverbank erosion (Hardy et al., 2000), catchment hydrology (Abbott et al., 1986) and floodplain ecology (Zonta et al., 2005). Generally, such applications require considering both the acceptable predictive accuracy and high spatiotemporal resolution while balancing the computational efficiency for real-time operations. For real-time simulations and forecasts, faster run times, and data assimilation are required to provide reliable results (Chen et al., 2013). Recent advances in global and continental hydrological models have shown very promising results thus far, which have provided critical information regarding surface runoff, streamflow, soil moisture, soil infiltration and evapotranspiration (Hong et al., 2007b; Kocalets et al., 2015; McAllister et al., 2018; Pasquier et al., 2019; Senatore et al., 2015; Werner et al., 2005; Wu and Johnston, 2007). One example is the application of the Coupled Routing and Excess Storage model embedded within the Ensemble Framework for Flash Flood Forecasting (CREST-EF5) framework (Clark et al., 2017; Flamig et al., 2020; Gourley et al., 2017; Wang et al., 2011). The CREST-EF5 framework integrates the distributed hydrological model, CREST, and 1D Kinematic Wave routing to simulate multiple excessive precipitation-triggered flash-flooding events in Oklahoma City and Houston at a continental scale (Gourley et al., 2017). As errors originally contained in the precipitation forcibly propagate through the hydrological model (Hong et al., 2006), the accuracy of precipitation datasets is also vitally crucial for hydrological modeling performance.

There are three common precipitation observations in the modern world—rain gauges, weather radars, and satellite-based remote sensing technologies (Li et al., 2013). Rain gauge measurement is traditionally the most straightforward in situ method to estimation the surface precipitation, which, for decades, has been regarded as the closest approximation to the true value at a point (Ciach and Krajewski, 1999; Hong et al., 2007a; Tang et al., 2016; Villarini et al., 2008). Many efforts have been made to interpolate the rain gauge data into a distributed precipitation field, and many versions of the optimal rainfall estimation procedures have been adapted by National Center for Environmental Prediction (NCEP) and National Weather Service (NWS) (Seo, 1998). However, the rain gauge network density varies spatially and is low over many developing regions (Huffman et al., 2001). Rather than estimating rainfall at a point in space, weather radar networks provide Quantitative Precipitation Estimates (QPE) covering much larger spatial domains (ranges up to 230 km), at spatial resolutions of the order of 1 km2 for each pixel. Since the 1990s, the Next Generation Weather Radar (NEXRAD) WSR-88D system has been improved and utilized, such as the dual-polarization capabilities in 2010 (Cifelli et al., 2011), for more advanced precipitation products including the MRMS system. The MRMS system integrates data from over 180 operational NEXRAD radars, over 7000 hourly rain gauges from the Hydrometeorology Automated Data System (HADS), the hourly High Resolution Rapid Refresh model analysis data and precipitation climatology(Hong and Gourley, 2015); it seamlessly covers the conterminous United States (CONUS) and Southern Canada at 1 km spatial resolution and a two minute temporal resolution using sophisticated algorithms and supplemental input data from ground gauges and environmental models (Zhang et al., 2016). However, the MRMS radar network still potentially suffers from radar miscalibration, reduced low-level coverage in mountainous areas and errors in the QPE algorithms. Earth observation

satellites provide the potential to estimate precipitation on a global scale (Chang and Hong, 2012). One such unprecedented effort, the NASA Global Precipitation Measurement (GPM) mission was launched in 2014 by building upon the success of previous Tropical Rainfall Measuring Mission (TRMM) from 1997 (Hou et al., 2014). To date, the GPM mission has used the Integrated Multisatellite Retrievals for GPM (IMERG) algorithm to generate the quasi-global precipitation products at 0.1 by 0.1 arc-degree spatial resolution and 30 minutes temporal resolution (Huffman et al., 2012, 2019a).

Today, given the availability of the above mentioned three precipitation data sources, the scientific community has exerted efforts on various precipitation product evaluations and intercomparisons, with particular foci over complex terrains and extreme events including the 2017 Hurricane Harvey event on the Mexico Gulf Coast (Eric and Zelinsky, 2018; Murphy, 2018; Pham et al., 2018; Thakur et al., 2018; Yang et al., 2019). Omaranian et al. (2019) compared the GPM IMERG final run precipitation estimates with NCEP Stage IV radar QPE and indicated that GPM IMERG could capture the pattern and trace the storm, but significantly overestimated the precipitation amount. Hayatbini et al. (2019) investigated the improved method for Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks- Cloud Classification System (PERSIANN-CCS) to match the cloud detection of GPM during Hurricane Harvey, which increases the possibility for PERSIANN-CCS to accurately detect extreme precipitation amount. Kao et al., (2019) again studied NCEP Stage IV QPE data for Hurricane Harvey using Probable Maximum Precipitation (PMP) estimation methods, and the study suggested a possible link between the extreme precipitation event and global climate change. Recently, a NASA report presented a precipitation estimations comparison between MRMS QPE and GPM IMERG products for Hurricane Harvey. The results indicate that GPM

IMERG had a coherent difference from MRMS QPE during Hurricane Harvey, where it underestimated precipitation in the storm core but overestimated it in the outer rainbands (Huffman et al., 2019b). All studies used different precipitation data as the ground truth and yielded various conclusions. Thus, it is necessary to evaluate the precipitation products based on high-density ground reference network and also further investigate how well they perform in a hydrologic context during this extreme event.

The overarching goal of this study is to investigate which precipitation product can better represent the true surface precipitation during the extreme event and further capture its cascading hydrological responses using a very high-density ground gauge network and an operational hydrological modeling framework. The specific objectives are to a) statistically compare interpolated rain gauge, MRMS QPE and GPM IMERG precipitation products with the local independently managed rain gauge precipitation observation in Harris County, TX and b) examine the corresponding hydrological response between interpolated rain gauge data, MRMS QPE and GPM IMERG products as forcing data for the hydrological simulation over the Spring Basin in the northern part of Harris County, TX. The study aims to answer the following research questions: 1) Is this type of extreme event detectable and quantifiable using remote sensing technology? 2) Can the hydrological model capture the extreme responses? 3) Which precipitation product performs better during extreme events?

This paper is organized as follows. Section 2.2 describes the study area, data used in this study, a short description of Ensemble Framework for Flash Flood Forecasting (EF5) hydrological model, and methodology. Section 2.3 inter-compares the MRMS, GPM IMERG, NCEP gauge-only precipitation products with Harris county rain gauge data and evaluates the

above precipitation products using the EF5 modeling framework. Section 2.4 concludes the study and proposes future directions.

2.2 Materials and methods

2.2.1 Study area

Figure 2.1 displays the impact area of Hurricane Harvey, Harris County, TX, and Spring Creek Basin. Despite the 79,000 square kilometers of impacted area by Hurricane Harvey, Harris County, TX was the most impacted, as almost half of the casualties from this event were from this area (Blake and Zelinsky, 2017). It is also the third most populated county in the USA, with 4.65 million people, and has an area of 4602 km2, where the dense population leads to the vulnerability to flood extremes and often relates to fatalities and significant economic loss. There are 147 rain gauges managed by the Harris County Flood Control District (HCFCD) and Hydrometeorological Automated Data System (HADS) from NWS in Harris County, which provided the 1-hour precipitation accumulation data for the study.

To further study the hydrological responses, Spring Basin was selected as the study watershed. Spring Basin is located at the northern end of Harris County, which contains four major rivers: Spring Creek, Willow Creek, Little Cypress Creek, and Cypress Creek, covering over 1960 km2 before entering Lake Houston. The elevation of Spring Basin ranges from 13 to 136 meters above sea level, with an average of 61 m. The slope varies from 0° to 39°, with an average of 1°. Spring Basin has urban structures concentrated at the southern and eastern part of the basin, agricultural ranches in the southwest portion, and forest in the northwest and middle of the basin along Spring Creek.

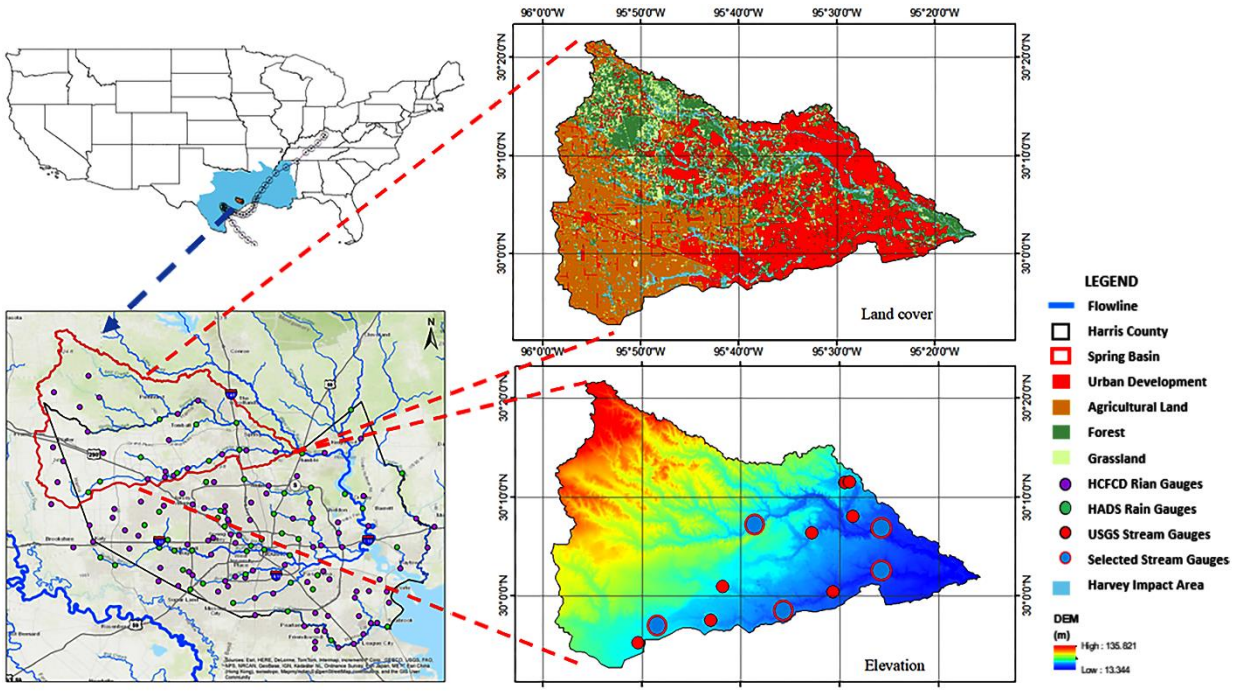


Figure 2.1. Study area showing Hurricane Harveyimpact (a), Harris County (b), Spring Basin, rain and stream gauges, and the landcover (c) as well as the topography (d) of Spring Basin.

2.2.2 Precipitation, stream flow, and geographic data

Five precipitation products (Table 2.1) were evaluated by the Harris County-managed HADS and HCFCD rain gauge observation data from 0:00 Central State Time (CDT) August 25th, 2017 to 23:00 CDT August 31st, 2017, with a total of 168 hourly time-steps. For the precipitation product evaluation, the ground rain-gauge should not overlap with those used by NSSL and NWS in developing or correcting the QPE estimates. Therefore, the HADS rain gauges (green dots in Figure 2.1) were removed from all analysis in this study. After quality control and HADS gauges removal, only 99 out of 147 rain gauge observations remained, where the gauges that had more than 60 out of 168 time-steps with continuous stationary values or NA values were eliminated. We obtained the NCEP and Environmental Modeling Center (EMC) national interpolated rain gauge-only hourly precipitation data from Earth Observing Laboratory (EOL) data archive (https://data.eol.ucar.edu/dataset/21.004) from August 15th to September 18th,

2017. This dataset utilizes measurements of about 3000 rain gauges across the CONUS (Kats, 2011) and uses the optimal estimation of rainfall fields methods to interpolate into 4 by 4 km gridded hourly precipitation data (Seo, 1998).

Two MRMS precipitation products from April 1st to September 30th, 2017, radar-based QPE (PCP_RATE), and 1-hour gauge bias-corrected radar precipitation accumulations (Q3GC_SHSR_1H), were obtained from the Iowa Environmental Mesonet NWS data archive (https://mesonet.agron.iastate.edu/nws/). PCP_RATE is the radar-based MRMS product that uses multiple R-Z relationships and is derived from MRMS Seamless Hybrid Scan Reflectivity (SGSR), which has a temporal resolution of 2 minutes and 0.01 by 0.01 arc-degree spatial resolution in real-time. Q3GC_SHSR_1H is the CoCoRaHS rain-gauge-corrected 1- hour radar QPE accumulation using a three steps method, which has the temporal resolution of 1 hour and 1 km2 spatial resolution with a 1.5-hour latency (Zhang et al., 2016).

Table 2.1. Summary of characteristics of the precipitation products in this study

ID	Products	Spatial Resolution	Temporal Resolution	Median (mm/hr)	Mean (mm/hr)	Maximum (mm/hr)
OBS	HCFCD rain gauge observation	Point data	1 hour	2.03	7.33	171.70
NCEP	NCEP hourly gauge only	4 km	1 hour	1.00	4.66	65.90
MRMS QPE	MRMS radar based QPE	0.01°	2 min	2.61	8.07	124.28
MRMS Corr	MRMS 1-hr gauge bias-corrected precipitation accumulation	0.01°	1 hour	1.90	6.70	105.60
V06AUncal	GPM IMERG satellite-based precipitation product	0.1°	30 min	3.82	6.55	66.80
V06ACal	GPM IMERG gauge calibrated precipitation product	0.1°	30 min	3.34	5.68	59.58

Two GPM IMERG Version 6 (V6) final-run precipitation products (Huffman et al., 2014) from August 15th to September 15th, 2017, PrecipitationUncal (V06AUncal), and Precipitation Cal (V06ACal), were obtained from the NASA GES DISC data archive

(https://disc.gsfc.nasa.gov/). The GPM IMERG system runs twice in near-real-time to produce early run and late run results, where the early run has the morphing scheme only propagated forward, and the late run has the morphing scheme applied both forward and backward (Huffman et al., 2019a). The IMERG final run has a 3.5-month latency, where the uncalibrated precipitation product (V06AUncal) is close to IMERG late run and then calibrated with the local rain gauge data to generate the calibrated precipitation product (V06ACal) (Huffman et al., 2019c). Both IMERG datasets have a 30-minute temporal resolution and 0.1 by 0.1 arc-degree spatial resolution.

Before any analysis, all precipitation products were aggregated or interpolated into 4 by 4 kilometers spatial resolution and hourly temporal resolution to make all data comparable. The MRMS product family has a higher spatial and temporal resolution, so MRMS QPE and MRMS Corr data were aggregated using an arithmetic mean and the 30-min data were summed to produce hourly precipitation accumulation. The GPM IMERG product family has lower spatial resolution, so V06AUncal, and V06ACal data were interpolated using the bilinear method and then aggregated into hourly time steps using an arithmetic mean.

Five U.S. Geological Survey (USGS) stream gauges, representing the upper, middle, and downstream branches of Spring and Cypress Creek, were selected to validate and calibrate the hydrological modeling process (Figure 2.1). The 15-min streamflow data of each gauge from April 1st to September 30th, 2017 were obtained from USGS National Water Information System (https://waterdata.usgs.gov/nwis). A high-resolution (15 arc second) hydrologically conditioned Digital Elevation Model (DEM), Flow Direction (FDR), Flow Accumulation (FAA), and major river network data were obtained from the HydroSHEDS (Lehner et al., 2008, https://www.hydrosheds.org/). The potential evapotranspiration (PET) data used in this study

were from USGS Famine Early Warning Systems Network (FEWS NET, https://earlywarning.usgs.gov/fews). The daily 1 by 1 arc-degree PET data were calculated from the Global Data Assimilation System (GDAS) using the Penman-Monteith method (Verdin et al., 2005). The US landcover data were obtained from Multi-Resolution Land Characteristics Consortium (MRLC, https://www.mrlc.gov/) and the 1 by 1 km EF5 parameters in CONUS were from the previous study by Vergara et al. (Vergara et al., 2016).

2.2.3 EF5 Modeling Framework and Hydrological Evaluation Method

EF5 is a framework built on multiple hydrological model cores including the Coupled Routing and Excess Storage (CREST) model version 2.0, co-developed by the University of Oklahoma and NASA Applied Science Team (Wang et al., 2011), and its grid-based water balance component is coupled with the kinematic wave water routing model (Chow et al., 1988). EF5 supports multiple water balance methods and comes with an automatic calibration module (Clark et al., 2017). EF5 was adapted as an operational tool across the NWS for flash flood forecasting by local NWS Forecast Offices in the Flooded Locations and Simulated Hydrographs Project (FLASH) (Gourley et al., 2017). The current modeling research was modified from the basic implementation from FLASH and was used here to evaluate the hydrological responses of remotely sensed observations and rain gauge interpolated precipitation products.

2.2.4 Statistical metrics

Seven common statistical metrics were used to evaluate the performances of different precipitation products and their performances in the hydrological model (Table 2.2). The correlation coefficient (CC) represents the degree of agreement between the precipitation estimates and the rain/stream gauge observation as the “ground truth.” Two metrics were selected to discover the error and bias between the precipitation products and observations,

which were the relative bias (RB) to describe the systematic bias as a ratio, and the root-mean-square error (RMSE) to measure the average error magnitude. Four additional metrics were calculated to evaluate the hydrological responses of different precipitation products, which were the conventional Nash–Sutcliffe coefficient of efficiency (NSCE), peak flow error (PE), peak time error (PTE), and runoff volume ratio (RR).

Table 2.2. List of statistical metrics used in this study

Statistic metrics	Equation[a]	Value range	Perfect value
Correlation coefficient (CC)	$CC = \frac{\sum_{n=1}^{N}(f_n - \bar{f})(r_n - \bar{r})}{\sqrt{\sum_{n=1}^{N}(f_n - \bar{f})^2}\sqrt{\sum_{n=1}^{N}(r_n - \bar{r})^2}}$	-1, 1	1
Relative bias (RB)	$RB = \frac{1}{N}\sum_{n=1}^{N}\frac{f_n - r_n}{r_n}$	-∞, +∞	0
Root-mean-square error (RMSE)	$RMSE = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(f_n - r_n)^2}$	0, +∞	0
Nash-Sutcliffe coefficient efficiency (NSCE)	$NSCE = 1 - \frac{\sum_{n=1}^{N}(f_n - r_n)^2}{\sum_{n=1}^{N}(r_n - \bar{r})^2}$	-∞, 1	1
Peak flow error (PE)	$PE = f_{max} - r_{max}$	-∞, +∞	0
Peak time error (PTE)	$PTE = t(r_{max}) - t(f_{max})$	-∞, +∞	0
Runoff volume ratio (RR)	$RR = \frac{\sum_{n=1}^{N} f_n}{\sum_{n=1}^{N} r_n}$	0, +∞	1

[a] Variables: n and N, sample index and a total number of samples, f represents the precipitation estimate products from gauge interpolation, radar, and satellite, r represents the reference observation including the HCFCD rain gauge and USGS stream gauge observations.

2.3 Results

2.3.1 Precipitation evaluation

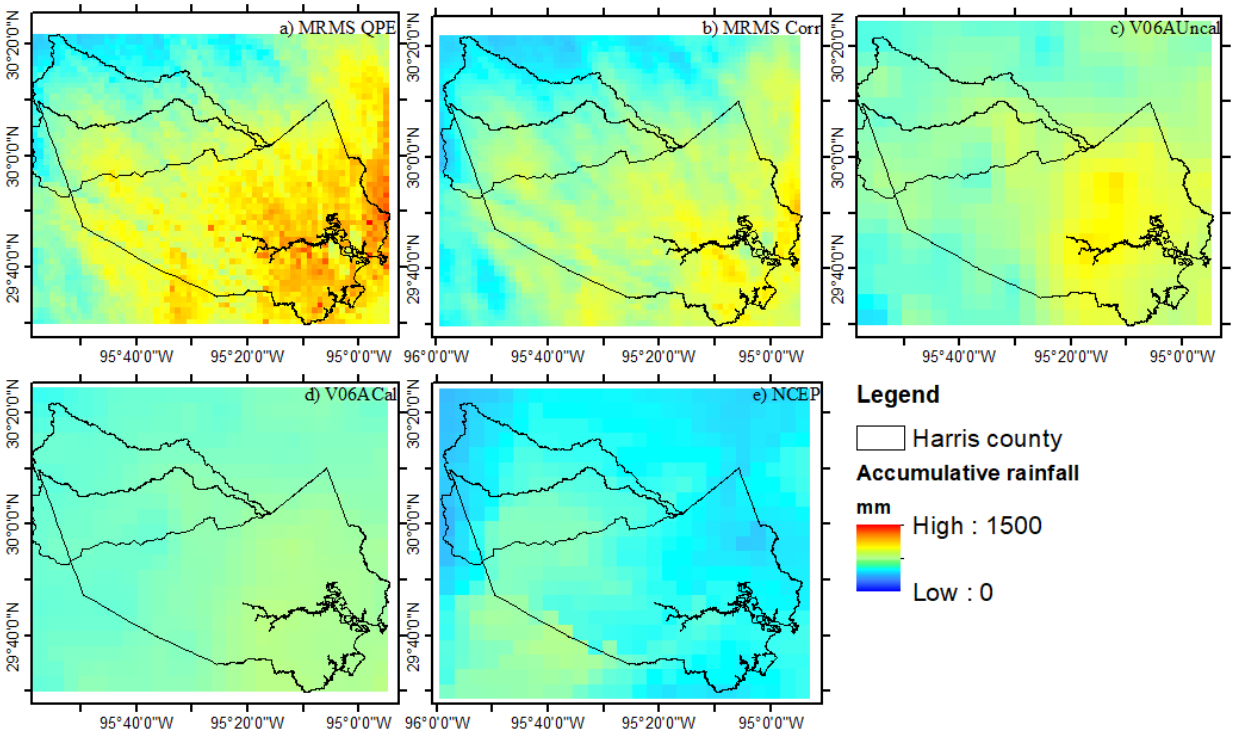


Figure 2.2. Accumulated precipitation from 25/08/2017 to 31/08/2017 during the Hurricane Harvey event in Harris County, TX. Hourly data from a) Multi-Radar Multi-Sensor (MRMS) quantitative precipitation estimates (QPE), b) MRMS Corr, c) V06AUncal, d) V06ACal, and e) NCEP gauge-only

Two series of comparisons were performed to evaluate the precipitation products, county-averaged analysis, and grid-based analysis, where the statistic results are listed in Table 3. The precipitation accumulation of Hurricane Harvey from all precipitation products is shown in Figure 2. This figure demonstrates that most precipitation products agreed that the southeastern part of Harris county received the highest precipitation amount, and the post-real-time corrections reduced the amount of precipitation for both MRMS and IMERG. In addition, V06ACal not only reduced the peak precipitation accumulation but also increased the minimum precipitation accumulation from V06AUncal. Contradictory to the majority agreement, the

NCEP gauge-only product shows the lower accumulative rainfall amount closer to the inner core of Hurricane Harvey (Figure 2.2).

The county-averaged comparisons were conducted by averaging the precipitation data from different products over the whole Harris county and Spring Basin area. Then, we averaged all 99 valid local HCFCD rain gauge data as the reference of the truth and calculated the statistics. For the grid-based comparisons and evaluations, we extracted all hourly precipitation rate data for the grids in which the 99 rain gauges were located from all five precipitation products and then calculated the statistics. Both local rain gauge corrections reduced the bias of MRMS Corr and V06ACal from positive 20% to negative 7% (positive 20% to negative 11% for grid-based) and from 56% to 33% (80% to 58% for grid-based), respectively, as shown in Table 2.3. However, the corrections only made minimal reductions in the RMSE, which could be attributed to the positive/negative biases being offset. The post-real-time correction of MRMS products has a slight overcorrection (Table 2.3 and Figure 2.3). In addition, the correlation coefficient between V06AUncal and V06ACal is exactly 1, possibly because GPM IMERG uses a fairly simple algorithm to calibrate data (Huffman et al., 2019c). Based on the first level statistical analysis, the MRMS product family has the highest correlation coefficient and the lowest RMSE compared with the local rain gauge observation, then NCEP gauge-only precipitation stands the second closest, and the GPM IMERG product family is the least accurate. The NCEP precipitation product has a small relative bias value in the grid-based analysis but much greater negative bias in the county-averaged analysis, which is further investigated in this study.

Table 2.3. Summary of the statistical evaluations of county-averaged and grid-based comparison of precipitation products at 4 km and hourly resolution

	County-Averaged Statistics			Grid-Based Statistics		
Precipitation Product ID	CC	RB (%)	RMSE (mm/hr)	CC	RB (%)	RMSE (mm/hr)
MRMS QPE	0.92	19.57	3.14	0.91	20.09	5.75
MRMS Corr	0.92	−7.18	3.04	0.93	-10.59	4.74
V06AUncal	0.79	55.97	5.02	0.45	80.45	11.98
V06ACal	0.79	32.80	5.47	0.45	57.61	11.87
NCEP	0.81	-28.88	5.57	0.61	3.48	10.84

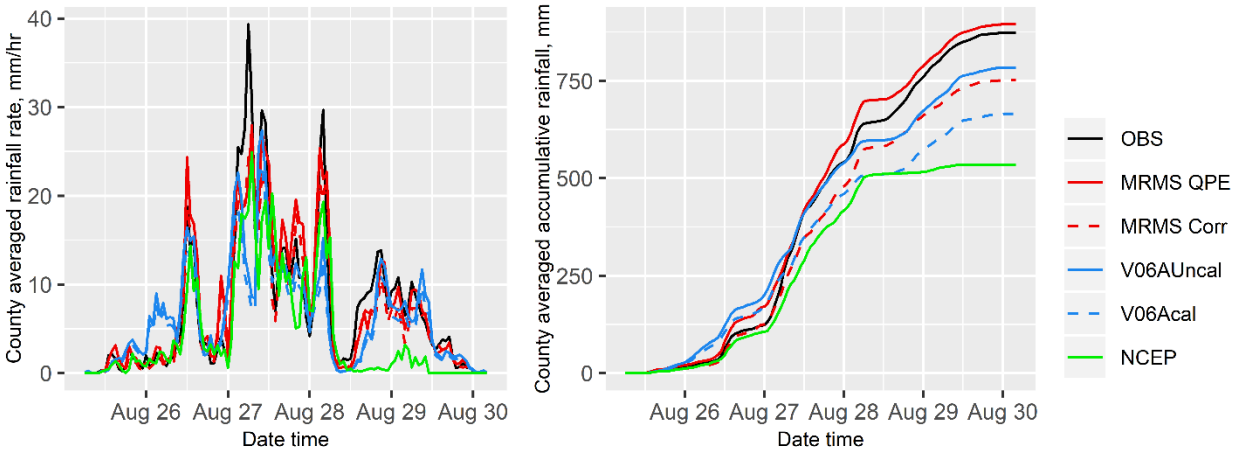


Figure 2.3. County averaged precipitation rate (left) and accumulative rainfall (right), where the MRMS family is in red (solid and dash), Global Precipitation Measurement Mission (GPM) family is in blue (solid and dash), National Centers for Environmental Prediction (NCEP) is in green, and HCFCD is in black

The county-averaged hourly rainfall rate and accumulated rainfall are shown in Figure 3. The MRMS product family has a strong correlation with the OBS (CC = 0.92), which caught most of the precipitation peaks, except for the highest one. IMERG V06 data overestimated at low precipitation intensities but underestimated at high precipitation intensities, which is consistent with the findings in the recent technical report (Huffman et al., 2019b). NCEP has an obvious unresponsive condition after the midday of 28th August 2017, which could be caused by the malfunctions of rain gauges that the algorithm utilized. This could possibly be the cause of the low precipitation accumulation of NCEP (Figure 2.2 and Figure 2.3). Generally speaking, the remotely sensed precipitation products (MRMS and IMERG) performed better than rain gauge interpolated product during Hurricane Harvey, as NCEP had a greater RMSE (Table 2.3). Before

the peak rainfall arrived in Harris county on early 08/27/2017, the OBS (HCFCD rain gauge observation) data had the best match with MRMS Corr, then NCEP, MRMS QPE, V06ACal, and V06AUncal was the last. After the peak rainfall, the ranking changed to MRMS QPE, V06AUncal, MRMS Corr, V06ACal, and NCEP. The uncalibrated precipitation products (MRMS QPE and V06AUncal) performed better than calibrated products (MRMS Corr and V06ACal) after the rainfall intensity picked up.

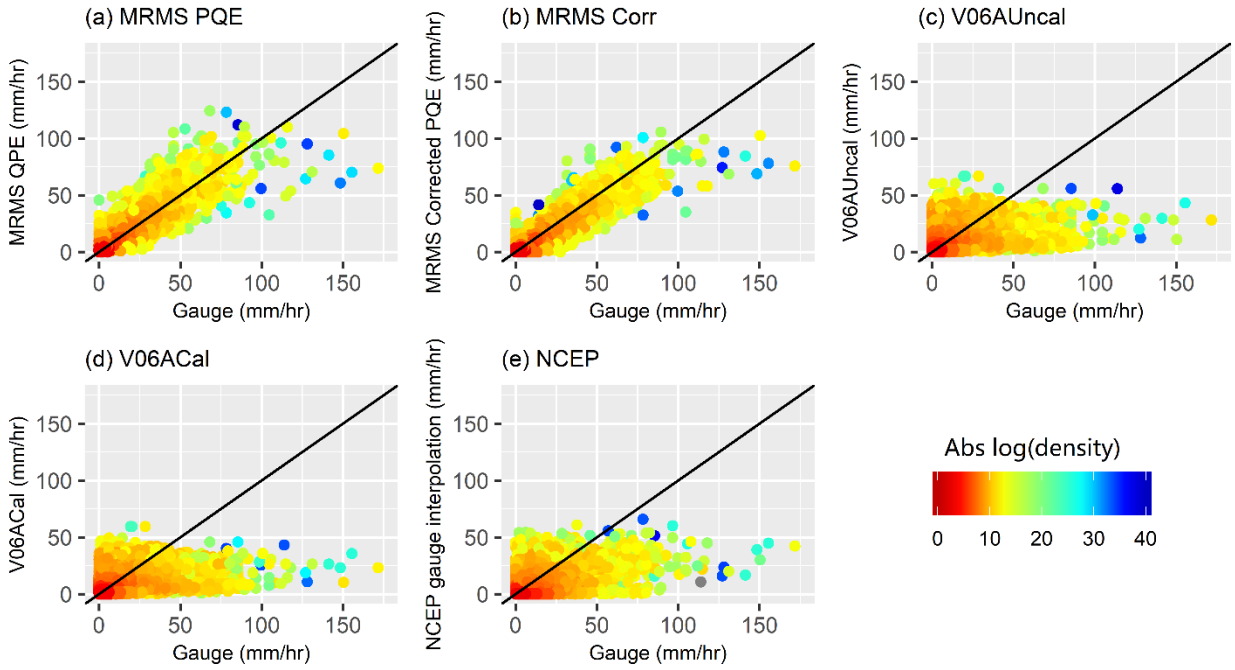


Figure 2.4. Grid-scale evaluation of hourly precipitation at 99 extracted 4 km grid cells between precipitation products and the rain gauge observations. Data from a) MRMS QPE, b) MRMS Corr, c) V06AUncal, d) V06ACal, and e) NCEP

As shown in Figure 2.4, the grid-based scatter plot has a similar finding as to the county-averaged analysis, where the MRMS product family performed the best during Hurricane Harvey. The NCEP product performed slightly better than the GPM IMERG product family. This difference indicates that rain gauge interpolation could potentially create large errors or patchiness that reduce the accuracy of the precipitation estimates over large spatial extents. The mean relative bias mathematically tends to generate a large positive value when the observation

value is small during overestimations, as the lowest possible value of relative bias is −1 for positive datasets such as precipitation. This mathematical nature resulted in the large positive bias of GPM IMERG products due to their overestimation during the low-intensity precipitation period (Figure 2.3).

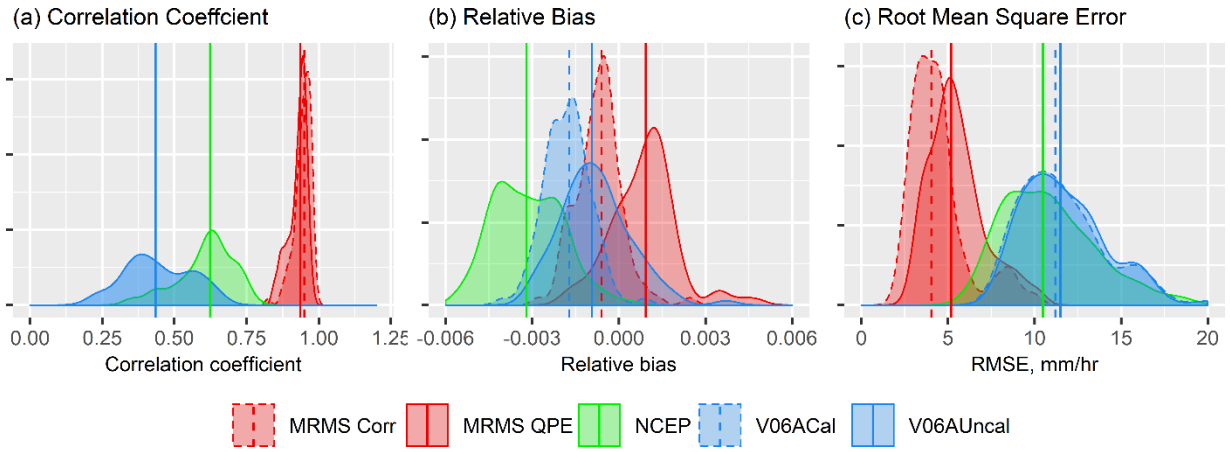


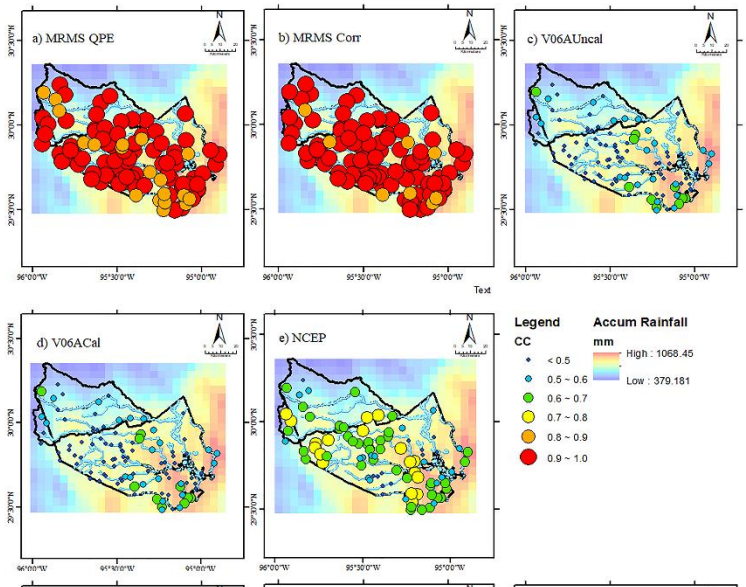
Figure 2.5. The distribution of the grid-scale evaluation statistics: a) CC, b) RB, and c) RMSE within 95% confidence interval and the vertical lines indicate the median of data

If focusing on the statistical results within the 95% confidence interval of all 99 grid points (Figure 2.5), one can find that the pattern differs where all precipitation products underestimated the precipitation rate during Harvey except for MRMS QPE. The MRMS Corrected product had a slight underestimation, but it was the closest to the perfect value. As shown in Figure 2.6, the GPM IMERG products had little temporal agreement with OBS in almost all 99 sites, and most of their higher RMSE values were concentrated close to the storm core, which indicates the accuracy of IMERG decreases as the precipitation intensity increases. NCEP had the same high RMSE concentration close to the storm core, possibly due to rain gauge malfunction. Even though the positive and negative bias offset gave the NCEP gauge-only product the best the average bias value (Table 2.3), the data quality is spatially inconsistent across the county.

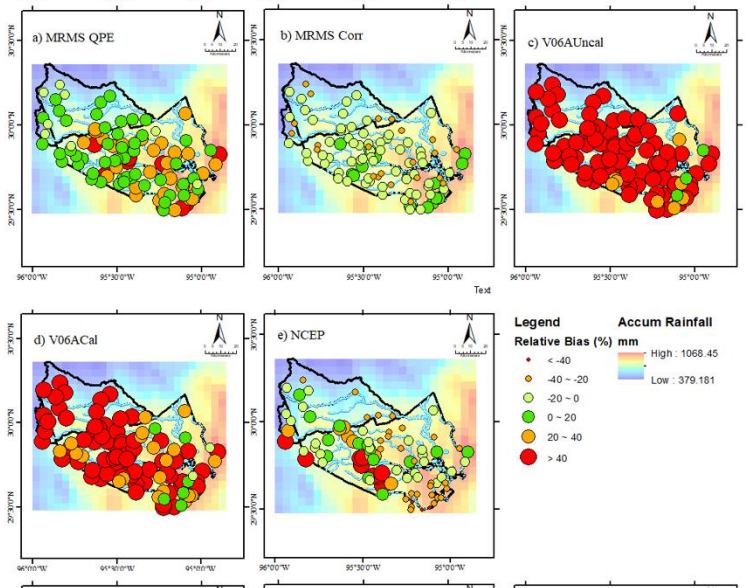
Among all the precipitation products, MRMS products shows the lowest RMSE and highest CC; however, the error has minor increases near the storm core.

In general, MRMS precipitation products show good agreement with HCFCD local rain-gauge observations, while IMERG and NCEP gauge-only products are comparable in quality (Table 2.3). This can be attributed to the following factors: first, Hurricane Harvey was an unprecedented precipitation event that potentially caused failures in the ground instruments that could impact methods relying on rain gauge interpolations and bias adjustments. Second, the GPM multi-satellite algorithm relies on passive sensors from satellites in low-earth and geostationary orbits to obtain high temporal resolution, while the radar QPE algorithm enjoys close proximity to the event and active sensors. Overall, MRMS products can best represent the surface precipitation field according to HCFCD rain gauge comparison, with high correlation coefficient (> 0.9) and low RMSE (~3 and ~5 mm per hour) at the county and grid-scale and all the post-real-time correction showed overcorrection for Hurricane Harvey.

CC:



Relative Bias:



RMSE:

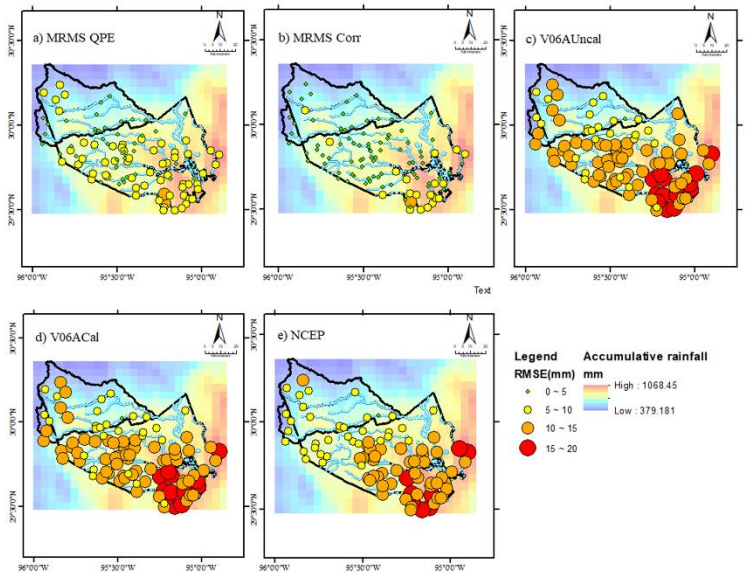


Figure 2.6. The grid-based statistic spatial distribution of different precipitation products during Hurricane Harvey

2.3.2 Hydrological evaluation

The evaluation of hydrological response using MRMS QPE, MRMS Corr, V06AUncal, V06ACal and NCEP products was carried out in the northwestern basin of Harris county, where all products had generally acceptable performances according to the previous section. Two rivers were studied using the EF5 modeling framework in this section (Table 2.4).

Table 2.4. Summary of stream gauges selections and calibration results

Stream Name	Location	USGS ID	Drainage Area (km^2)	Overbank Flow in Harvey?	Calibration Results			
					NSCE	CC	RB (%)	RMSE (m^3/s)
Spring	Midstream	08068275	483	No	0.99	0.94	8.78	21.23
Spring	Downstream	08068500	1059	No	0.99	0.91	-8.34	19.99
Cypress	Upperstream	08068720	280	No	0.90	0.90	34.64	5.15
Cypress	Midstream	08068800	540	Yes	0.91	0.87	7.36	4.93
Cypress	Downstream	08069000	738	Yes	0.95	0.97	-3.53	20.15

The first river, Spring Creek, has a larger channel network with a total of 179 km of open channels and a large natural floodplain. The Spring Creek watershed remains mostly underdeveloped and natural, except for downstream, where the Woodland Township and the city of Tomball are located. Due to the lack of urban development, the flood risk of Spring Creek is comparatively less, and there were no signs of overbank flow from two USGS stream gauges on this river during Harvey. This river was studied to evaluate the performance of different hydrological responses of MRMS, GPM IMERG, and NCEP precipitation products.

The second river, Cypress Creek, is a smaller river with 137 km of open water channel and a well-developed drainage area. The middle and downstream portion of Cypress Creek has experienced intensive urbanization in the past 20 to 30 years, hosting a population greater than 350,000 according to the 2010 U.S. census, while only the upstream area remains as agricultural land (Figure 2.1). Perhaps exacerbated by the impacts of urbanization, both the middle and

downstream USGS gauges showed overbank flow during Harvey (Figure 2.9), despite the effort of a 200,000 m^2 detention basin and multiple detention ponds located around the upper stream of Cypress Creek. This river was studied to examine the early warning capability of the coupled hydrological modeling system when the ground instruments were damaged or malfunctioning during the extreme event.

Table 2.5. Summary of hydrological simulation of Spring Creek

	Spring Creek 08068275					Spring Creek 08068500				
	MRMS QPE	**V06A Uncal**	**MRMS Corr**	**V06ACal**	**NCEP**	**MRMS QPE**	**V06A Uncal**	**MRMS Corr**	**V06A Cal**	**NCEP**
NSCE	0.62	0.9	0.81	0.83	0.84	0.98	0.88	0.97	0.91	0.93
RB (%)	50.64	23.69	33.4	42.03	7.93	2.9	-7.96	-1.64	8.25	-15
CC	0.96	0.96	0.97	0.96	0.94	0.99	0.96	0.99	0.96	0.98
RMSE (m^3/s)	167	84	116	111	107	59	161	86	144	125
Peak Error (m^3/s)	1066	97	736	326	392	-298	-787	-475	-581	-542
Peak Time Error (min)	33	-28	31	-28	-52	13	40	12	40	27
RR	1.51	1.24	1.33	1.42	1.08	1.03	0.92	0.98	1.08	0.85

First, the EF5 model was forced by MRMS Corrected precipitation data, considered as the most accurate data source during non-extreme events, to benchmark the model parameters from 1st April 2017 to 15th August 2017, with two 15-minute USGS streamflow observations (midstream and downstream of Spring Creek) using the Differential Evolution Adaptive Metropolis (DREAM) algorithm (Vrugt et al., 2009). The model then proceeded to warm up for the same period and the same forcing precipitation data (MRMS Corr) was used during the benchmark generating process to create a uniform initial condition for all precipitation product evaluations. Table 4 displays good agreement between simulated streamflow and the observations during the calibration period, with very high NSCE (0.99) and CC (~ 0.93), as well as a small RMSE (~20 m^3/s), slight overestimation at midstream (8.78%) and underestimation at downstream (-8.34%). The MRMS, GPM IMERG, and NCEP rain gauge interpolated precipitation products were then set as the model forcing data to simulate the hydrological

responses from 15/08/2017 to 15/09/2017 separately for Hurricane Harvey. All products were read at their original spatial and temporal resolution by the model, as listed in Table 2.1.

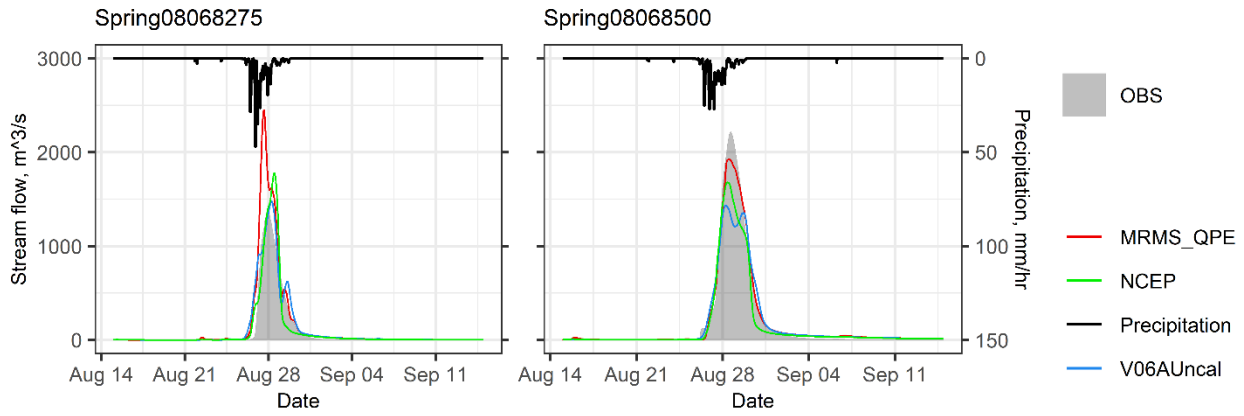


Figure 2.7. Comparison of EF5 simulated streamflow with USGS stream gauge observation during Hurricane Harvey at midstream (left) and downstream (right) of Spring Creek

In the hydrological simulations, all precipitation products generally had acceptable performances with a NSCE greater than 0.8, except for one, and a high CC (Table 2.5). All products overestimated the streamflow at the midstream of Spring Creek, and MRMS QPE had a large spike of simulated streamflow at the peak due to data sensitivity (Figure 2.7).

Moreover, the total runoff ratio (RR) of MRMS QPE at the midstream was large (Table 2.5), which could be due to the error propagation from the MRMS QPE as its errors were mostly located at the upstream of Spring Creek (Figure 2.6), which could partially cause the spike as well. All products yielded reasonable simulation hydrographs that could capture the flow peak with high CCs (> 0.94), and MRMS products were able to simulate the peak with over 30 minutes of lead time, which is crucial and beneficial for flood early warnings. At the downstream gauge, all products underestimated the streamflow peak and MRMS Corrected, V06AUncal and NCEP products had negative bias compared to the USGS stream gauge data. One observation of the simulation statistics is that the near-real-time remote sensing precipitation products performed almost equally well as the post-real-time corrected products, where the differences of

NSCE (< 0.1), bias (< 0.2), and RMSE (< 50 m3/s) between the two groups were minimal (Table 2.5). We can conclude that MRMS QPE and IMERG V06AUncal (equivalent to IMERG late run) are sufficient to drive hydrological model to provide flood warning information. As the near-real-time products can provide timeliness, the remote sensing technologies can significantly increase the accuracy and reliability of global flood early warning systems. As shown in Figure 2.7 and the statistic metrics in Table 2.5, the MRMS product family performed best if not considering the sensitivity effect in the midstream, followed by NCEP gauge only and GPM IMERG, according to the metrics of NSCE, RB, CC, and RMSE.

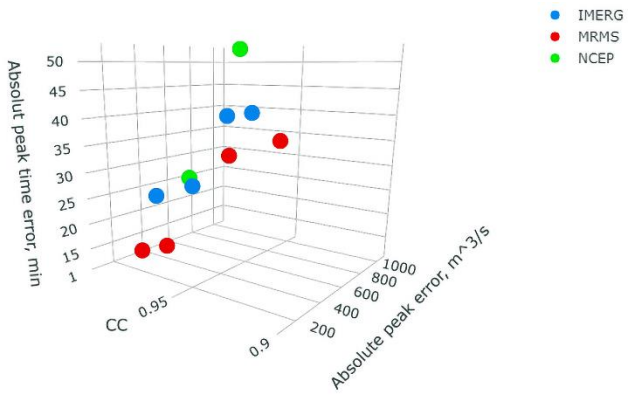


Figure 2.8. The 3-dimensional scatter plot using CC, PE, and PTE as variables for all three precipitation product families. The perfect point is the left lower corner

In Figure 2.8, the three relatively independent statistics are plotted in a 3-dimensional scatter plot, where the lower-left corner is the perfect point with the correlation coefficient of 1 and no flow peak nor peak time error. As shown in Figure 2.8, the MRMS products are closer to the perfect point overall, followed by GPM IMERG and then NCEP gauge-only, indicating that the performance of GPM IMERG and NCEP gauge-only are comparable for Hurricane Harvey in Spring Basin. Since products that are not based on rain gauge corrections show hydrological performance superior to the gauge-forced product within Spring Basin, we can infer that the gauge-interpolated precipitation product is reliable only when the gauge network is well

functioning. Downstream of Spring Creek, the NCEP gauge-only product yielded a hydrograph which had significant deficit in total runoff (RR=0.85), which shows the disadvantage of using the NCEP product as the underestimation occurred near the storm core and this error propagated to the hydrological simulation.

In summary, MRMS performed the best in capturing the hydrological response in Spring Creek, compared to USGS stream gauge observations during Hurricane Harvey, followed by GPM IMERG and NCEP gauge-only precipitation products which had comparable hydrological responses. The post-real-time corrected remote sensing precipitation products did not provide significant improvement in hydrological responses, which justifies the global real-time operational flood warning system based on the near-real-time products.

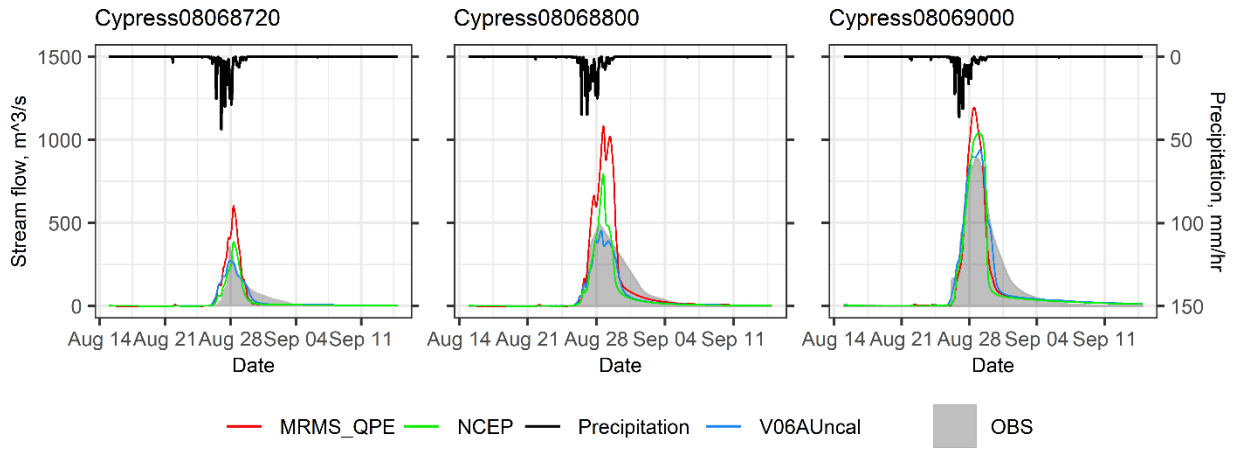


Figure 2.9. Comparison of EF5 simulated streamflow with USGS stream gauges during Hurricane Harvey at upstream (left), midstream (middle), and downstream (right) of Cypress Creek

For Cypress Creek, the EF5 model was calibrated and warmed up using the same methodology as with Spring Creek, enabling comparison with three USGS stream gauges located at upper, middle and downstream reaches. Table 2.4 shows the statistic results during the warm-up period at Cypress Creek stream gauges, which have high NSCE (> 0.9) and CC (> 0.87) values for all gauge locations during non-extreme periods, which indicates that the EF5

simulation had good agreement with the stream gauge observations. The CREST-EF5 simulation has a tendency towards slight overestimation at the upstream and gradually changed to underestimation at the downstream, which matches the findings of the previous study (Xue et al., 2016).

Figure 2.9 shows the hydrographs at three stream gauges along Cypress Creek and EF5 simulated results using MRMS, IMERG, and NCEP products during Hurricane Harvey. At the upstream location (USGS 08068720), the observed hydrograph has an obvious long and gentle receding limb, which was caused by the 650,000 m^2 Warren Reservoir located 6.4 km ahead of the upstream gauge. The slow-release and regulation of the water caused the long receding limb and was not considered in this study. Since the Cypress Creek drainage area has been under rapid development during the past three decades, HCFCD reported that overbank flow was the common problem of the region and extensive water detention-related engineering jobs were completed. Additionally, as shown in Figure 9, both the mid and downstream observed hydrographs also have longer receding limbs. However, during this intense event, the engineered infrastructure did not prevent the overbank flow and flooding in the middle and downstream of Cypress Creek, where both observed hydrographs (USGS 08068800 and 08069000) have an unnatural plateau. Furthermore, the linear appearance of the rising and falling limbs at the midstream gauge was a result of no data being recorded from noon 27/08/2019 until 01/09/2019 due to gauge malfunction.

Due to the lack of information on Warren Reservoir and its operational details during Hurricane Harvey, EF5 simulation results could not capture the reduction in peak flow magnitude nor the long receding limb from the slow release of detained stormwater. In Figure 2.9, all simulated hydrographs have a very steep receding limb, which represents a more typical

hydrological response in urban areas. Therefore, it is inferred that EF5 simulated the urban hydrological conditions without the reservoir interferences, which is sufficient for flood conditions in the related warning scenarios. For the upstream gauge (USGS 08068720), MRMS QPE could closely capture the first peak almost at the same time as the stream gauge observation, and the mismatch of the second peak could be caused by the structural flood control. IMERG performed poorly for Cypress Creek, as its simulated hydrograph was rather "smooth" when not considering structural flood control, unlike typical urban hydrographs. At this point, we can conclude that MRMS products performed the best in the urban hydrological condition, and we can use MRMS products to regenerate a well-approximated hydrograph when the stream gauge is damaged or over bank flow occurs. The simulated streamflow could be utilized for many applications, including but not limited to flood warning, flood risk analysis, flood inundation calculation, and flood control performance assessment.

In summary, the EF5 modeling framework can simulate close-to-reality streamflow using MRMS precipitation products during the extreme precipitation event, especially when the ground stream gauges are damaged, or no gauge is implemented, or when overbank flow occurs. IMERG products might not be suitable for simulating extreme events in small basins (e.g., < 200 km^2), but given its global coverage, it is sufficient for major rivers and sub-basins, where most CREST and satellite precipitation studies were found successful when applied to different basins globally (Huang et al., 2014; Tang et al., 2016; Zhang et al., 2015). Overall, the EF5 modeling framework combined with current remote sensing technologies (IMERG and MRMS) can provide a robust flood-early-warning system for real-time operational uses at the regional, continental, and even global scale.

2.4 Discussion

Three results from the previous section require further explanation. First, the result in Figure 2.3 indicates that after the peak rainfall intensity occurred, the uncalibrated precipitation products (MRMS QPE and V06AUncal) outperformed the calibrated precipitation products (MRMS Corr and V06ACal). This might be caused by a combination of overcorrection of the underestimation of GPM IMERG during the high-intensity precipitation period, possible rain gauge malfunctions, and post-real-time calibration error. Due to the unprecedented nature of Hurricane Harvey, it could cause errors to the post-real-time correction algorithms at such an intensity. Therefore, the current precipitation correction algorithms for MRMS and GPM IMERG are less valid for unprecedented events like Hurricane Harvey. We also suspect that the insensitive precipitation observation from the NCEP data is caused by the damage to instruments by the flood or objects and the mechanical saturations caused by the intensive rainfall. However, the true causes of the suspected rain gauge malfunction are unknown.

Second, the results from Figure 2.7 show a spike on the MRMS QPE simulated hydrograph. This could be caused by the sensitivity to the high quality and high temporal resolution of MRMS QPE product with the combination of smaller channels in the upstream. When simulated with MRMS Corr, the resulting streamflow had a dramatically smoother spike as the temporal resolution is hourly.

Third, the Cypress Creek hydrological evaluation yields an untypical hydrograph at the upper stream from IMERG produced simulation. However, as stated in the previous section, EF5 did not consider the impact of flood control infrastructure. This result could be caused by the low spatial resolution of IMERG combining with the smaller drainage area which undermined the precipitation representativeness. It was a fact that there were less than 4 grid cells from IMERG

products within the drainage area of the upstream gauge. At the downstream location, the IMERG yielded a more reasonable hydrograph, as there were 12 grid cells covering the drainage area.

2.5 Conclusions

The results of this study indicate that the remote sensing technologies and gauge interpolation method could all detect the unprecedented extreme rainfall associated with Hurricane Harvey, as well as capturing the cascading hydrological responses. This study first focused on statistically comparing the MRMS QPE, MRMS Corr, IMERG final V06AUncal and V06ACal, as well as NCEP gauge-only interpolated precipitation products with the very dense HCFCD local rain gauges. Then, hydrological responses were evaluated using the EF5 modeling framework in Harris County and Spring Basin of Texas, USA. The findings and results from this study can be potentially applicable to other subtropical zones impacted by tropical cyclones or low-lying flood-prone areas that are similar to southeast Texas, particularly in extreme events.

The main conclusions from the cross-evaluation of MRMS, IMERG, and NCEP gauge-only precipitation products at county- and grid-based scales during Harvey, are summarized below:

1. MRMS precipitation products are the best remote sensing rainfall measurements that perform most comparably to the local dense network rain gauge observations. IMERG and NCEP rain gauge-interpolated precipitation products are comparable to each other statistically, but IMERG has the advantage of global coverage, and is not limited to the national radar and local rain gauge network coverage.

2. The post-real-time corrections for remote-sensing-based precipitation products were not necessarily valid for the unprecedented precipitation event and caused overcorrections to MRMS and IMERG, as overcorrection occurred for both product families.

3. IMERG products tended to overestimate the low–moderate precipitation intensity but underestimate the highest precipitation intensities. The NCEP product showed significant

underestimation, especially near the storm core region, due to possible instrumental failure during the record Harvey event, implying its high dependence on the functionality and reliability of the ground instruments during extreme events.

In terms of the hydrological evaluation, Spring Basin was selected because no precipitation product had significant errors within the basin from the above analysis. Two rivers were analyzed separately for their differences in urban and natural hydrological environments. The main conclusions are as follows:

1. Consistent with statistical evaluations, MRMS performs the best, showing comparable simulations with USGS stream gauge observation in Spring Basin during Hurricane Harvey, followed by IMERG and NCEP with acceptable performances.

2. The current remote-sensing-based, near-real-time precipitation products are sufficient to capture the extreme precipitation and its cascading hydrological responses. Providing the advantages of timeliness and vast spatial coverage in national and global scale, the user community is encouraged to integrate the latest remote sensing products into their operational flood disaster warning systems for the public to be informed, and to reduce and mitigate the risk of extreme precipitation events.

3. The EF5 modeling framework can capture the hydrological responses during such unprecedented extreme precipitation events; and, more powerfully, such a system can be integrated with the latest remote sensing forcing data (i.e., MRMS and IMERG) into national and even global modeling frameworks to alternatively compliment the vast ungauged regions.

This study proved the value of MRMS precipitation products for extreme precipitation detectability and accuracy, as well as their capability in hydrological prediction when combined

with the EF5 model framework, which further confirmed the success of the FLASH project run by the NOAA National Severe Storm Laboratory (NSSL) and the University of Oklahoma (http://flash.ou.edu). Even though the GPM IMERG's performance during Hurricane Harvey is second to MRMS, its simulated hydrological response is sufficient to provide flood magnitude and peak-timing warnings and to potentially build an operational flood early warning system at the global scale given satellite products' global coverage. Furthermore, it is reasonable to expect the evolving IMERG products will keep improving for hydrological and water resource applications. Concurrently, as the recent remote sensing technologies have progressed to accurately capture an unprecedented rainfall event, the EF5 modeling framework will need further improvement to provide not only the streamflow estimation but also the flood inundation extents and even water-depth over inundated urbans. Ultimately, future interdisciplinary building blocks are encouraged to connect extreme rainfall, hydrological responses and consequent flood risk analysis, as well as loss quantifications, in order to maximize the socio-economic value of the latest remote sensing observations for the general public, nationally and globally.

Reference

Abbott, M.B., Bathurst, J.C., Cunge, J.A., O'Connell, P.E., Rasmussen, J.V., 1986. An introduction to the European Hydrological System — Systeme Hydrologique European, "SHE", 1: History and philosophy of a physically-based, distributed modeling system. Journal of Hydrology 87, 45–59.

Adhikari, P., Hong, Y., Douglas, K.R., Kirschbaum, D.B., Gourley, J., Adler, R., Robert Brakenridge, G., 2010. A digitized global flood inventory (1998–2008): compilation and preliminary results. Nat Hazards 55, 405–422. https://doi.org/10.1007/s11069-010-9537-2

Ashley, S.T., Ashley, W.S., 2008. Flood fatalities in the United States. Journal of Applied Meteorology and Climatology 47, 805–817. https://doi.org/10.1175/2007JAMC1611.1.

Barredo, J.I., 2007. Major flood disasters in Europe: 1950–2005. Nat Hazards 42, 125–148. https://doi.org/10.1007/s11069-006-9065-2

Benito, G., Lang, M., Mariano Barriendos, M.C.L., Frances, F., Ouarda, T., Thorndycraft, V.R., 2004. Use of systematic, palaeoflood and historical data for the improvement of flood risk estimation. Review of scientific methods." Natural Hazards 31, 623–643.

Blake, E.S., Zelinsky, D.A., 2017. Tropical cyclone report Hurricane Harvey. National Hurricane Center, Miami, FL.

Chang, N.-B., Hong, Y. (Eds.), 2012. Multiscale hydrologic remote sensing: perspectives and applications. Taylor & Francis, Boca Raton.

Chen, H., Yang, D., Hong, Y., Gourley, J.J., Zhang, Y., 2013. Hydrological data assimilation with the Ensemble Square-Root-Filter: use of streamflow observations to update model states for real-time flash flood forecasting. Advances in Water Resources 59, 209–220.

Chow, V.T., Maidment, D.R., Mays, L.W., 1988. Applied Hydrology, International. ed. McGraw-Hill, Singapore.

Ciach, G.J., Krajewski, W.F., 1999. Radar-rain gauge comparisons under observational uncertainties. Journal of Applied Meteorology 38, 1519–1525. https://doi.org/10.1175/1520-0450(1999)038

Cifelli, R., Chandrasekar, V., Lim, S., Kennedy, P.C., Wang, Y., Rutledge, S.A., 2011. A new Dual-Polarization radar rainfall algorithm: Application in Colorado precipitation events. Journal of Atmospheric and Oceanic Technology 28, 352–364. https://doi.org/10.1175/2010JTECHA1488.1.

Clark, R.A., Flamig, Z.L., Vergara, H., Hong, Y., Gourley, J.J., Mandl, D.J., Frye, S., Handy, M., Patterson, M., 2017. Hydrological Modeling and Capacity Building in the Republic of Namibia. Bulletin of the American Meteorological Society 98, 1697–1715. https://doi.org/10.1175/BAMS-D-15-00130.1

Cohen, S., Praskievicz, S., Maidment, D.R., 2018. Featured Collection Introduction: National Water Model. J Am Water Resour Assoc 54, 767–769. https://doi.org/10.1111/1752-1688.12664

Dutta, D., Herath, S., Musiake, K., 2006. An application of a flood risk analysis system for impact analysis of a flood control plan in a river basin. Hydrological Processes 20, 1365–1384.

Dutta, D., Teng, J., Vaze, J., Lerat, J., Hughes, J., Marvanek, S., 2013. Storage-based approaches to build floodplain inundation modeling capability in river system models for water resources planning and accounting. Journal of Hydrology 504, 12–28.

Eric, B.S., Zelinsky, D.A., 2018. National Hurricane Center tropical cyclone report Hurricane Harvey. National Hurricane Center, Miami, Florida.

Flamig, Z.L., Vergara, H., Gourley, J.J., 2020. The Ensemble Framework For Flash Flood Forecasting (EF5) v1.2: Description and Case Study (preprint). Hydrology. https://doi.org/10.5194/gmd-2020-46

Gallegos, H.A., Schubert, J.E., Sanders, B.F., 2009. Two-dimensional, high-resolution modeling of urban dam-break flooding: A case study of Baldwin Hills, California. Advanced in Water Resources 32, 1323–1335.

Gourley, J.J., Flamig, Z.L., Vergara, H., Kirstetter, P.-E., Clark, R.A., Argyle, E., Arthur, A., Martinaitis, S., Terti, G., Erlingis, J.M., Hong, Y., Howard, K.W., 2017. The FLASH Project: Improving the Tools for Flash Flood Monitoring and Prediction across the United States. Bulletin of the American Meteorological Society 98, 361–372. https://doi.org/10.1175/BAMS-D-15-00247.1

Hardy, R.J., Bates, P.D., Anderson, M.G., 2000. Modelling suspended sediment deposition on a fluvial floodplain using a two-dimensional dynamic finite element model. Journal of Hydrology 229, 202–218.

Hayatbini, N., Hsu, K., Sorooshain, S., Zhang, Y., Zhang, F., 2019. Effective cloud detection and segmentation using a gradient-based algorithm for satellite imagery: Application to improve PERSIANN-CCS. Journal of Hydrometeorology 901–913. https://doi.org/10.1175/JHM-D-18-0197.1.

Hong, Y., Adler, R.F., Hossain, F., Curtis, S., Huffman, G.J., 2007a. A first approach to global runoff simulation using satellite rainfall estimation. Water Resources Research 43. https://doi.org/10.1029/2006WR005739.

Hong, Y., Gochis, D., Cheng, J.-T., Hsu, K., Sorooshian, S., 2007b. Evaluation of PERSIANN-CCS rainfall measurement using the NAME event rain gauge network. Journal of Hydrometeorology 8, 469–482. https://doi.org/10.1175/JHM574.1.

Hong, Y., Gourley, J.J., 2015. Radar hydrology: principles, models, and applications. Taylor & Francis, Boca Raton.

Hong, Y., Hsu, K.-L., Moradkhani, H., Sorooshian, S., 2006. Uncertainty quantification of satellite precipitation estimation and Monte Carlo assessment of the error propagation into hydrologic response. Water Resource Research 42. https://doi.org/10.1029/2005WR004398.

Hou, A.Y., Kakar, R.K., Neeck, S., Azarbarzin, A.A., Kummerow, C.D., Kojima, M., Oki, R., Nakamura, K., Iguchi, T., 2014. The Global Precipitation Measurement Mission. Bulletin of American Meteorological Society 701–722.

Huang, Y., Chen, S., Cao, Q., Hong, Y., Wu, B., Huang, M., Qiao, L., 2014. Evaluation of Version-7 TRMM multi-satellite precipitation Analysis product during the Beijing extreme heavy rainfall event of. Water 6, 32–44. https://doi.org/10.3390/w6010032.

Huffman, G., Bolvin, D., Braithwaite, D., Hsu, K., Joyce, R., Xie, P., 2014. Integrated Multi-satellitE Retrievals for GPM (IMERG), in: Vers. 4.4. NASA's Precipitation Processing Center.

Huffman, G.J., Adler, R.F., Morrissey, M.M., Bolvin, D.T., Curtis, S., Joyce, R., McGavock, B., Suzzkind, J., 2001. Global precipitation at one-degree daily resolution from multisatellite observations. Journal of Hydrometeorology 2, 36–50.

Huffman, G.J., Bolvin, D.T., Braithwaite, D., Hsu, K., Joyce, R., Kidd, C., Nelkin, E.J., Sorooshain, S., Tan, J., Xie, P., 2012. Developing the Integrated Multi-Satellite Retrievals for GPM (IMERG), in: EGU General Assembly. EGU, Vienna, p. 6921.

Huffman, G.J., Bolvin, D.T., Braithwaite, D., Hsu, K., Joyce, R., Kidd, C., Nelkin, E.J., Sorooshian, S., Tan, J., Xie, P., 2019a. Algorithm Theoretical Basis Document (ATBD) Version 06 NASA Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG).

Huffman, G.J., Bolvin, D.T., Nelkin, E.J., Jackson, T., Braithwaite, D., Hsu, K., Joyce, R., Kidd, C., Sorooshian, S., Xie, P., 2019b. Early results for Version 06 IMERG. Presentation, Singapore, Singapore.

Huffman, G.J., Bolvin, D.T., Nelkin, E.J., Stocker, E.F., Tan, J., 2019c. V06 IMERG Release Notes.

Kao, S.-C., ASCE, M., DeNeale, S.T., Watson, D.B., 2019. Hurricane Harvey highlights: Need to assess the adequacy of Probable Maximum Precipitation estimation methods. Journal of Hydrologic Engineering 24. https://doi.org/doi:05019005.

Kats, S., 2011. NCEP/EMC U.S. Gage Only Hourly Precipitation Data Version 1.0. UCAR/NCAR - Earth Observing Laboratory.

Kocalets, I.V., Kivva, S.L., Udovenko, O.I., 2015. Usage of the WRF/DHSVM model chain for simulation of extreme floods in mountainous areas: a pilot study for the Uzh River Basin in the Ukrainian Carpathians. Natural Hazards 75, 2049–2063. https://doi.org/10.1007/s11069-014-1412-0.

Lehner, B., Verdin, K., Jarvis, A., 2008. New Global Hydrography Derived From Spaceborne Elevation Data. Eos Trans. AGU 89, 93. https://doi.org/10.1029/2008EO100001

Li, Z., Yang, D., Yang, H., 2013. Multi-scale evaluation of high-resolution multi-sensor blended global precipitation products over the Yangtze River. Journal of Hydrology 157–169.

McAllister, M., Gochis, D., Bariage, M.J., Dugger, A.L., FitzGerald, K., Karsten, L., McCreight, J.L., 2018. The community of WRF-Hydro Modeling system Version 5 melding with the National Water Model: Enhancements and education, in: AGU Fall Meeting 2018. AGU, Washington D.C. USA.

Merz, B., Kreibich, H., Lall, U., 2013. Multi-variate flood damage assessment: a tree-based data-mining approach. Natural Hazards and Earth System Sciences 13, 53–64.

Murphy, J.D., 2018. Service assessment August–September 2017 Hurricane Harvey (US DOC). NOAA National Weather Service, Silver Spring, Maryland.

Nigussie, T.A., Altunkaynak, A., 2019. Modeling the effect of urbanization on flood risk in Ayamama Watershed, Istanbul, Turkey, using the MIKE 21 FM model. Nat Hazards 99, 1031–1047. https://doi.org/10.1007/s11069-019-03794-y

Omaranian, E., Sharif, H.O., Tavakoly, A.A., 2019. How well can Global Precipitation Measurement (GPM) capture hurricanes? Case study: Hurricane Harvey. Remote sensing 10. https://doi.org/10.3390/rs10071150.

Pasquier, U., He, Y., Hooton, S., Goulden, M., Hiscock, K.M., 2019. An integrated 1D–2D hydraulic modeling approach to assess the sensitivity of a coastal region to compound flooding hazard under climate change. Natural Hazards 98, 915–937. https://doi.org/10.1007/s11069-018-3462-1.

Pham, N.T.T., Nguyen, Q.H., Ngo, A.D., Le, H.T.T., Nguyen, C.T., 2018. Investigating the impacts of typhoon-induced floods on agriculture in the central region of Vietnam by using hydrological models and satellite data. Natural Hazards 92, 189–204. https://doi.org/10.1007/s11069-018-3202-6.

Sampson, C.C., Smith, A.M., Bates, P.D., Neal, J.C., Alfieri, L., Freer, J.E., 2015. A high-resolution global flood hazard model: A HIGH-RESOLUTION GLOBAL FLOOD

HAZARD MODEL. Water Resour. Res. 51, 7358–7381. https://doi.org/10.1002/2015WR016954

Senatore, A., Mendicino, G., Gochis, D.J., Yu, W., Yates, D.N., Kunstmann, H., 2015. Fully coupled atmosphere-hydrology simulations for the central Mediterranean: Impact of enhanced hydrological parameterization for short and long time scales. Journal of Advances in Modeling Earth Systems 7, 1693–1715. https://doi.org/10.1002/2015MS000510.

Seo, D., 1998. Real-time estimation of rainfall fields using rain gage data under fractional coverage conditions. Journal of Hydrology 25–36.

Smith, K., Ward, R., 1998. Floods: Physical processes and human impact. John Wiley, Chichester.

Tang, G., Zeng, Z., Long, D., Guo, X., Yong, B., Zhang, W., Hong, Y., 2016. Statistical and hydrological comparisons between TRMM and GPM Level-3 products over a midlatitude basin: Is Day-1 IMERG a good successor for TMPA 3B42V7. Journal of Hydrometeorology 121–137.

Teng, J., Jakeman, A.J., Vaze, J., Croke, B.F.W., Dutta, D., Kim, S.S.H., 2017. Flood inundation modeling: A review of methods, recent advances and uncertainty analysis. Environmental Modelling & Software 90, 201–216.

Thakur, M.K., Kumar, T.V.L., Dwivedi, S., Narayanan, M.S., 2018. On the rainfall asymmetry and distribution in tropical cyclones over Bay of Bengal using TMPA and GPM rainfall products. Natural Hazard 94, 819–832. https://doi.org/10.1007/s11069-018-3426-5.

van Oldenborgh, G.J., van der Wiel, K., Sebastian, A., Singh, R., Arrighi, J., Otto, F., Haustein, K., Li, S., Vecchi, G., Cullen, H., 2018. Corrigendum: Attribution of extreme rainfall from Hurricane Harvey, August 2017 (2017 *Environ. Res. Lett.* **12** 124009). Environ. Res. Lett. 13, 019501. https://doi.org/10.1088/1748-9326/aaa343

Verdin, J., Funk, C., Senay, G., Choularton, R., 2005. Climate science and famine early warning. Philosophical Transactions of the Royal Society B: Biological Science 360, 2155–2168. https://doi.org/10.1098/rstb.2005.1754.

Vergara, H., Kirstetter, P.-E., Gourley, J.J., Flamig, Z.L., Hong, Y., Arthur, A., Kolar, R., 2016. Estimating a-priori kinematic wave model parameters based on regionalization for flash flood forecasting in the Conterminous United States. Journal of Hydrology 541, 421–433. https://doi.org/10.1016/j.jhydrol.2016.06.011

Villarini, G., Mandapaka, P.V., Krajewski, W.F., Moore, R.J., 2008. Rainfall and sampling uncertainties: A rain gauge perspective. Journal of Geophysical Research 113. https://doi.org/10.1029/2007JD009214.

Vrugt, J.A., Braak, C.J.F., Gupta, H.V., Robinson, B.A., 2009. Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling? Stochastic Environmental Research and Risk Assessment 23, 1101–1026. https://doi.org/10.1007/s00477-008-0274-y.

Wang, J., Hong, Y., Li, L., Gourley, J.J., Khan, S.I., Yilmaz, K.K., Adler, R.F., Policelli, F.S., Habib, S., Irwn, D., Limaye, A.S., Korme, T., Okello, L., 2011. The coupled routing and excess storage (CREST) distributed hydrological model. Hydrological Sciences Journal 56, 84–98. https://doi.org/10.1080/02626667.2010.543087

Werner, M., Reggiani, P., Roo, A.D., Bates, P., Sprokkereef, E., 2005. Flood forecasting and warning at the River Basin and at the European scale. Natural Hazards 36, 25–42.

Wu, K., Johnston, C.A., 2007. Hydrologic response to climatic variability in a Great Lakes Watershed: A case study with the SWAT model. Journal of Hydrology 337, 187–199. https://doi.org/10.1016/j.jhydrol.2007.01.030.

Xue, X., Zhang, K., Hong, Y., Gourley, J.J., Kellogg, W., McPherson, R.A., Wan, Z., Austin, B.N., 2016. New multisite cascading calibration approach for hydrological models: case study in the Red River Basin using the VIC model. Journal of Hydrologic Engineering 21. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001282.

Yang, L., Smith, J., Liu, M., Baeck, M.L., 2019. Extreme rainfall from Hurricane Harvey (2017): Empirical intercomparisons of WRF simulations and polarimetric radar fields. Atmospheric Research 223, 114–131.

Zhang, J., Howard, K., Langston, C., Kaney, B., Qi, Y., Tang, L., Grams, H., Wang, Y., Cocks, S., Martinaitis, S., Arthur, A., Cooper, K., Brogden, J., Kitzmiller, D., 2016. Multi-Radar Multi-Sensor (MRMS) Quantitative Precipitation Estimation: Initial Operating Capabilities. Bulletin of the American Meteorological Society 97, 621–638. https://doi.org/10.1175/BAMS-D-14-00174.1

Zhang, W., Villarini, G., Vecchi, G.A., Smith, J.A., 2018. Urbanization exacerbated the rainfall and flooding caused by hurricane Harvey in Houston. Nature 563, 384–398.

Zhang, Y., Hong, Y., Wang, X., Gourley, J.J., Xue, X., Saharia, M., Ni, G., Wang, G., Huang, Y., Chen, S., Tang, G., 2015. Hydrometeorological Analysis and Remote Sensing of Extremes: Was the July 2012 Beijing Flood Event Detectable and Predictable by Global Satellite Observing and Global Weather Modeling Systems? Journal of Hydrometeorology 16, 381–395. https://doi.org/10.1175/JHM-D-14-0048.1

Zonta, R., Collavini, F., Zaggia, L., Zuliani, A., 2005. The effect of floods on the transport of suspended sediments and contaminants: A case study from the estuary of the Dese River (Venice Lagoon, Italy). Environment International 31, 948–958.

Chapter 3. A comprehensive flood inundation mapping for Hurricane Harvey using an integrated hydrological and hydraulic model

Abstract

As climate change will increase the frequency and intensity of precipitation extremes, especially for tropical storms causing severe coastal flooding, there is a clear need for an integrated hydrology and hydraulic system that has the benefit of utilizing the modern computation technology to model the hydrologic conditions over a long period and the flow dynamic representations when extreme hydrometeorological events occur. This system coupling provides comprehensive information (flood wave, inundation extents and depths) about coastal flood events for emergency management and risk minimization. This study provides an integrated hydrologic and hydraulic coupled modeling system that is based on the Coupled Routing and Excessive Storage (CREST) model and the Australia National University- Geophysics Australia (ANUGA) model to simulate flood. Forced by the near-real-time Multi-Radar Multi-Sensor (MRMS) Quantitative Precipitation Estimates (QPEs), this integrated modeling system was applied during the 2017 Hurricane Harvey event to simulate the streamflow, the flood extent, and the inundation depth. The results were compared with post-event Water High Mark (WHM) survey data and its interpolated flood extent by the United States Geological Survey (USGS) and the Federal Emergency Management Agency (FEMA) flood insurance claims, as well as a satellite-based flood map, the National Water Model (NWM) and the Fathom (LISFLOOD-FP) model simulated flood map. The proposing hydrologic and hydraulic model simulation indicated that it could capture 87% of all flood insurance claims within the study area, and the overall error of water depth was 0.91 meters, which is comparable to the mainstream operational flood models (NWM and Fathom).

3.1 Introduction

Flooding triggered by excessive precipitation is the second-deadliest and most common natural hazard in the United States and worldwide (Ashley and Ashley, 2008; Barredo, 2007; Benito et al., 2004; Smith and Ward, 1998), which accounts for 43% of total United Nation (UN) recorded natural disasters from 1995 to 2015 (Wallemacq et al., 2015). The Gulf Coast and the South Atlantic Coast of the USA are profoundly affected by extreme precipitation from tropical cyclones and their resulting floods (Adhikari et al., 2010), which is responsible for around 25,000 fatalities in the USA since 1942 (Rappaport, 2014, 2000). For pluvial floods, the precipitation rate, duration, the land use of the region, topography, and antecedent soil moisture conditions are the main factors to determine the overall flood severity (Brauer et al., 2020). Recent studies indicate that the frequency and intensity of extreme rainfall and tropical cyclones will increase (van Oldenborgh et al., 2018) and the propagation of the cyclones will decrease due to climate change (Kossin, 2018). It is thus likely that the future flood risk and its consequential socio-economic damage will escalate. On top of the changes in the tropical cyclone characteristics, rising sea level in a warming climate can intensify coastal flooding (Wing et al., 2019). There is a clear need for tools that can facilitate the current and future flood risk mitigation.

One of those tools is an improved, real-time flood prediction system, which supports multi-disciplinary decision making including but not limited to: first-responder preparedness, temporary flood defense planning, insurance budgeting, and supply chain management. Flood prediction and inundation mapping have been studied for over a century, and two traditional groups of research efforts have gained most of the attention of the research community: observations and hydraulic models (Teng et al., 2017). The observation methods include ground

measurements, surveys, and remote-sensing technology (Shen et al., 2019b, 2019a; Syvitski and Brakenridge, 2013). As remote-sensing technologies can identify the flood extent over a vast area, while other observation methods can only provide data at single points, flood mapping aided by remote sensing has gained popularity in recent years. However, due to limitations such as data latency, the observation-based results are more often used as an input or as a benchmark to validate and calibrate hydrological and hydraulic models (Teng et al., 2017).

Hydraulic models include one-dimensional, two-dimensional, and three-dimensional methods that use the physical equations and laws to describe fluid motion where the degree of complexity varies. The one-dimensional hydraulic model is considered the most straightforward representation of floodplain flow, as it simulates the open surface water flow with the assumption that the water flows in the same direction. The flow velocity is then averaged over the channel cross-section (Brunner, 2016). The 2D model assume the water as a shallow ditch where no flow occurs vertically, and the shallow water equation is solved from the depth-averaging Navier-Stokes equations (Roberts et al., 2015). In some individual cases which need detailed information for engineering solutions, such as dam breaks, tsunamis, or embankment failures, 3D hydraulic models are implemented. However, for most of the floodplain analysis and simulation, the 2D shallow-water approximation is considered adequate after proper model construction and validation (Alcrudo, 2004). The fully solved the 2D St. Venant shallow water equations using the finite volume and the finite element methods are considered to have a higher complexity among the 2D hydraulic models (Bates and De Roo, 2000). The storage cell or the cellular automata approach by solving Manning's equation with finite difference methods, such as LISFLOOD-FP model, was suggested to be a good approximation to the physics-based model and the computation time was reduced by 30 times (Ghimire et al., 2013). It was further tested to

prove the method was as efficient as other classes of models implementing HPC techniques (Bates et al., 2010).

Integrating hydrologic model and hydraulic models has the benefit of utilizing present-day computational resources to model dynamic representations of extreme hydrometeorological events (Anselmo et al., 1996). A recent study (Tanaka et al., 2018) has integrated three difference models: a distributed hydrological model called Geomorphology-Based Hydrological Model (GBHM), a 1D hydraulic model named Mike11, and a 2D hydraulic model called Local Inertial Equation (LIE). The study found the integrated framework yielded good agreement between with the observation data of the stream discharge, as well as the lake water level over four years span. It was able to simulate a significant flooding event in 2000 over the study area. The authors also indicated that the framework could simulate sediment movement downstream in the future research plan.

Since 2016, the Ensemble Framework For Flash Flood Forecasting (EF5) integrated the Coupled Routing and Excess Storage (CREST) distributed hydrological model, a 1D hydraulic model, with kinematic wave channel routing, to successfully simulate multiple extreme precipitation-triggered flash flooding events in Oklahoma City and Houston at a continental scale implementation (Flamig et al., 2020; Gourley et al., 2017). While outside the USA, the CREST was coupled with an 1D fully distributed linear reservoir routing scheme and found success over a study in China (Shen et al., 2017). The National Water Center (NWC) led an effort to integrate the WRF-Hydro hydrologic model and the Height Above Nearest Datum (HAND) inundation mapping method into the new National Water Model (NWM; Cohen et al., 2018). The HAND method performed a simulation of a 2016 Texas flooding event with good agreement with remote-sensing observations and less computation cost (Zhang et al., 2018). A recent study by

(Wing et al., 2019), compared the flood mapping performances for Hurricane Harvey between NWM+HAND and Fathom, a LISFLOOD-FP based hydraulic modeling system that contains all major hydrological and hydraulic components to describe the water dynamics. The study results indicated that Fathom outperformed NWM+HAND according to all statistical metrics and could better capture the pluvial and coastal flooding phenomena. The public would greatly benefit from a comprehensive and accurate flood extent and inundation depth predictions for tropical cyclone events to improve local risk management.

A "grand challenge for hydrology" was raised by (Wood et al., 2011) to provide hyperresolution hydrological prediction capacities to the public, as the society critically demands the high spatial-temporal resolution forecasting for floods and droughts. New remote-sensing technology provides accurate, and high spatially and temporally resolved observations over the globe, which helps to advance physics-based models for atmospheric, hydrologic and hydraulic forecasting. However, the remote-sensing products are not error-free and these errors in precipitation products can be further propagated to the hydrologic & hydraulic modeling results (Hong et al., 2006); therefore the accuracy of the input precipitation data is crucial for flood applications. The Multi-Radar Multi-Sensor system (MRMS), which utilizes data from over 180 NEXRAD radars and covers the conterminous United States at 1 km spatial resolution with a 2-min update frequency (Zhang et al., 2016), has been shown to have the best performance during Hurricane Harvey event relatively to NASA's Global Precipitation Mission (GPM) Integrated Multi-satellitE Retrievals (IMERG) v6 product, and National Centers for Environmental Prediction (NCEP) gridded gauge only precipitation production; and has a good agreement with the Harris County Flood Control District (HCFCD) rain gauge data (Chen et al., 2020; Li et al., 2020). This study, again, uses Hurricane Harvey as the study case, since it was considered a 100

to 500-year flood event, which caused the local streams' return period reduced 20% to 35% after the event (McDonald and Naughton, 2019; Vu and Mishra, 2019).

The overarching goal of the study is to first introduce the newly developed hydrology and hydraulic model CREST- inundation Mapping And Prediction (CREST-iMAP) as a latest addition of the well-documented CREST hydrologic modeling family; and to further test and compare this new coupled system with some other the-state-of-the-art flood mapping models in an extreme storm event setting. Additionally, the physics-based 2D hydrological-hydraulic CREST-iMAP model provides ease on scale-up implementation with less data preparation requirement and streamflow-only calibration processes. More specified objectives of this study are to (a) design and develop the CREST-iMAP which couples the hydrologic and hydraulic components, while bypasses the river channel data requirement; (b) evaluate the flood extent by comparing the modeling results with the observed USGS flood map, satellite-based (SAR) flood map, Fathom flood map, NWM+HAND flood map, and the FEMA flood insurance claims map for Hurricane Harvey, and (c) evaluate the flood inundation depths between the USGS Water-High-Marks (WHMs), Fathom flood map, NWM+HAND flood map, and CREST-iMAP simulated flood map using statistic methods. This paper is organized as follows. Section 2 describes the study area, data, the description of the CREST hydrologic model and 1D hydraulic model, Fathom, and the evaluation methodology. Section 3 discusses the results of the inter-comparison of the flood maps from the USGS interpolation, SAR data, the Fathom simulation, the NWM+HAND simulation, and the CREST-iMAP based simulation. Section 4 concludes the study and proposes future directions.

3.2 Methods and data

3.2.1 Study area and data

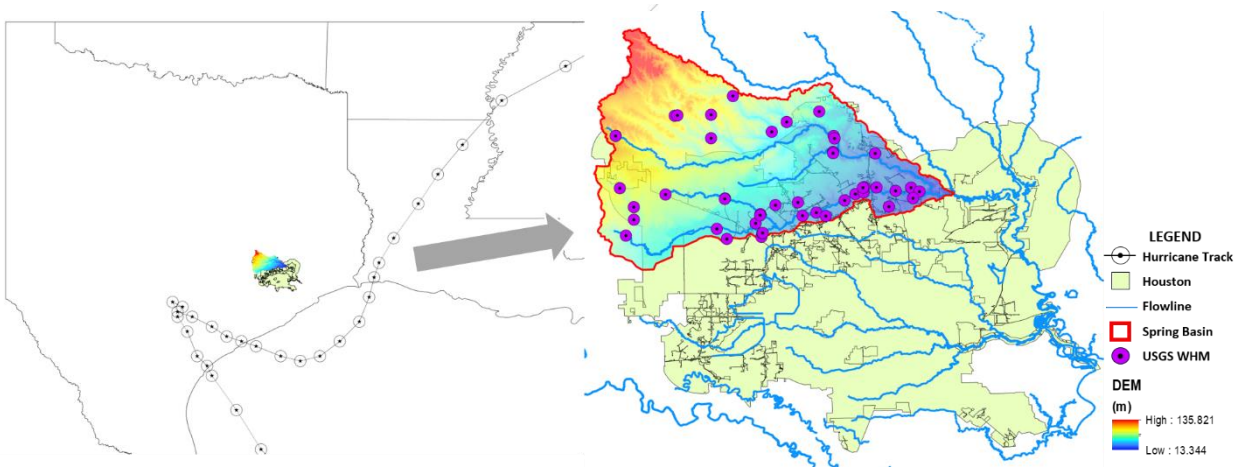


Figure 3.1. Study area showing Hurricane Harvey storm-track, City of Houston Spring Basin, WHM locations as well as the topography of Spring Basin

Hurricane Harvey made the first landfall on northern San Jose island, TX, on August 26th, 2017, and moved along the Texas Coast for almost 5 days before the second landfall on August 30th, 2017 (Eric and Zelinsky, 2018). Hurricane Harvey poured over 1,500 mm of water on the Great Houston area (Brauer et al., 2020). Figure 3.1 displays the Hurricane track, the boundary of Houston, the Spring basin, major flowlines, and the USGS Water-High-Marks (WHMs) within the Spring basin. The Spring basin, located in the northwestern part of the Great Houston area, was selected for this study. The Spring basin is among the most impacted area during Hurricane Harvey (Murphy, 2018). It has mixed landcover types across the basin (Chen et al., 2020), which was believed to be an underperforming area by another study (Wing et al., 2019). There are four major rivers within the Spring basin: Spring Creek, Willow Creek, Little Cypress Creek, and Cypress Creek, entering Lake Houston and 55 USGS WHM sites measured for the Hurricane Harvey event. The WHM sites are more concentrated along Cypress Creek and the southeastern part of the basin because the area has more urban development.

The MRMS radar-based QPE was obtained from the Iowa Environmental Mesonet NWS data archive (https://mesonet.agron.iastate.edu/nws/). The 15-min streamflow data of five stream gauges across the basin from April 1st to September 3rd, 2017, were obtained from the USGS National Water Information System (https://waterdata.usgs.gov/nwis). The hyper-resolution (10 meters) Digital-Elevation-Model (DEM) was obtained from USGS Earth Explorer (https://earthexplorer.usgs.gov/). The 250-meter resolution hydrologically-conditioned DEM was obtain from HydroSHEDS (https://www.hydrosheds.org/; Lehner et al., 2008). And the 10-meter resolution CREST parameters were derived and calculated using methods described in (Vergara et al., 2016) from the USGS Soil Survey Geographic Database (SSURGO). The 1-km resolution CREST parameters were passed from the beginning phase of FLASH project to input to the hydrological modeling portion of the CREST-iMAP. The US landcover data was obtained from Multi-Resolution Land Characteristics Consortium (MRLC, https://www.mrlc.gov/), and it was extended to arrive at gridded Manning's roughness coefficients as described in (Liu et al., 2019). The USGS Water High Mark (WHM) data and FEMA property claim data during Hurricane Harvey were obtained from HydroShare (https://hydroshare.org; Arctur et al., 2018).

3.2.2 CREST inundation mapping and prediction (CREST-iMAP) framework

CREST is a grid-based, distributed hydrological model that was developed by the University of Oklahoma and NASA Applied Science Team (Wang et al., 2011). The EF5 framework later included CREST as one of the water balance modeling cores and coupled it with the kinematic wave channel routing method, and the framework can provide hydrological simulation at continental and global scales (Clark et al., 2017). Researchers have adapted the EF5 framework for high-resolution flash flood forecasting in the USA (Flamig et al., 2020). It now serves as the backbone of the Flooded Locations and Simulated Hydrographs project (FLASH), which was

transitioned to the NWS in Nov. 2016 and has evolved the tools for operational flash flood forecasting (Gourley et al., 2017). The NWS Weather Forecast Offices (WFOs) have reported that having FLASH data was extremely useful by allowing staff to focus in on threats and upgrade warnings more rapidly and timely (Murphy, 2018). This study uses a modified EF5 framework from the implemented version in the FLASH project, which is used as water balance component of the CREST-iMAP. CREST is coupled with the Australia National University and Geoscience of Australia (ANUGA) hydraulic model (Nielson et al., 2005). The ANUGA model was built based on a finite volume method for solving the 2D Shallow Water Equation (Roberts et al., 2015), to simulate the floodplain flow movement and assume that water depth is much less than the water movements in x and y directions (Teng et al., 2017). However, the ANUGA model only simulates the 2D water depth distribution and flow velocity, while the water-soil and water-atmosphere interactions are not considered in the modeling framework. Therefore, an additional component of water balancing is needed to meet that need.

In this study, the EF5-CREST model was one-way and offline coupled with the ANUGA model to comprehensively provide flood information, including streamflow, flood extent, and inundation depth. Figure 3.2 illustrates the schematic flow chart of CREST-iMAP coupling mechanics. CREST-iMAP receives the forcing precipitation data, such as radar/satellite QPE, machines learning modeled or numerical modeled Quantitative Precipitation Forecasts (QPFs), or interpolated rainfall field. The hydrological model simulates and generates excessive rainfall, soil moisture and streamflow, where excessive rainfall as well as soil moisture are further used to drive and set the initial soil condition for the hydraulic model (ANUGA), and the streamflow is the 1-dimensional output from the CREST-iMAP. The water balance module has 17 parameters to describe the physical interactions between water, soil, and air (Wang et al., 2011), while the

hydraulic model has only the Manning's roughness coefficient that governs the water flow (Roberts et al., 2015). The simulated streamflow is the variable used to calibrate the model parameters based on comparisons with USGS stream gauge records. The hydraulic model receives the excessive rainfall data field from each time step and simulates two variables in a single raster file to illustrate the flood extent and the flood inundation depth to the end user.

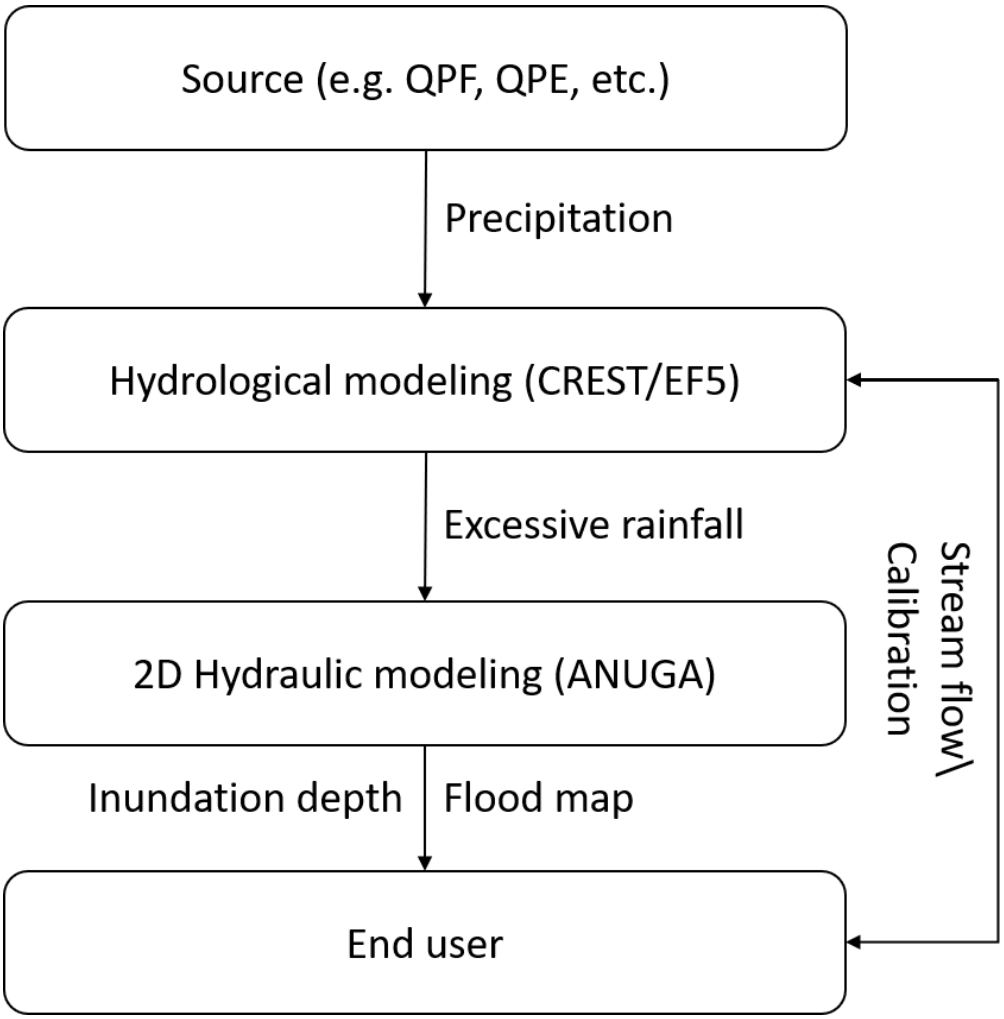


Figure 3.1. The schematic data flowchart of CREST-iMAP. One component (specified in rectangular boxes) becomes the input of another, where arrows represent the data flow

In this study, the MRMS near-real-time QPE was input to CREST-iMAP at 2-min temporal resolution. To imitate the nested real-time operational scheme, the water balance module calculates the excessive rainfall at 250-meter resolution and simulated streamflow on all water

channels using 250-meter DEM and 1-km parameters, which were calibrated with five USGS stream gauges described in (Chen et al., 2020), using DREAM methods (Vrugt, 2016; Vrugt et al., 2009). The calibrated excessive rainfall was then read by a hydraulic module at 2-min temporal resolution, and the internal calculation time step is 10 seconds on an approximated 10 by 10 m^2 area triangular mesh. Throughout the hydrologic-hydraulic modeling processes, the 1 km resolution MRMS QPE is first derived into 250 m excessive rainfall, and then calculated for water depth and water fluxes at 10-m resolution. The model output was extracted every 15 min and further interpolated to the 10-m resolution raster files. For more accurate simulation of the flooding impact of Hurricane Harvey, the water balance module utilized a warmup period from 01/04/2017 to 25/08/2017 and coupled the simulation period from 25/08/2017 to 03/09/2017. This approach is an attempt to use the native DEM with no manipulation and bypass the detailed river channel information, which is not available at all locations and problematic for an operational flood prediction system (Lejot et al., 2007; Merwade et al., 2006; Orzech et al., 2011). All flow direction and flow accumulation data were derived from the original DEM and relied on its quality. Therefore, CREST-iMAP can reduce the data requirement for model applications comparing to traditional 2D hydraulic modeling (Brunner, 2016). CREST-iMAP is designed to operate at the real-time and provide timely flood information, therefore, the comparing group of flood maps are from the real-time operational system over the world.

3.2.3Fathom model (LISFLOOD-FP)

Fathom-Global is a framework that combines multiple continental-scale hydraulic model implementations (Wing et al., 2017), among which was first described by (Sampson et al., 2015). Fathom utilizes the LISFLOOD-FP as the computational core, which solves the local inertial form of the shallow water equations using a 2D regular grid (Bates et al., 2010; Bates and De

Roo, 2000; de Almeida and Bates, 2013). In this study, the Fathom simulated flood map was generated using the river discharge data as the forcing, which was the USGS stream gauge observations at the related inflow points in the stream network (Wing et al., 2019). The pluvial flood was also considered by inputting the US Weather Prediction Center (WPC) Stage IV rainfall data onto the grid cells. The infiltration was considered using a simplified infiltration capacity method, using the information from the Harmonized World Soil Database derived by a modified Hortonian infiltration equation (Morin and Benyamini, 1977). The simulation also included levee information, burned in channels, and storm surge simulation to capture the multiple phenomena contributing to the Hurricane Harvey flood at 10-meter spatial resolution. This data is from an operational Fathom-US version, which was built upon and updated from a high resolution global flood hazard model (Sampson et al., 2015; Wing et al., 2019), which continuously updates the flood frequency analysis at all catchments in the world. The model was operating at the operational level to produce the flood map.

3.2.4 NWM coupled with HAND

As a part of the National Flood Interoperability Experiment (Maidment, 2017), the NOAA National Water Center has been exploring to couple the NWM and HAND to provide the flood map and prediction. The HAND method created a normalized HAND DEM from the original DEM, which indicates the height of each grid above its nearest flow channel. Then the NWM streamflow data at each drainage reach is converted to the water stage long the channel grids by reversing the flow-stage rating curve, and then any HAND DEM grid values that is small than its nearest water channel stage is considered “wet” and the inundation depth equals to the difference between the stage and the HAND DEM cell value. This simple and conceptual approach was executed at one third arc second spatial resolution for every day during and after Hurricane

Harvey (28/08/2017 – 03/09/2017). A maximum extent flood simulation was generated based on the daily inundation data and was downloaded from HydroShare platform (Arctur et al., 2018). NWM v1.1 was deployed for operational uses to NCEP in late 2016 until October 2017, the v1.2 was delivered. This data is based on the NWM v1.1, which includes over 49 thousand catchments and the Harris County area had high positive bias but high correlation coefficient when comparing to USGS during an official assessment during the NWM training seminar in 2017 (Gochis et al., 2017).

3.2.5 Reference flood data

Following Hurricane Harvey, the USGS field team visited multiple impact areas and collected over 2,000 HWMs with the official guidelines (Feaster and Koenig, 2017; Koenig et al., 2016). The HWM ground survey was conducted by measuring the GPS elevation of the mud, debris, and water stain lines on the side of buildings, trees, fences, poles, and other structures. (Watson et al., 2018) utilized over 2,000 HWMs and 47 peak stage heights and interpolated them into a flooded water plain at 19 locations, aided by a LiDAR-derived DEM with 1.4 to 3.0-m resolution. This interpolation result was believed to be the best reconstruction of the flood extent and has been used as a benchmark in another study (Wing et al., 2019). However, this data set is not error-free, and (Watson et al., 2018) indicated the uncertainty ranged from 0.01 to 0.55 meters at specific points. This data was obtained from the USGS data release (https://doi.org/10.5066/F7VH5N3N).

Remote sensing technology provide another potential reference for flood extend as there are clear Synthetic Aperture Radar (SAR) images on August 29th and 30th of 2017, that capture the Hurricane Harvey and its flooding impact on Harris county, TX (Shen et al., 2019a). As satellite data can objectively illuminate the earth surface, theoretically, it has the potential to be the

unbiased ground truth for the flood extent only if the information can be extracted accurately. In this study the flood map produced from Radar-Produced Inundation Diary (RAPID) system, which was derived from Sentinel-1 SAR data captured on August 29th, 2017 and then went through binary classification, morphological processing, compensation, and machine learning correction. Yet, the automated algorithms to detect water surface through vegetations and urban structure are not available when RAPID was put online, so RAPID is not capable to extract flop map from SAR data at its current version. The RAPID system was calibrated using a 2016 flooding event in China, before processing the data for Hurricane Harvey. The data was obtained from the University of Connecticut Hydrometeorology and Hydrologic Remote Sensing Group, RAPID flood map archive (https://rapid-nrt-flood-maps.s3.amazonaws.com/index.html#RAPID_Archive_Flood_Maps/20170829/flooding_S1A_IW_GRDH_1SDV_20170829T002645_20170829T002710_018131_01E74D_3220/). RAPID system is an operational system designed to quickly extract flood maps from SAR images during or after the event, therefore many amateur water identification technologies are not incorporated in the system. The RAPID flood map was interpolated using the Floodwater Depth Estimation Tool version 2 (FwDET v2; Cohen et al., 2019), to generate the flood depth by subtracting the elevation surface created by the edge of the flooded polygon by the 10-meter resolution DEM.

3.2.6 Statistical metrics

Two levels of statistical tests were used in the evaluation. First, the reference data from 3.2.5 were used to test the extent to which the models capture the spatial patterns of flooding. In this case, we used standard binary pattern measures listed in Table 3.1.

The Probability of Detection (POD) measures the model's ability to capture the referencing flood extent or the proportion of the reference flood extent that was replicated by the model. The

False Alarm Rate (FAR) reflects the model's tendency to overestimate the reference flood extent or the proportion of the modeled flood area that was classified as positive while the reference data was classified as negative. The Critical Success Index (CSI) measures the performance of the model estimates compared to the reference flood extent, which accounts for both overprediction and underprediction by the model.

Table 3.1. List of statistical metrics used in this study

Name	Equation[a]	Value range	Perfect value
Probability of detection	$\mathrm{POD} = \frac{F_1 \wedge R_1}{F_1 \wedge R_1 + F_0 \wedge R_1}$	0,1	1
False alarm ratio	$\mathrm{FAR} = \frac{F_1 \wedge R_0}{F_1 \wedge R_0 + F_1 \wedge R_0}$	0,1	0
Critical success index	$\mathrm{CSI} = \frac{F_1 \wedge R_1}{F_1 \wedge R_1 + F_0 \wedge R_1 + F_1 \wedge R_0}$	0,1	1
Correlation coefficient	$CC = \frac{\sum_{n=1}^{N}(f_n - \bar{f})(r_n - \bar{r})}{\sqrt{\sum_{n=1}^{N}(f_n - \bar{f})^2}\sqrt{\sum_{n=1}^{N}(r_n - \bar{r})^2}}$	-∞, 1	1
Root-mean-square error (RMSE)	$RMSE = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(f_n - r_n)^2}$	0, +∞	0

[a] Variables: F and f represent the model simulation results of binary classification and values respectively; R and r represent the reference data of binary classification and values respectively; 1 and 0 means positive (wet) and negative (dry) classifications; n and N represent sample index and a total number of samples.

Second, the 50 USGS WHMs in the study area were used as the reference to calculate the difference produced by the model simulations. Two metrics were used (Table 3.1), where the Correlation Coefficient measures the relationships between model-simulated inundation depth and WHMs and Root Mean Squared Error (RMSE) measures the average magnitude of the errors

from the model simulations. These traditional statistics test if the models capture the pattern and accurate water inundation depth.

3.3 Results and discussion

3.3.1 Flood extent evaluation

A display of flood maps from the USGS mapping, SAR data interpolation, NWM+HAND, Fathom (LISFLOOD-FP), and CREST-iMAP are listed in Figure 3.3. To visually compare all data sets, all data were cropped within the USGS HWMs interpolation boundaries, and the max values of each pixel in the modeled time series were taken from the model outputs. For the ease of observing, only the inundation depth pixel value that is larger than 1 inch (0.0254 m) was displayed in Figure 3.3.

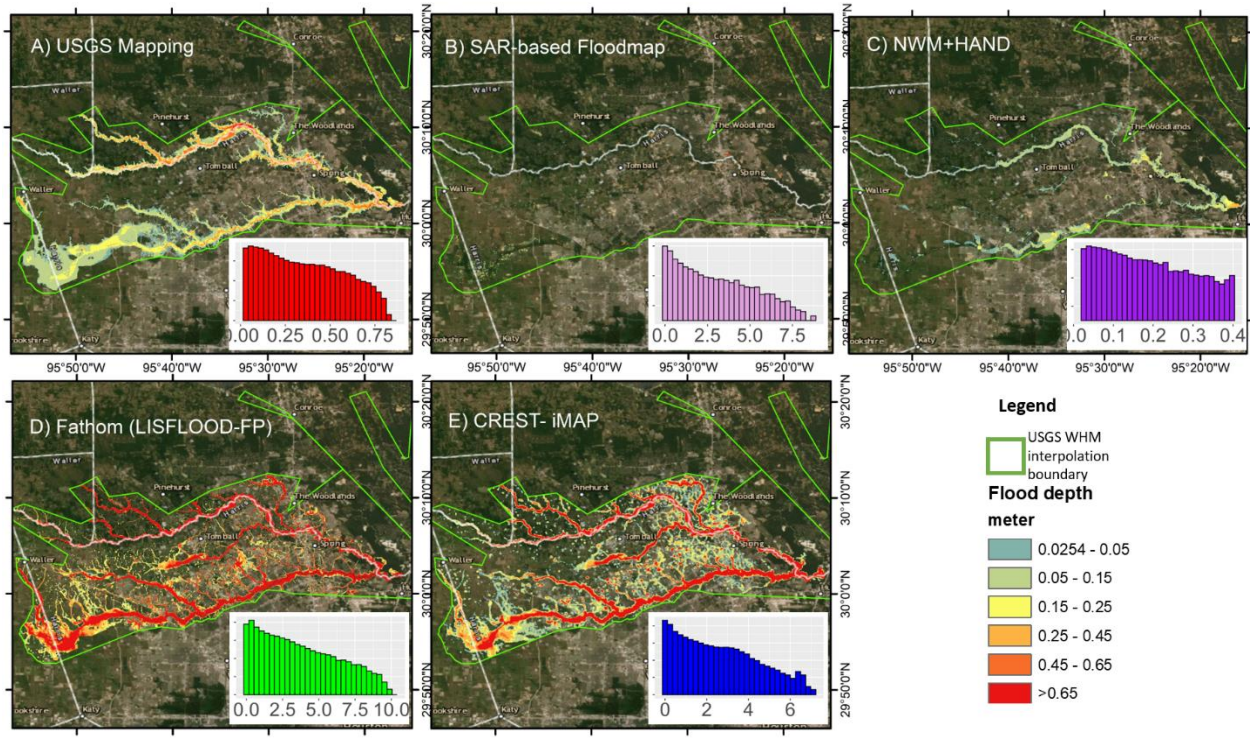


Figure 3.3. The flood extent and depth of A) USGS inundation mapping, B) SAR image flood mapping, C) NWM+HAND, D) Fathom model, and E) CREST-iMAP model, and their flood depth distributions

The satellite-based flood mapping was not able to capture most of the flood inundation compared to other methods. Therefore, the satellite-based flood mapping failed in this study area and will not be included for further analysis. However, the SAR data captures a small area of

flood at the southwestern corner of the boundary (green line), which indicates that the SAR and RAPID system performs slightly better at the ranch area instead of mountainous or urban area. The NWM+HAND method has a very limited flood extent but better than the satellite-based flood mapping. Especially, the NWM+HAND method could not capture the large flooded area at the upper stream of the Cypress Creek (lower-left corner of the study boundary), where USGS mapping, Fathom and CREST-iMAP all found inundation in the area. The USGS mapping only captures the fluvial flooding along the three main streams (Spring Creek, Cypress Creek, Little Cypress Creek). In contrast, the two modeling methods can capture more inundation along smaller channels. The major flooding occurred at the southwestern side of the basin, which was captured by the two simulations and the USGS interpolation. However, by looking at the flood depth distribution at subplots and the color scales of each map, the simulated inundation depths are quite different as the Fathom simulation output appears to have more "red" pixels than the other methods and its maximum inundation depth reaches over 10 meters. To note that the NWM+HAND method produced eight 'noise' pixels that have the values around 11 meters while the majority of the pixel values are less than 0.4 meters, which can be caused by the DEM error or during the data processing. One observation of the different flood maps is that the max flood depth ranges from 0.4 meter to over 10 meters, where the USGS mapping, Fathom, and CREST-iMAP all include the channel water depth, while SAR and NWM+HAND do not. The vary of flood depth can be caused by the different DEM treatments by each automated operating system, even though all automated methods use the DEM from USGS National Elevation Dataset (NED). Fathom has river channels "burned" into the DEM, while CREST-iMAP uses the native DEM and SAR flood map interpolation and NWM+HAND use ditch-filled DEM. The benchmark

USGS flood map uses its own lidar detected data as DEM, which is unique and different from all other methods. The more detailed flood depth analysis will be done in 3.3.2.

Table 3.2. The comparison results of NWM+HAND, Fathom (LISFLOOD-FP), and CREST- iMAP to the benchmark USGS flood mapping. The POD, FAR, and CSI were described in section 2 and Table 1, with the threshold of 1 inch (0.0254 meters)

Name	NWM+HAND	Fathom	CREST-iMAP
POD	0.22	0.72	0.72
FAR	0.13	0.40	0.45
CSI	0.21	0.49	0.45

Assuming the USGS flood mapping is the ground truth and is set as the benchmark of the study, and only over 1 inch (0.0254 meters) of water depth is considered as an inundated pixel, the comparison results are listed in Table 3.2. The previous study indicated that Spring Creek and San Jacinto River area were one of the poor performance basins (Wing et al., 2019). This study confirmed the performance where the POD indicates 72% of the area matched the benchmark, and the CSI was only 0.49. The CREST-iMAP has a comparable performance, where the POD was 72%, and CSI was 0.45. Both models produced false alarms, and their FARs were over 40%, which was caused by the large underrepresentation of pluvial flooding between the two main streams according to the benchmark USGS flood map (Figure 3.4. Blue area). Compared to other two models, the NWM+HAND approach dramatically underperformed giving only 22% detection and the CSI is only 0.21. As the NWM+HAND simulated flooded area is very small, the FAR is as low as 13%.

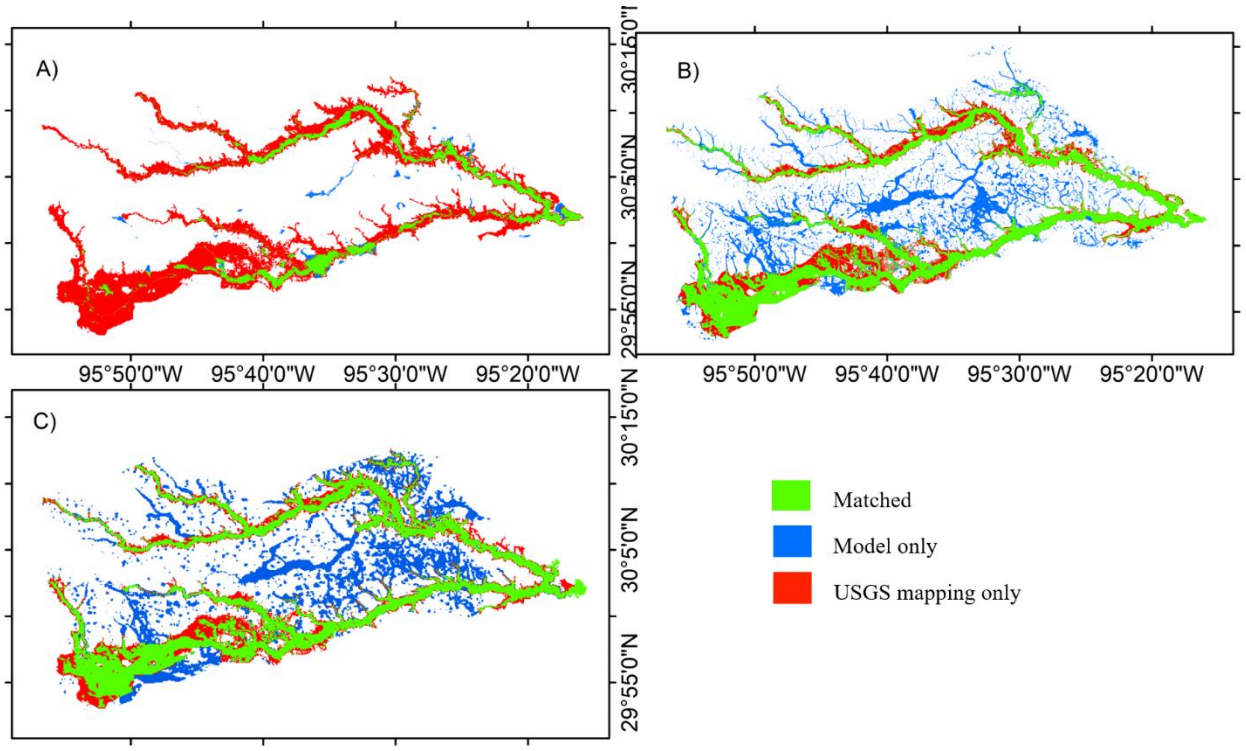


Figure 3.4. Maps were displaying the intersection of the (A) NWM+HAND, (B) Fathom, and (C) CREST-iMAP flood extents with those from the USGS flood mapping

Figure 3.4 demonstrates the spatial distribution of the match between the model simulations and the USGS flood map. The NWM+HAND shows clear underestimation where only the center lines of the two main downstreams appear to have inundations, and all the upper streams flooding are not captured. Since NWM-HAND only considers the overbank flow, there is no pluvial flood showing between two main streams, which leads to the low FAR value. Besides the large false alarms from Fathom and CREST-iMAP models, the general performance over two major streams appears to be good. CREST-iMAP performed better than Fathom on Spring Creek (northern stream), and Fathom performed better at the upper part of Cypress Creek (southern stream) than CREST-iMAP. Both model simulations did not fully capture the east end of the upper Cypress Creek flood, where Cypress City is located (29.98N 95.74W). The CREST-iMAP simulation showed a more inundated area in the eastern portion of the study area, which is more

developed and lower in elevation. Fathom simulation showed more flooding at the middle and upper streams of the study area, especially along the flowlines and small water channels.

It is reasonable to raise speculations about how under-representative the USGS flood mapping was for Hurricane Harvey. First, the methodology of the USGS flood mapping does not consider any physical water movement nor hydrological cycling dynamics. Second, both models showed a significant amount of false alarms in the area between two major rivers, which indicates the benchmark data might be underrepresenting. To investigate further, we counted the number of FEMA flood/water damage insurance claims that landed on the "wet" pixels of different flood maps. The results are illustrated in Figure 3.5.

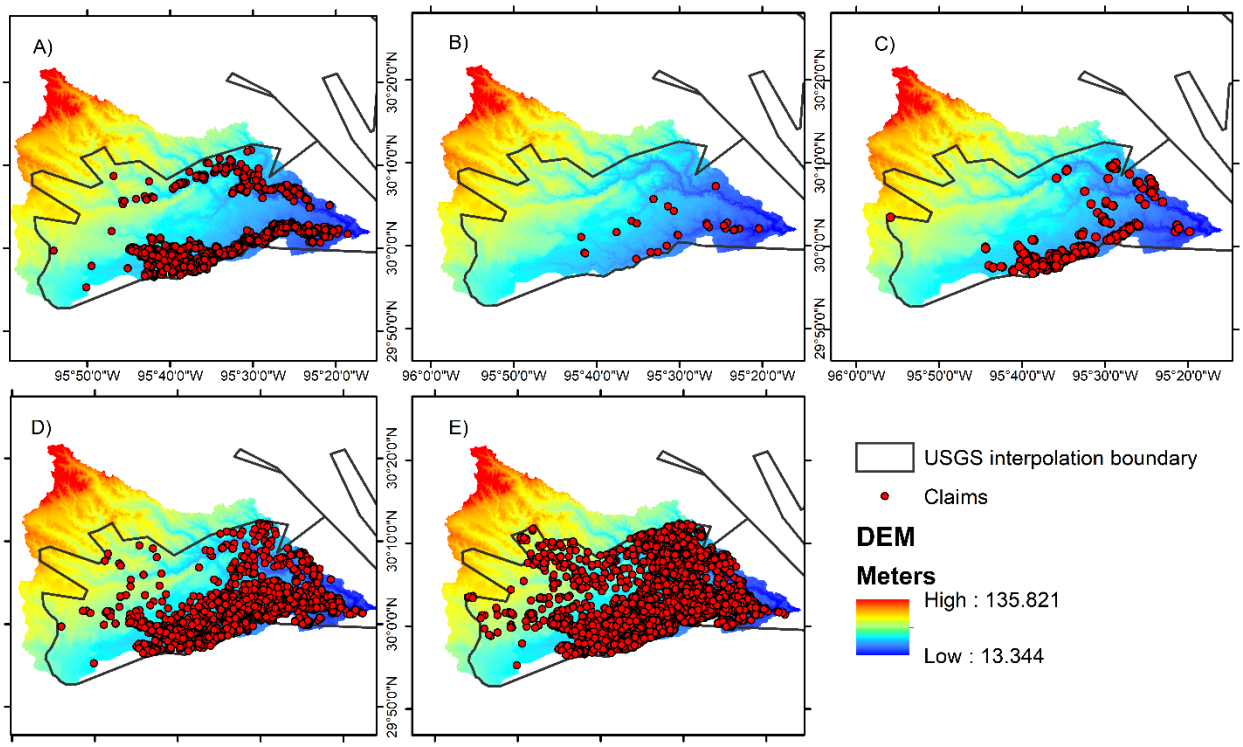


Figure 3.5. The filed FEMA flood/water damage insurance claims that land within the "wet" area from each flood mapping sources: A) USGS flood mapping, B) satellite-based flood extent, C) NWM+HAND, D) Fathom model simulation, and E) CREST-iMAP model simulation

The flood/water damage insurance claims included water damages, flood, water-related electrical damages, and property rupture or cracks, with a total of 10,459 cases within the study area. Since it is difficult to determine how much water could cause water damage, any pixel has value more than 0 meter was considered as a wet area for all datasets. The satellite-based flood map is very sparse, so there were only 107 (1.0%) claims collocated with the flood extent. The NWM+HAND is proven to underestimate the flooding by Hurricane Harvey, and there were only 1,692 (16.2%) claims inside the simulated flood extent. For the USGS flood map, there were 4,610 (44.1%) claims inside its flood extent approximation, which is mainly aligning along the two main river channels, as the USGS interpolation heavily focused on the fluvial flooding. The Fathom simulated flood extent has a comparable number of claims as USGS mapping of 4,284 (41.0%). However, the claims associated with flooded pixels in the Fathom simulation are visually more widespread compared to the USGS map, where it shows many more claims between the two main rivers (Figure 3.5c). We speculate that since no model simulation showed as much flood inundation area at the City of Cypress as the USGS mapping, concentrated flood/water damage claims (1,896 out of 10,459 claims) were partially counted in the Fathom simulation. Therefore, even though Figure 3.5c visually shows more widespread claims, the USGS flood map has a higher count of claims associated to inundated pixels. The CREST-iMAP simulated flood extent contains the most claims of 9,085 (86.9%), which indicates the benefit of the full pluvial flood simulation scheme can better capture the extreme precipitation triggered flood inundation. From Figure 3.5e, the claim points are well spread across the study area, and especially in the northern river, Spring Creek, the model simulation captured more flood/water damage insurance claims than all other methods.

Overall, it is difficult to conclude which flood prediction method is better than others or which methods can better be used as the benchmark data that represent the ground truth during the event. Theoretically, the satellite-based (SAR) flood map should have provided the "ground truth" as the data was derived from the snapshot when Sentinel-1 passed over the area on 2017/08/29 (Shen et al., 2019a). According to the CREST-iMAP simulated results, the study area had the most wet pixels at 11:45 am on 2017/08/29 during Hurricane Harvey, which is consistent with the date of the peak USGS gaged and EF5 simulated streamflow in a previous study (Chen et al., 2020). However, the SAR flood map shows the SAR and RAPID method only captures very small flood inundation at the southwestern ranches of the study area. As the water detection through vegetation and urban buildings from SAR data is not available for an automated system like RAPID, the lack of flood extent detection exposes the limitation of the satellite-based automated flood mapping system when detecting through large objects is not available. The USGS flood map has been used as the benchmark in other studies, but it is unrealistic to have no inundation away from the river channels, and only less than half (44.1%) of the flood insurance claims landed within its flood extent. The NWM+HAND method underperforms but the data was generated by NWM v1.1, and as of today, the model has developed to v2.0 and v2.1, which we can believe it should have better performs now. However, this NWM+HAND is still limited to the stage-flow conversion, DEM manipulation, and other systematic errors (Johnson et al., 2019). The Fathom and CREST-iMAP simulated flood extents showed promising results that can capture the pluvial flood. However, Fathom's simulation of flood extent captures less inundation at the downstream of the basin and around Spring Creek, which leads to fewer insurance cases within its flood extent (41%) compared to USGS mapping and CREST-iMAP simulation. The Fathom system uses the return period of the streams as the

main forcing to the model, and the infiltration capacity method might undermine the effects of the overland flow and subsurface flow. Meanwhile, LISFLOOD-FP uses local inertial method to simplify the St. Venant Shallow Water Equation, which requires individual flood surveyed flood extent for each catchment for calibration to reach optimal result, but it is hard to achieve for a continental scale, automated operational flood model (Horritt and Bates, 2002). The CREST-iMAP shows the broadest coverage over insurance claims in the basin (80+%). At the same time, the FAR (45%) is similar to the Fathom simulation (FAR=40%) compared with the USGS flood map, which means that the CREST-iMAP simulation has a comparable amount of overestimation as the other modeling approach in this study. Therefore, it is reasonable speculation that the CREST-iMAP simulation can cover the majority of the flood insurance claim is due to the better description of the water movements during Hurricane Harvey, by using the fully solved Shallow Water Equation instead of the simplified local inertial form in Fathom or a conceptual interpolation used in HAND. Both Fathom and CREST-iMAP do not represent the mid and upper Cypress creek flood very well, compared to USGS flood map, which is where the City of Cypress is located. This shows a certain degree of inability of simulating urban flooding by automated flood modelling, since Cypress city was heavily flooded during Hurricane Harvey based on news and social media posts.

The flood extent analysis shows that the CREST-iMAP model at its current setting has a comparable result as the operational Fathom system, by comparing with the benchmark USGS flood map, which also matches well with FEMA flood insurance claims. The flood extent from SAR and NWM+HAND are shown to be underrepresentative.

3.3.2 Flood inundation depth analysis

The model simulated inundation depth values were extracted from the same locations as the 50 WHM sites in the study area. Figure 3.6 shows the water depths at each location from different data sources.

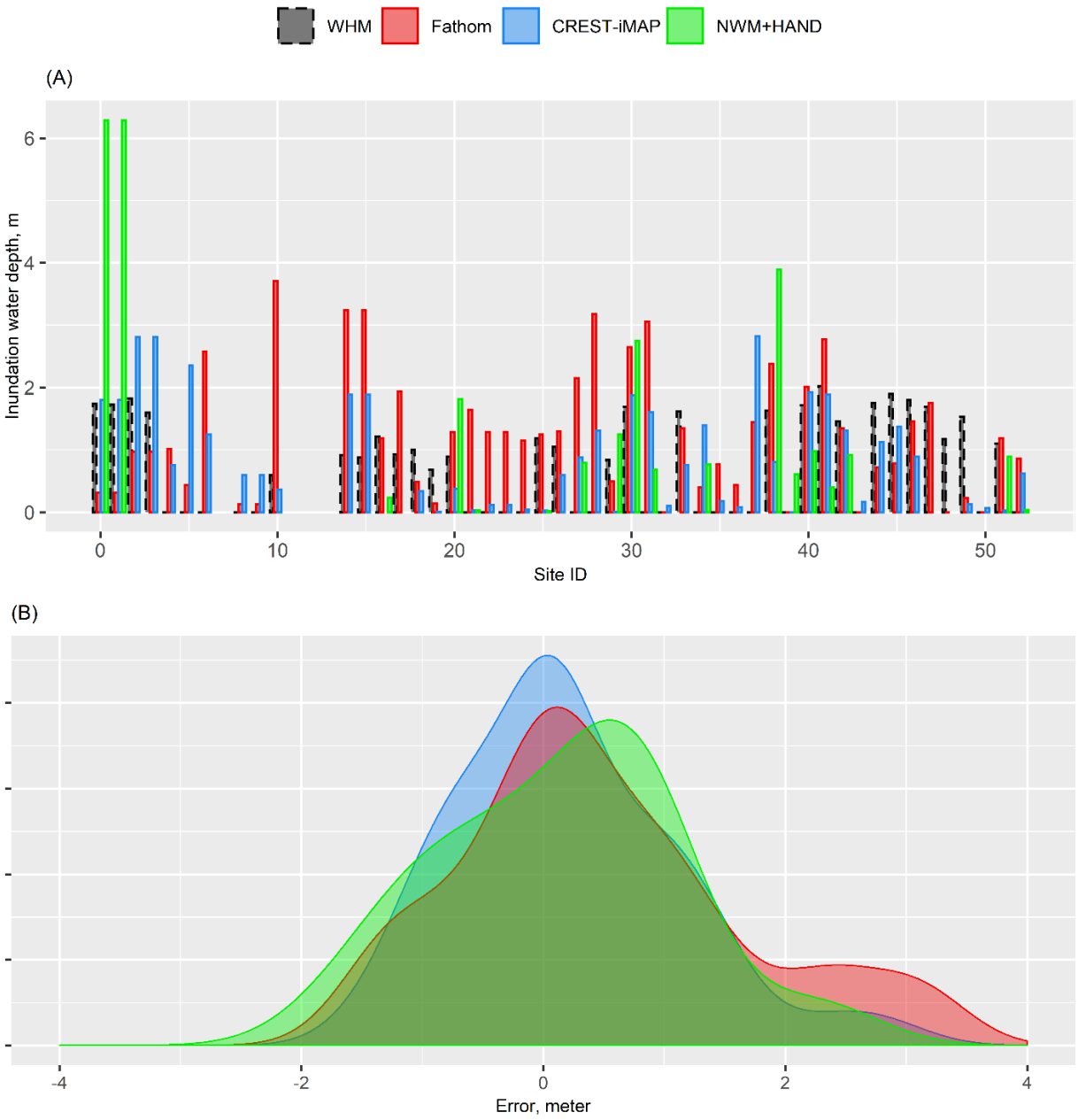


Figure 3.6. The (A) flood inundation depth at 50 USGS WHM locations and (B) error distributions of NWM+HAND simulation, Fathom simulation, and CREST-iMAP simulation

The first observation from the result is that the Fathom simulated flood depth has much higher extreme values compared to other sources, which is consistent with the results in Figure 3.3 where Fathom has more red area and its inundation value has the largest range among all flood approximations. The error distribution of Fathom simulation also indicates that there are multiple pixels overestimate the flood depth by 2 to 4 meters. Second, considering NWM+HAND does not have pixel values at 31 locations, the rest of the locations appears to slightly overestimate the flood depth, with two extreme values (error > 7 meter) that beyond the range of the figure. Third, no flood inundation data source aligns perfectly with the USGS WHM measurement and CREST-iMAP is the only model that limits the majority of the error withing ±2 meters. Therefore, no modeling method can approximate the actual ground survey in this study, which leaves room for much improvement in the field. The statistical analysis was done to analyze and compare the simulated data with the USGS WHM as the reference (Table 3.3).

Table 3.3. The traditional statistical analysis of HWM+HAND simulation, Fathom (LISFLOOD-FP) simulation, and CREST-iMAP simulation. The CC and RMSE were described in Section 3.2 and Table 3.1

Name	NWM+HAND	Fathom	CREST-iMAP
CC	0.51	0.12	0.38
RMSE, m	1.81	1.26	0.91
# of NAs	31	1	1

The results indicate that all flood inundation depth approximations have poor correlation with the USGS WHM records (<0.51) and about 1 meter of error. Despite the fact that there are 31 no value pixels extract at WHM locations from the NWM+HAND simulation result, its depth is the most correlated to the ground survey with CC of 0.51, but the least accurate with RMSE of 1.81

meters. The CREST-iMAP model simulation has a better performance, with CC of 0.36 and RMSE of 0.91 meters. The Fathom simulation ranks second-best with RMSE of 1.26 but worst on CC of 0.12. The scatter plot (Figure 7) shows the concentration of Fathom simulation results close to the left upper corner, which indicates the majority of overestimation occurs when the inundation water depth is less than 1 meter. All flood simulation methods yield overestimation when the WHM value is 0 (a very small flood inundation), and Fathom tends to have greater flood depth values than CREST-iMAP and NWM+HAND. Majority of the overestimation of CREST-iMAP occur at the locations with 0 value WHMs and many underestimations occur at the locations when WHMs are between 0.5 to 1.5 meters.

Overall, the correlations between the USGS WHM and different approximations were low, where there is no clear pattern along the isoline in this scatter plot. It is partially due to the inaccuracy of the flood inundation simulations, as well as the USGS WHM measurements, which is based on human observation and claims the uncertainty is between ±0.015 meters (0.05 ft) to ±0.12 meter (0.4 ft) (Feaster and Koenig 2017). Another reason could be the location information of the WHM sites is not accurate enough, which the USGS WHM data set provide the longitude and latitude coordinates at accuracy of 5 digits after the decimal point. However, all flood simulation models were operated under hyper resolution (3 m to 10 m). Therefore, only 10-5 arc-degree accuracy is not enough to precisely extract the right pixel from the modeled flood inundation result, which could cause the inconsistency that we found this study. The previous Hurricane Harvey study (Wing et al. 2019) also argued that the Fathom model error compare to WHM was close to 1.19 meters and justified that the ~1-meter model deviation from the WHM was acceptable and informative to flood prevention and preparations for local first

responders. This study provides another flood inundation approximation method that marginally improved the error (RMSE) to ~0.9 meter.

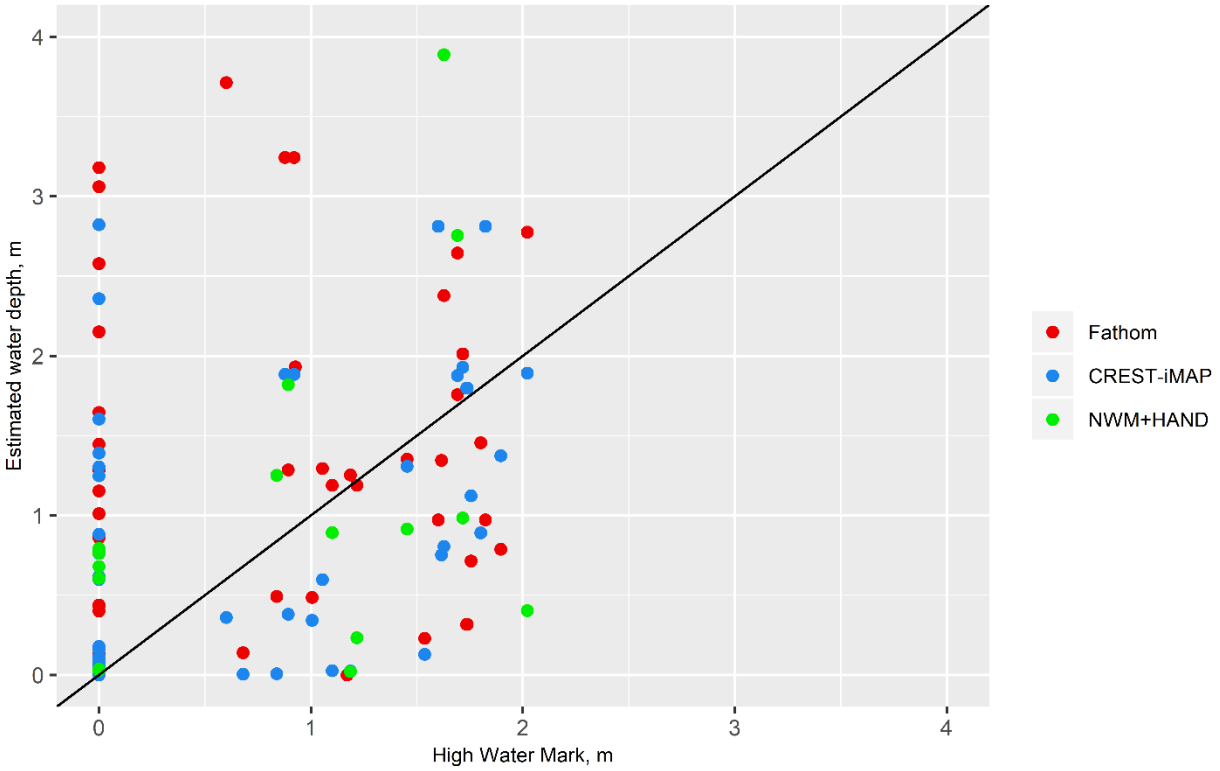


Figure 3.7. The scatter plot of NWM+HAND (green), Fathom (LISFLOOD-FP) simulation (red), and CREST-iMAP simulation (blue) compared with 50 USGS WHM records as the reference

3.4 Conclusion

Validation of a flood map for a single event is challenging since such an error-free reference data that comprehensively reflects a flood event is not available. This study provides a model application that utilizes a coupling scheme that uses excessive rainfall as the input for the hydraulic component. The results prove that the CREST-iMAP framework can well capture the flood extent and the spatial pattern of the flood extent, that is comparable with current automated operational flood-monitoring systems in the world. However, the error (RMSE) is approximately 0.9 meter using the traditional statistical method. This study compares the five different sources of flood inundation approximations for the flood induced by Hurricane Harvey in Spring Basin located northern Harris County, TX, which are RAPID extraction from SAR data, USGS WHM and stream stage interpolation, NWM+HAND model simulation, Fathom (LISFLOOD-FP) model simulation, and CREST-iMAP model simulation. The main conclusions of this study are: first, the CREST-iMAP modeling methods can capture the flood extent under extreme precipitation as well, if not better than, as other sources; second, satellite-based (SAR) flood observation and NWM+HAND model simulation severely underperformed during Hurricane Harvey and HAND only considers the fluvial flood; third, this study cannot conclude the most reliable method to capture the flood inundation during the extreme event as no method can completely reproduce the flood extent and the inundation depth errors are not negligible.

This study provides a hydrological and hydraulic coupled approach to simulate flood with less data requirement in an automated and operational setup using the CREST-iMAP framework, which yields acceptable result during the extreme precipitation-driven flood event: Hurricane Harvey. The current case study has 41 million computing pixels and a computational efficiency of 0.02 second per time step (10 seconds) using 2 nodes (40 computing cores), but a more

systematic computational efficiency study for the CREST-iMAP will be needed in the future. Even though the simulated flood depth is not perfect, there is plenty of decision-supportive information to potentially establish a high-resolution flood prediction system based on the weather radar network for not only southeastern Texas but along all high-precipitation intensity and flood-prone areas across the globe, as the model is physics-based, compatible with global data, built with parameter transformation module. Furthermore, as the CREST-iMAP framework can comprehensively provide streamflow, flood extent, flood inundation depth, and soil moisture outputs, it can easily connect to other interdisciplinary building blocks (e.g. insurance, supply chain, utility management) to further analysis and predict the consequent socio-economic impact from any flood event. Ultimately, an action-directing smart system can be built upon reliable flood predictions to guide public safety decision-making during flood hazards.

Reference

Adhikari, P., Hong, Y., Douglas, K.R., Kirschbaum, D.B., Gourley, J., Adler, R., Robert Brakenridge, G., 2010. A digitized global flood inventory (1998–2008): compilation and preliminary results. Nat Hazards 55, 405–422. https://doi.org/10.1007/s11069-010-9537-2

Alcrudo, F., 2004. A state of the art review on mathematical modelling of flood propagation.

Anselmo, V., Galeati, G., Palmieri, S., Rossi, U., Todini, E., 1996. Flood risk assessment using an integrated hydrological and hydraulic modelling approach: a case study. Journal of Hydrology 175, 533–554. https://doi.org/10.1016/S0022-1694(96)80023-0

Arctur, D., Boghichi, E., Tarboton, D., Maidment, D., Bales, J., Idaszak, R., Seul, M., Castronova, A.M., 2018. Hurricane Harvey 2017 Collection. https://doi.org/10.4211/hs.2836494ee75e43a9bfb647b37260e461

Ashley, S.T., Ashley, W.S., 2008. Flood fatalities in the United States. Journal of Applied Meteorology and Climatology 47, 805–817. https://doi.org/10.1175/2007JAMC1611.1.

Barredo, J.I., 2007. Major flood disasters in Europe: 1950–2005. Nat Hazards 42, 125–148. https://doi.org/10.1007/s11069-006-9065-2

Bates, P.D., De Roo, A.P.J., 2000. A simple raster-based model for flood inundation simulation. Journal of Hydrology 236, 54–77. https://doi.org/10.1016/S0022-1694(00)00278-X

Bates, P.D., Horritt, M.S., Fewtrell, T.J., 2010. A simple inertial formulation of the shallow water equations for efficient two-dimensional flood inundation modelling. Journal of Hydrology 387, 33–45. https://doi.org/10.1016/j.jhydrol.2010.03.027

Benito, G., Lang, M., Mariano Barriendos, M.C.L., Frances, F., Ouarda, T., Thorndycraft, V.R., 2004. Use of systematic, palaeoflood and historical data for the improvement of flood risk estimation. Review of scientific methods." Natural Hazards 31, 623–643.

Brauer, N.S., Basara, J.B., Homeyer, C.R., McFarquhar, G.M., Kirstetter, P.E., 2020. Quantifying Precipitation Efficiency and Drivers of Excessive Precipitation in Post-Landfall Hurricane Harvey. Journal of Hydrometeorology 21, 433–452. https://doi.org/10.1175/JHM-D-19-0192.1

Brunner, G.W., 2016. HEC-RAS, River Analysis System Hydraulic Reference Manual (User Manual). US Army Corps of Engineers, Hydrologic Engineering Center, Davis, California.

Chen, M., Nabih, S., Brauer, N.S., Gao, S., Gourley, J.J., Hong, Z., Kolar, R.L., Hong, Y., 2020. Can Remote Sensing Technologies Capture the Extreme Precipitation Event and Its Cascading Hydrological Response? A Case Study of Hurricane Harvey Using EF5 Modeling Framework. Remote Sensing 12, 445. https://doi.org/10.3390/rs12030445

Clark, R.A., Flamig, Z.L., Vergara, H., Hong, Y., Gourley, J.J., Mandl, D.J., Frye, S., Handy, M., Patterson, M., 2017. Hydrological Modeling and Capacity Building in the Republic of Namibia. Bulletin of the American Meteorological Society 98, 1697–1715. https://doi.org/10.1175/BAMS-D-15-00130.1

Cohen, S., Praskievicz, S., Maidment, D.R., 2018. Featured Collection Introduction: National Water Model. J Am Water Resour Assoc 54, 767–769. https://doi.org/10.1111/1752-1688.12664

Cohen, S., Raney, A., Munasinghe, D., Loftis, J.D., Molthan, A., Bell, J., Rogers, L., Galantowicz, J., Brakenridge, G.R., Kettner, A.J., Huang, Y.-F., Tsang, Y.-P., 2019. The Floodwater Depth Estimation Tool (FwDET v2.0) for improved remote sensing analysis of coastal flooding. Nat. Hazards Earth Syst. Sci. 19, 2053–2065. https://doi.org/10.5194/nhess-19-2053-2019

de Almeida, G.A.M., Bates, P., 2013. Applicability of the local inertial approximation of the shallow water equations to flood modeling: APPLICABILITY LOCAL INERTIAL. Water Resour. Res. 49, 4833–4844. https://doi.org/10.1002/wrcr.20366

Eric, B.S., Zelinsky, D.A., 2018. National Hurricane Center tropical cyclone report Hurricane Harvey. National Hurricane Center, Miami, Florida.

Feaster, T.D., Koenig, T.A., 2017. Field manual for identifying and preserving High-Water Mark data (Open-File Report No. 67). U.S. Geological Survey, Reston, Virginia.

Flamig, Z.L., Vergara, H., Gourley, J.J., 2020. The Ensemble Framework For Flash Flood Forecasting (EF5) v1.2: Description and Case Study (preprint). Hydrology. https://doi.org/10.5194/gmd-2020-46

Ghimire, B., Chen, A.S., Guidolin, M., Keedwell, E.C., Djordjević, S., Savić, D.A., 2013. Formulation of a fast 2D urban pluvial flood model using a cellular automata approach. Journal of Hydroinformatics 15, 676–686. https://doi.org/10.2166/hydro.2012.245

Gochis, D.J., Dugger, A.L., Yu, W., Yates, D.N., Sampson, K., Barlage, M., Pan, L., Zhang, Y., McCreight, J.L., RafieeiNasab, A., Karsten, L., Read, L., Gaydos, A., McAllister, M., Mills, J., Towler, E., Grim, J., FitzGerald, K., 2017. The NOAA National Water Model: Research to Operations to Research (Presentation). CUAHSI, Boulder, CO.

Gourley, J.J., Flamig, Z.L., Vergara, H., Kirstetter, P.-E., Clark, R.A., Argyle, E., Arthur, A., Martinaitis, S., Terti, G., Erlingis, J.M., Hong, Y., Howard, K.W., 2017. The FLASH Project: Improving the Tools for Flash Flood Monitoring and Prediction across the United States. Bulletin of the American Meteorological Society 98, 361–372. https://doi.org/10.1175/BAMS-D-15-00247.1

Hong, Y., Hsu, K.-L., Moradkhani, H., Sorooshian, S., 2006. Uncertainty quantification of satellite precipitation estimation and Monte Carlo assessment of the error propagation into hydrologic response. Water Resource Research 42. https://doi.org/10.1029/2005WR004398.

Horritt, M.S., Bates, P.D., 2002. Evaluation of 1D and 2D numerical models for predicting river flood inundation. Journal of Hydrology 268, 87–99. https://doi.org/10.1016/S0022-1694(02)00121-X

Johnson, J.M., Munasinghe, D., Eyelade, D., Cohen, S., 2019. An integrated evaluation of the National Water Model (NWM)–Height Above Nearest Drainage (HAND) flood mapping methodology. Nat. Hazards Earth Syst. Sci. 19, 2405–2420. https://doi.org/10.5194/nhess-19-2405-2019

Koenig, T.A., Bruce, J.L., O'Connor, J., McGee, B.D., Holmes, Jr, R.R., Hollins, R., Forbes, B.T., Kohn, M.S., Schelleken, M.F., Martin, Z.W., Peppler, M.C., 2016. Identifying and preserving High-Water Mark data, in: U.S. Geological Survey Techniques and Methods. U.S. Geological Survey, Reston, Virginia, p. 47.

Kossin, J.P., 2018. A global slowdown of tropical-cyclone translation speed. Nature 558, 104–107. https://doi.org/10.1038/s41586-018-0158-3

Lehner, B., Verdin, K., Jarvis, A., 2008. New Global Hydrography Derived From Spaceborne Elevation Data. Eos Trans. AGU 89, 93. https://doi.org/10.1029/2008EO100001

Lejot, J., Delacourt, C., Piégay, H., Fournier, T., Trémélo, M.-L., Allemand, P., 2007. Very high spatial resolution imagery for channel bathymetry and topography from an unmanned mapping controlled platform. Earth Surf. Process. Landforms 32, 1705–1725. https://doi.org/10.1002/esp.1595

Li, Z., Chen, M., Gao, S., Hong, Z., Tang, G., Wen, Y., Gourley, J.J., Hong, Y., 2020. Cross-Examination of Similarity, Difference and Deficiency of Gauge, Radar and Satellite Precipitation Measuring Uncertainties for Extreme Events Using Conventional Metrics and Multiplicative Triple Collocation. Remote Sensing 12, 1258. https://doi.org/10.3390/rs12081258

Liu, Z., Merwade, V., Jafarzadegan, K., 2019. Investigating the role of model structure and surface roughness in generating flood inundation extents using one- and two-dimensional hydraulic models. J Flood Risk Management 12, e12347. https://doi.org/10.1111/jfr3.12347

Maidment, D.R., 2017. Conceptual Framework for the National Flood Interoperability Experiment. J Am Water Resour Assoc 53, 245–257. https://doi.org/10.1111/1752-1688.12474

McDonald, W.M., Naughton, J.B., 2019. Impact of hurricane Harvey on the results of regional flood frequency analysis. J Flood Risk Management 12. https://doi.org/10.1111/jfr3.12500

Merwade, V.M., Maidment, D.R., Goff, J.A., 2006. Anisotropic considerations while interpolating river channel bathymetry. Journal of Hydrology 331, 731–741. https://doi.org/10.1016/j.jhydrol.2006.06.018

Morin, J., Benyamini, Y., 1977. Rainfall infiltration into bare soils. Water Resour. Res. 13, 813–817. https://doi.org/10.1029/WR013i005p00813

Murphy, J.D., 2018. Service assessment August–September 2017 Hurricane Harvey (US DOC). NOAA National Weather Service, Silver Spring, Maryland.

Nielson, O.M., Roberts, S.G., Gray, R., D., McPherson, A., Hitchman, A., 2005. Hydrodynamic modelling of coastal inundation. Presented at the MODSIM 2005 International Congress on Modelling and Simulation, Modelling and Simulation Society of Australia & New Zealand, pp. 518–523.

Orzech, M.D., Reniers, A.J.H.M., Thornton, E.B., MacMahan, J.H., 2011. Megacusps on rip channel bathymetry: Observations and modeling. Coastal Engineering 58, 890–907. https://doi.org/10.1016/j.coastaleng.2011.05.001

Rappaport, E.N., 2014. Fatalities in the United States from Atlantic Tropical Cyclones: New Data and Interpretation. Bulletin of the American Meteorological Society 95, 341–346. https://doi.org/10.1175/BAMS-D-12-00074.1

Rappaport, E.N., 2000. Loss of Life in the United States Associated with Recent Atlantic Tropical Cyclones. Bull. Amer. Meteor. Soc. 81, 2065–2074. https://doi.org/10.1175/1520-0477(2000)081<2065:LOLITU>2.3.CO;2

Roberts, S.G., Nielson, O.M., Gray, D., Sexton, J., Davis, G., 2015. ANUGA User Manual 2.0. (Manual). Commonwealth of Australia (Geoscience Australia) and the Australian National University.

Sampson, C.C., Smith, A.M., Bates, P.D., Neal, J.C., Alfieri, L., Freer, J.E., 2015. A high-resolution global flood hazard model: A HIGH-RESOLUTION GLOBAL FLOOD HAZARD MODEL. Water Resour. Res. 51, 7358–7381. https://doi.org/10.1002/2015WR016954

Shen, X., Anagnostou, E.N., Allen, G.H., Robert Brakenridge, G., Kettner, A.J., 2019a. Near-real-time non-obstructed flood inundation mapping using synthetic aperture radar. Remote Sensing of Environment 221, 302–315. https://doi.org/10.1016/j.rse.2018.11.008

Shen, X., Hong, Y., Zhang, K., Hao, Z., 2017. Refining a Distributed Linear Reservoir Routing Method to Improve Performance of the CREST Model. J. Hydrol. Eng. 22, 04016061. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001442

Shen, X., Wang, D., Mao, K., Anagnostou, E., Hong, Y., 2019b. Inundation Extent Mapping by Synthetic Aperture Radar: A Review. Remote Sensing 11, 879. https://doi.org/10.3390/rs11070879

Smith, K., Ward, R., 1998. Floods: Physical processes and human impact. John Wiley, Chichester.

Syvitski, J.P.M., Brakenridge, G.R., 2013. Causation and avoidance of catastrophic flooding along the Indus River, Pakistan. GSAT 23, 4–10. https://doi.org/10.1130/GSATG165A.1

Tanaka, T., Yoshioka, H., Siev, S., Fujii, H., Fujihara, Y., Hoshikawa, K., Ly, S., Yoshimura, C., 2018. An Integrated Hydrological-Hydraulic Model for Simulating Surface Water Flows of a Shallow Lake Surrounded by Large Floodplains. Water 10, 1213. https://doi.org/10.3390/w10091213

Teng, J., Jakeman, A.J., Vaze, J., Croke, B.F.W., Dutta, D., Kim, S.S.H., 2017. Flood inundation modeling: A review of methods, recent advances and uncertainty analysis. Environmental Modelling & Software 90, 201–216.

van Oldenborgh, G.J., van der Wiel, K., Sebastian, A., Singh, R., Arrighi, J., Otto, F., Haustein, K., Li, S., Vecchi, G., Cullen, H., 2018. Corrigendum: Attribution of extreme rainfall from Hurricane Harvey, August 2017 (2017 *Environ. Res. Lett.* **12** 124009). Environ. Res. Lett. 13, 019501. https://doi.org/10.1088/1748-9326/aaa343

Vergara, H., Kirstetter, P.-E., Gourley, J.J., Flamig, Z.L., Hong, Y., Arthur, A., Kolar, R., 2016. Estimating a-priori kinematic wave model parameters based on regionalization for flash flood forecasting in the Conterminous United States. Journal of Hydrology 541, 421–433. https://doi.org/10.1016/j.jhydrol.2016.06.011

Vrugt, J.A., 2016. Markov chain Monte Carlo simulation using the DREAM software package: Theory, concepts, and MATLAB implementation. Environmental Modelling & Software 75, 273–316. https://doi.org/10.1016/j.envsoft.2015.08.013

Vrugt, J.A., Braak, C.J.F., Gupta, H.V., Robinson, B.A., 2009. Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling? Stochastic Environmental Research and Risk Assessment 23, 1101–1026. https://doi.org/10.1007/s00477-008-0274-y.

Vu, T.M., Mishra, A.K., 2019. Nonstationary frequency analysis of the recent extreme precipitation events in the United States. Journal of Hydrology 575, 999–1010. https://doi.org/10.1016/j.jhydrol.2019.05.090

Wallemacq, P., Delbiso, T., Below, R., Wahlström, M., McClean, D., Landelle, S., Leoni, B., House, R., 2015. The human cost of weather related disasters 1995-2015. United Nation Office for Disaster Risk Reduction, Brussels, Belgium.

Wang, J., Hong, Y., Li, L., Gourley, J.J., Khan, S.I., Yilmaz, K.K., Adler, R.F., Policelli, F.S., Habib, S., Irwn, D., Limaye, A.S., Korme, T., Okello, L., 2011. The coupled routing and excess storage (CREST) distributed hydrological model. Hydrological Sciences Journal 56, 84–98. https://doi.org/10.1080/02626667.2010.543087

Watson, K.M., Harwell, G.R., Wallace, D.S., Welborn, T.L., Stengel, V.G., McDownwell, J.S., 2018. Characterization of Peak Streamflows and Flood Inundation of Selected Areas in Southeastern Texas and Southwestern Louisiana from the August and September 2017 Flood Resulting from Hurricane Harvey (Scientific Investigations Report No. 44). U.S. Geological Survey, Reston, Virginia.

Wing, O.E.J., Bates, P.D., Sampson, C.C., Smith, A.M., Johnson, K.A., Erickson, T.A., 2017. Validation of a 30 m resolution flood hazard model of the conterminous United States: 30

m RESOLUTION FLOOD MODEL OF CONUS. Water Resour. Res. 53, 7968–7986. https://doi.org/10.1002/2017WR020917

Wing, O.E.J., Sampson, C.C., Bates, P.D., Quinn, N., Smith, A.M., Neal, J.C., 2019. A flood inundation forecast of Hurricane Harvey using a continental-scale 2D hydrodynamic model. Journal of Hydrology X 4, 100039. https://doi.org/10.1016/j.hydroa.2019.100039

Wood, E.F., Roundy, J.K., Troy, T.J., van Beek, L.P.H., Bierkens, M.F.P., Blyth, E., de Roo, A., Döll, P., Ek, M., Famiglietti, J., Gochis, D., van de Giesen, N., Houser, P., Jaffé, P.R., Kollet, S., Lehner, B., Lettenmaier, D.P., Peters-Lidard, C., Sivapalan, M., Sheffield, J., Wade, A., Whitehead, P., 2011. Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water: OPINION. Water Resour. Res. 47. https://doi.org/10.1029/2010WR010090

Zhang, J., Howard, K., Langston, C., Kaney, B., Qi, Y., Tang, L., Grams, H., Wang, Y., Cocks, S., Martinaitis, S., Arthur, A., Cooper, K., Brogden, J., Kitzmiller, D., 2016. Multi-Radar Multi-Sensor (MRMS) Quantitative Precipitation Estimation: Initial Operating Capabilities. Bulletin of the American Meteorological Society 97, 621–638. https://doi.org/10.1175/BAMS-D-14-00174.1

Zhang, J., Huang, Y.-F., Munasinghe, D., Fang, Z., Tsang, Y.-P., Cohen, S., 2018. Comparative Analysis of Inundation Mapping Approaches for the 2016 Flood in the Brazos River, Texas. J Am Water Resour Assoc 54, 820–833. https://doi.org/10.1111/1752-1688.12623

Chapter 4. A flood predictability study for Hurricane Harvey with the CREST-iMAP model using high resolution Quantitative Precipitation Forecast and U-Net deep learning precipitation nowcast

Abstract

Flooding is one of the most hazardous natural disasters, and it commonly causes fatalities and social-economical damages. The advances of modeling techniques and data in flood prediction have found success operationally, and this paper presents a more comprehensive flood prediction of Hurricane Harvey in 1-hour lead-time that is not limited to 1D streamflow forecast but also 2D flood extent and 3D inundation depth. It uses high-resolution quantitative precipitation forecasts (QPFs, from operational Rapid Refresh-RAP, and High Resolution Rapid Refresh-HRRR models) and deep learning nowcasts (AI nowcasts). The results show that the QPFs have a well-known displacement issue and the AI nowcast cannot predict the precipitation intensity, and an attempt to combine the two methods (AI hybrid) failed to improve the overall accuracy. However, the 2D flood extent predictions with the HRRR and AI hybrid forcings can provide information indicating the future flooded area with about 50% accuracy (hit rate). In contrast, the AI nowcast reveals minimal displacement errors but underpredicts precipitation intensity. The deep learning method also indicates that the binary tests with low threshold, which are commonly employed in the deep learning field, neglect the importance of precipitation intensity errors for extreme event studies.

4.1 Introduction

The increase of frequency and magnitude of extreme weather and precipitation caused by climate change has received much attention (Meehl et al., 2000; USGCRP, 2017; van Oldenborgh et al., 2018). Further, 'tropicalization', a global trend towards a heavier precipitation climate, has been modeled in Europe (Gobiet et al., 2014). Consequentially, this change would lead to more flood events in the future, which is the second-deadliest and most common natural hazard in the United States and world (Ashley and Ashley, 2008; Barredo, 2007; Benito et al., 2004; Smith and Ward, 1998). According to Brauer et al., (2020), the rate and duration of precipitation, land use and soil moisture content, as well as topography all have impacts on the severity of flooding. Unfortunately, tropicalization combining with rising sea levels and ocean temperatures will only intensify flooding especially along coastal areas (Wing et al., 2019). To quantify and mitigate flood risks, many tools have been developed to simulate flooding in post-event fashion (Bates et al., 2010; Bates and De Roo, 2000; de Almeida and Bates, 2013; Srinivasan and Arnold, 1994; Wang et al., 2011; Wood et al., 1992; Xue et al., 2016), and in real-time simulation (Cohen et al., 2018; Gochis et al., 2017; Sampson et al., 2015; Wing et al., 2019, 2017). Instead of post-event analysis or real-time simulation, this study explores the options for forecasting flood events, which could potentially provide enough lead-time for emergency response officials or even the general public to proactively respond to possible future flood risks.

Almost 40 years ago, Georgakakos and Hudlow, (1984) discuss the use of Quantitative Precipitation Forecasts (QPFs) from the Numerical Weather Prediction (NWP) models for hydrological forecasting; the operational Limited-Area Fine Mesh (LFM) model (Newell and Deaver, 1981) of the National Weather Service at the time had a grid spacing of 127 km, while

the on-demand Movable Fine Mesh (MFM) model had a grid spacing of up to 60 km. While promising, the lack of spatial specificity needed for hydrological forecasting and the relatively low accuracy of QPF limited the application of model-based QPF for driving hydrological models. With the advancement of computing and observation technologies, the NWP models have improved dramatically in recent decades in terms of their spatial and temporal resolutions, accuracy, and coverage (Trenberth, 2010). The current operational High-Resolution Rapid Refresh (HRRR, Benjamin et al., 2016; Lee et al., 2019) has a horizontal grid spacing of 3 km with hourly updated forecasts over the contiguous United States (CONUS). Convection-allowing, cloud-resolving NWP models have been run at 1-km grid spacing (Loken et al., 2017; Xue et al., 2013). However, accurate QPF remains a challenge because of many complex factors involved in precipitation processes (Ebert and McBride, 2000; Golding, 2000). Major errors of NWP QPFs include spatial displacement errors, where the location of the precipitating storm is in the wrong location (Ebert and McBride, 2000), and errors in intensity especially for extreme events (Cuo et al., 2011). For short-range precipitation forecasting, the proper initialization of existing precipitation systems within the model initial condition is also an important issue (Kain et al., 2010; Sun et al., 2014).

Precipitation nowcasting typically refers to very short-range (1-6 hr) forecasting of precipitation, and traditionally nowcasting is mostly done by extrapolating observed radar reflectivity fields. Ligda (1953) demonstrated the possibility of providing reasonable forecasts based on the persistence and movement of radar echoes. This technology has experienced a recent technological boost with the combination of High Performance Computing (HPC), Artificial Intelligence (AI), resulting in successful applications in real-world problems including precipitation nowcasting (Agrawal et al., 2019; LeCun et al., 2015). In 2019, Google Research

published a study showing the advantage of using deep learning nowcasts to predict 1-hour lead-time precipitation over the CONUS, outperforming traditional optical flow extrapolation method, persistence, and 1-hour HRRR forecasts (Agrawal et al., 2019). Moreover, many recent studies also show successes in deep learning precipitation nowcasting. Franch et al. (2020) introduce a new deep learning model, TAASRAD19, to forecast 1-hour lead-time radar reflectivity which is then converted to precipitation rate, yielding a Critical Success Index (CSI) of 0.5 over a 9-year sample. While most AI nowcast studies evaluate binary coverage as the performance test, Kumar et al. (2020) demonstrate their deep learning model reduces the Root-Mean-Square Error (RMSE) between forecast and satellite precipitation estimates to 0.8 mm/hr on 30 min lead-time and up to 1.4 mm/hr on 150 min lead-time. Using Global Precipitation Mission (GPM) Integrated Multi-satellitE Retrievals (IMERG) datasets, they identify the most significant error source being underestimation with extreme events. However, the hydrological performance of AI nowcasts is generally underexplored.

Flood prediction has a history of using Quantitative Precipitation Estimates (QPEs), precipitation nowcasts, and NWP QPFs to drive hydrological models (Cuo et al., 2011; Golding, 2000; Hapuarachchi et al., 2011). Since there is a lag between the precipitation peak and the streamflow peak (Bedient et al., 1988), some operational systems are capable of predicting the flood without precipitation forecasts. For example, the Scottish Environment Protection Agency uses radar and in-situ precipitation data to provide streamflow forecasts (Werner and Cranston, 2009); the National Center of Environmental Prediction (NCEP) adapts the National Water Model (NWM) (Cohen et al., 2018; Gochis et al., 2017), while the National Severe Storm Laboratory (NSSL) utilizes the Ensemble Framework For Flash Flood Forecasting (EF5) as part of the FLASH project (Flamig et al., 2020; Gourley et al., 2017). Both NWM and EF5 use the

Multi-Radar Multi-Sensor (MRMS) QPE to drive streamflow forecasts for the entire CONUS and outer territories. To provide longer range and more comprehensive flood predictions, it is necessary to utilize nowcasts or QPFs, and the flood predictions can be extended beyond streamflow prediction to also include two-dimensional (2D) flood extent and flood depth forecasting. However, due to the accuracy issues of QPF, the skill of longer flood range prediction is still limited, especially at the scales of urban watersheds (Hapuarachchi et al., 2011). More recent hydrological evaluation of operational HRRR QPF indicates that even though this advanced product is capable of forecasting mesoscale convective systems well, QPFs associated with smaller scale precipitation systems often contain significant errors making hydrological prediction errors to become too large (Lee et al., 2019; Seo et al., 2018).

This study is an attempt to utilize advanced NWP QPFs and AI precipitation nowcasting for the prediction of flooding triggered by the landfalling Hurricane Harvey in 2017. The comprehensive predictions include 1D stream discharge, 2D flood extent, and 3D flood depth, using the Coupled Routing and Excessive STorage – inundation MApping and Prediction (CREST-iMAP) model. This work evaluates the ability of state-of-art operational QPFs to predict not only streamflow but also flood extents and flood depths, it also presents one of the first studies to evaluate the hydrological performance of AI precipitation nowcast techniques. The precipitation forecasts and flood prediction results are compared with MRMS QPE, US Geological Survey (USGS) stream gauge data, MRMS QPE simulated flood extent, Federal Emergency Management Agency (FEMA) flood insurance claims, and USGS Water-High Mark post-event survey data. The rest of this paper is organized as follows. Section 2 describes the QPF datasets, study area, CREST-iMAP model, and the deep learning architecture for AI nowcasting. Section 3 demonstrates the results of the predicted streamflow hydrographs, flood

extent, and inundation depth compared to the benchmark and other ground observations. Section 4 discusses the major findings from the results. Section 5 concludes and proposes future studies.

4.2. Methodology

4.2.1 Forecast rainfall

This study relies on the applications of operational NWP models that provides the QPF products, which are archived in NOAA's National Operational Model Archive and Distribution System. The Rapid Refresh (RAP) was first developed in May 2012, built upon the first hourly updated operational NWP system in the world, Rapid Update Cycle (RUC), back in 1998 (Benjamin et al., 2016, 2004). After two major improvements in 2014 and 2016, RAP now consists of multiple meteorological data and models to increase accuracy, such as the NOAA Gridpoint Statistical Interpolation (GSI), a version of Weather Research and Forecasting (WRF) regional model, and the National Center of Environmental Prediction (NCEP) Unified Post Processor (UPP). The current RAP QPF product can provide hourly updated precipitation forecasts out to 18 hours at 13-km spatial resolution over the entire North America. In this study, the 1-hour lead-time data was obtained from the NOAA's National Center for Environmental Information (NCEI) website (https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/rapid-refresh-rap). The HRRR was first released in September 2014 and finished its first upgrade in 2016, then was deployed as a new operational model in July 2018 (Lee et al., 2019). HRRR is a nested model that is heavily relied on RAP data assimilation, which covers the Continental United States (CONUS) and provides hourly forecasts up to 18 hours in the future at 3-km spatial resolution (Benjamin et al., 2016). In this study, the 1-hour lead-time HRRR v2 forecast data obtained from the University of Utah HRRR data archive (http://home.chpc.utah.edu/~u0553130/Brian_Blaylock/cgi-bin/hrrr_download.cgi). All forecast products are 1-hour lead-time since the deep learning precipitation nowcasting method in this

study can only produce a 1-hour forecast due to the limitation of computational resources. The basic statistics of precipitation forecast datasets are listed in Table 4.1.

Table 4.1. List of precipitation estimation and forecasting productions and the basic statistics during Hurricane Harvey

	Temporal resolution	Spatial resolution	Precipitation rate, mm/hr		Total precipitation (25/8/2017 to 31/8/2017), mm		
Name			Max	Mean	Max	Mean	Min
MRMS QPE	2 mins	1 km	202.60	6.16	1793	830	287
RAP QPF	1 hour	13 km	49.83	7.45	1469	1081	730
HRRR QPF	1 hour	3 km	236.87	6.65	1640	928	502
AI Nowcast	6 mins	1 km	71.27	3.18	851	190	80
AI Hybrid	6 mins	1 km	236.87	7.10	2633	930	419

4.2.2 Deep learning rainfall nowcast

The advancement in machine learning (deep learning) recently has been successful when applied to the short-range nowcasting of precipitation (Shi et al., 2015). AI precipitation nowcasting, as a data-driven and localized method, is adapted to the local environment and requires no prior knowledge about weather systems from the model developer, which predicts next-step precipitation distribution by a series of previous frames from a well-trained model structure. In this study, we select the U-Net model structure for the following reasons. First, it is a widely adopted structure for precipitation estimation and prediction (Li et al., 2020b; Sadeghi et al., 2020). Second, the lightweight framework necessitates a fewer number of parameters compared to other models, also requiring less system memory to train the model. Before model training, some preprocessing steps are conducted for unifying precipitation estimates from multiple events. Due to the nature of the non-Gaussian distribution of precipitation per frame,

one way to normalize precipitation data is via logarithmic transformation. Similar to Sønderby et al. (2020), precipitation data is transformed and normalized based on the following equation.

$$R = \log(x + 0.01)\,/4,$$

where x is the input rainfall rate, and R is the normalized rainfall rates. Subsequently, rainfall data are grouped every 10 frames at 6-min increments to target next-hour precipitation forecasts. The Mean Square Error (MSE) and Adam, a deep-learning optimizer (Kingma and Ba, 2017), are chosen as the objective function and optimizer in this setting. The initial learning rate (α)[1] is set at 0.001 but is scheduled to decrease exponentially with training process to avoid blowing up. Due to the limitation of the Graphics Processing Unit (GPU) capacity, this study can only predict a 1-hour lead time given the size of the study area. In this study, the U-Net structured deep learning model is trained by the MRMS QPEs for 16 precipitation events listed in NOAA Storm Report (Table 2) that caused flooding in the Houston area from 2015 to 2019, except Hurricane Harvey, with a total of over 30,000 images. The intensive-precipitation-focused deep learning model is then forecast 1-hour lead-time from 24/08/2017 to 02/09/2017 (10 days). The experiment is conducted using a Nvidia GTX 960M GPU card.

AI-based nowcasting can effectively capture the spatiotemporal correlation with observations, yet it may misrepresent the event magnitude, which is a common weakness of AI nowcast (Kumar et al., 2020). On the other hand, the physical simulations are apt at identifying storm cores and thus the event rainfall magnitude. It is presumably advantageous to take the respective advantages of both to produce a hybrid product. In this study, a conventional non-

[1] Learning rate is a hyperparameter that controls how much change the deep-learning model in response to estimated error each time the model weights are updated. Some considers the learning rate is the most important hyperparameter for neural network configuring.

parametric probability matching scheme is taken at each nowcast time step to increase the rainfall magnitude of the AI-produced product close to the HRRR forecasts. The basic idea is to conserve the ranks of the initial rainfall rate for a given frame, and its empirical cumulative density function (CDF) is modified to match the target CDF.

Table 4.2 List of heavy precipitation events used for deep-learning training

Index	Start location	Flood date	County	Fatality	Damage (million dollars)
1	Orr St.	31/10/2015	Harris, TX	2	1.7
2	SW Houston St.	18/03/2016	Harris, TX	0	0
3	NW Hockley St.	18/04/2016	Harris, TX	8	51
4	Hooks homepark	01/06/2016	Harris, TX	0	0.005
5	SW Almeda Rd.	18/01/2017	Harris, TX	0	0.5
6	S Houston St.	29/03/2017	Harris, TX	2	0.62
7	NE McNair St.	04/06/2017	Harris, TX	0	0.01
8	NE Little York Rd.	24/06/2017	Harris, TX	0	0.001
9	Satsuma Dr.	09/07/2017	Harris, TX	0	0
10	S Deer Park	21/09/2017	Harris, TX	0	0.005
11	NE Spring St.	04/07/2018	Harris, TX	0	0
12	SE Englewood St.	08/12/2018	Harris, TX	0	0
13	NW Huffman	07/05/2019	Harris, TX	0	0.25
14	NW Katy	09/05/2019	Harris, TX	0	0.05
15	E Wallis Rd.	23/08/2019	Harris, TX	0	0
16	NE Humble St.	19/09/2019	Harris, TX	2	565

4.2.3 Observed rainfall, flood benchmark, and data

The observed precipitation product is chosen to be the MRMS radar-only, 2-min QPE, which showed high correlations and small error compared to the Harris County Flood Control District rain gauge network in previous studies (Chen et al., 2020; Li et al., 2020a). This data is obtained from the Iowa Environmental Mesonet NWS data archive (https://mesonet.agron.iastate.edu/nws/). The benchmark flood map is simulated by CREST-iMAP using MRMS 2-min QPE as the forcing data. To create the benchmark flood map, the CREST-iMAP goes through a warmup period from 01/04/2017 to 24/08/2017 using the MRMS 1-hour gauge corrected QPE and then simulates the Hurricane Harvey induced flood from 24/08/2017 to 02/09/2017 (10 days), driven by the MRMS 2-min QPE. The CREST-iMAP is calibrated against the USGS gauge 08069000 (Cypress Creek) and 08068500 (Spring Creek) from 01/04/2017 to 24/08/2017 to optimize the water balance parameters. The 10-meter resolution Manning's roughness coefficient field is derived from the landcover data from Multi-Resolution Land Characteristics Consortium (MRLC, https://www.mrlc.gov/), using a look-up table from the literature (Liu et al., 2019; McCuen, 2005).

Other benchmark data include the two USGS stream gauge data downstream of Spring (08068500) and Cypress Creek (0809000), which are obtained from the USGS Water Information System (USGS WIS, https://waterdata.usgs.gov/). The USGS Water-High-Marks (WHM) data are manually surveyed post-event water depths by measuring the residual water stain or mudlines left on buildings and other infrastructure, which is considered as the true ground observations of flood depth (Feaster and Koenig, 2017). The National Flood Insurance Program was passed by the US Congress in 1968 and placed under Federal Emergency Management Agency (FEMA) since 1979 (Bedient et al., 1988). The FEMA flood claim dataset

for Hurricane Harvey includes the coordinates of the claims that experienced water damage, as well as water-related electric damage, rain damage, and property cracks caused by water. The USGS WHM and FEMA flood insurance claim datasets are obtained from the HydroShare Hurricane Harvey collection (Arctur et al., 2018, https://www.hydroshare.org/resource/2836494ee75e43a9bfb647b37260e461/).

4.2.4 The CREST-iMAP model

CREST-iMAP is a hydrological and hydraulic coupled model developed by the Hydrometeorology and Remote Sensing Laboratory at the University of Oklahoma, and is designed to operate in near-real-time. It is the newly added member of a well-documented CREST modeling family, which has been widely applied to multiple operations and applications over the world (Clark et al., 2017; Flamig et al., 2020; Gourley et al., 2017; Wang et al., 2011). The CREST-iMAP coupled the water balance component of CREST and the 2D hydraulic routing using a fully solved Shallow Water Equation with a finite volume method. The model can be nested with the CREST/EF5 framework for near-real-time operation or as a standalone operation. The model receives precipitation data as the driving input over topographical datasets and the CREST water balance component generates excessive rainfall, which is then routed through the unstructured triangular mesh to generate channel flow rate, flood extent, and flood depth. The model calibration is conducted for the period from 2017/04/01 to 2017/08/24, using the DREAM algorithm (Vrugt, 2016; Vrugt et al., 2009) to optimize all water-balance parameters by targeting five different USGS stream gauges at midstream (08068275) and downstream (08068500) of Spring Creek, as well as upperstream (08068720), midstream (08068800), and downstream (08069000) of Cypress Creek.

4.2.5 Study area

Harris County, TX was the most impacted area during Hurricane Harvey with 103 casualties, over 40,000 people evacuated, and over 30,000 water rescues conducted *(Murphy, 2018)*. Figure 1 shows the impacted area of Hurricane Harvey, the hurricane track, and the study area of this research. The statistical analysis of different QPFs is conducted throughout the entire Harris County area, and the flood analysis is conducted at the Spring basin, which is located at the northern part of Harris County.

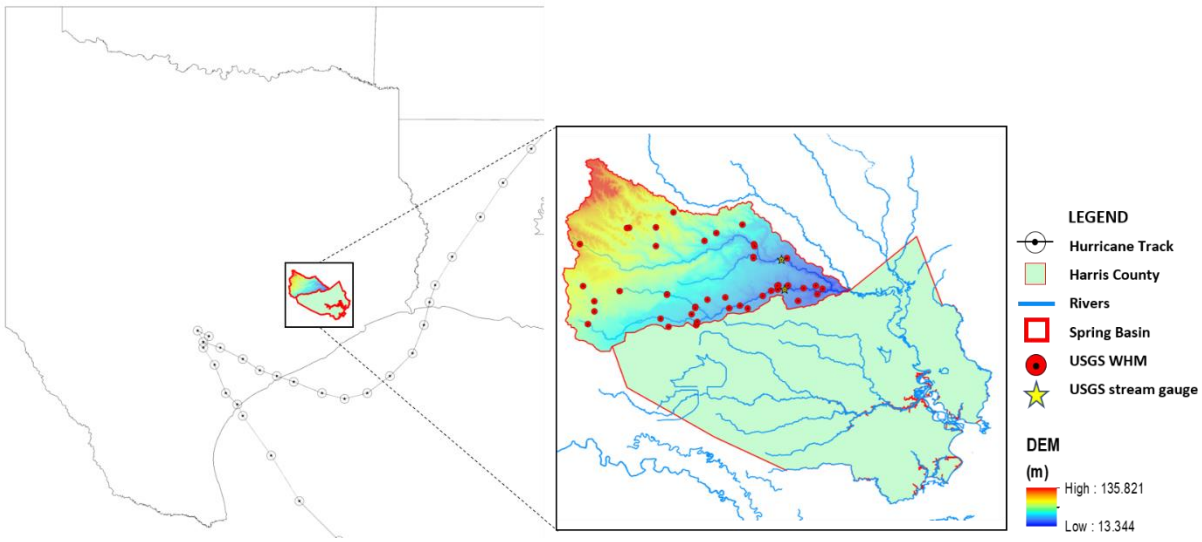


Figure 4.1. The Harris County, TX and Spring Basin, showing the track of Hurricane Harvey, WHM locations, USGS gauge locations, and the DEM of Spring Basin

4.2.6 Statistic metrics

To compare each forecasting product and evaluate the flood predictability, three levels of statistical tests are conducted in this study. First, the benchmark precipitation in 2.3 is used to compare QPFs and AI-based nowcasts using statistical measures. Second, the CREST-iMAP simulated hydrographs are analyzed using not only standard statistic metrics but also hydrological modeling efficiency measures. Third, the CREST-iMAP simulated flood extents are

analyzed using standard binary pattern measures. All three levels of statistical metrics are listed in Table 4.3.

Table 4.3. List of statistical metrics used in this study

Statistic metrics	Equation[a]	Value range	Perfect value
Correlation coefficient (CC)	$CC = \frac{\sum_{n=1}^{N}(f_n - \bar{f})(r_n - \bar{r})}{\sqrt{\sum_{n=1}^{N}(f_n - \bar{f})^2}\sqrt{\sum_{n=1}^{N}(r_n - \bar{r})^2}}$	-1, 1	1
Relative bias (RB)	$RB = \frac{1}{N}\sum_{n=1}^{N}\frac{f_n - r_n}{r_n}$	-∞, +∞	0
Root-mean-square error (RMSE)	$RMSE = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(f_n - r_n)^2}$	0, +∞	0
Probability of detection (POD)	$POD = \frac{F_1 \wedge R_1}{F_1 \wedge R_1 + F_0 \wedge R_1}$	0,1	1
False alarm ratio (FAR)	$FAR = \frac{F_1 \wedge R_0}{F_1 \wedge R_0 + F_1 \wedge R_0}$	0,1	0
Critical success index (CSI)	$CSI = \frac{F_1 \wedge R_1}{F_1 \wedge R_1 + F_0 \wedge R_1 + F_1 \wedge R_0}$	0,1	1
Nash-Sutcliffe coefficient efficiency (NSCE)	$NSCE = 1 - \frac{\sum_{n=1}^{N}(f_n - r_n)^2}{\sum_{n=1}^{N}(r_n - \bar{r})^2}$	-∞, 1	1
Peak flow error (PE)	$PE = f_{max} - r_{max}$	-∞, +∞	0
Peak time error (PTE)	$PTE = t(r_{max}) - t(f_{max})$	-∞, +∞	0

[a] Variables: n and N, sample index and a total number of samples, f represents the precipitation forecast products from the numerical modeling and AI nowcasts, r represents the reference including the MRMS QPE and USGS stream gauge observations, F and f represent the model simulation results of binary classification and values respectively; R and r represent the reference data of binary classification and values respectively; 1 and 0 means positive (wet) and negative (dry) classifications.

The correlation coefficient (CC) over a time series measures the strength of an estimate to capture the temporal pattern of the observation. The Relative bias (RB, fraction) measures the

errors of the estimate as a fraction of the observation value. The Root-mean-square error (RMSE, mm/hr for precipitation, and m^3/s for streamflow) measures the distance between the estimates and the observation. The Probability of detection (POD) measures the ability to predict the benchmark flood extent. The False Alarm Rate (FAR) reflects the tendency to overpredict the benchmark flood extent. The Critical Success Index (CSI) measures the overall performance of the flood predictions compared to the benchmark. The Nash-Sutcliffe coefficient of efficiency (NSCE, unitless) measures the effectiveness of the model prediction compared to the stream gauge observation. The peak flow error (PE, m^3/s) calculates the difference between predicted and observed peak flow. The peak time error (PTE, hour) calculates the arriving time difference between predicted and observed peak flow.

4.3 Results

4.3.1 Precipitation analysis

To observe the first level differences of all the precipitation forecast products, the accumulated precipitation from 25/08/2017 to 31/08/2017 is plotted in Figure 4.2.

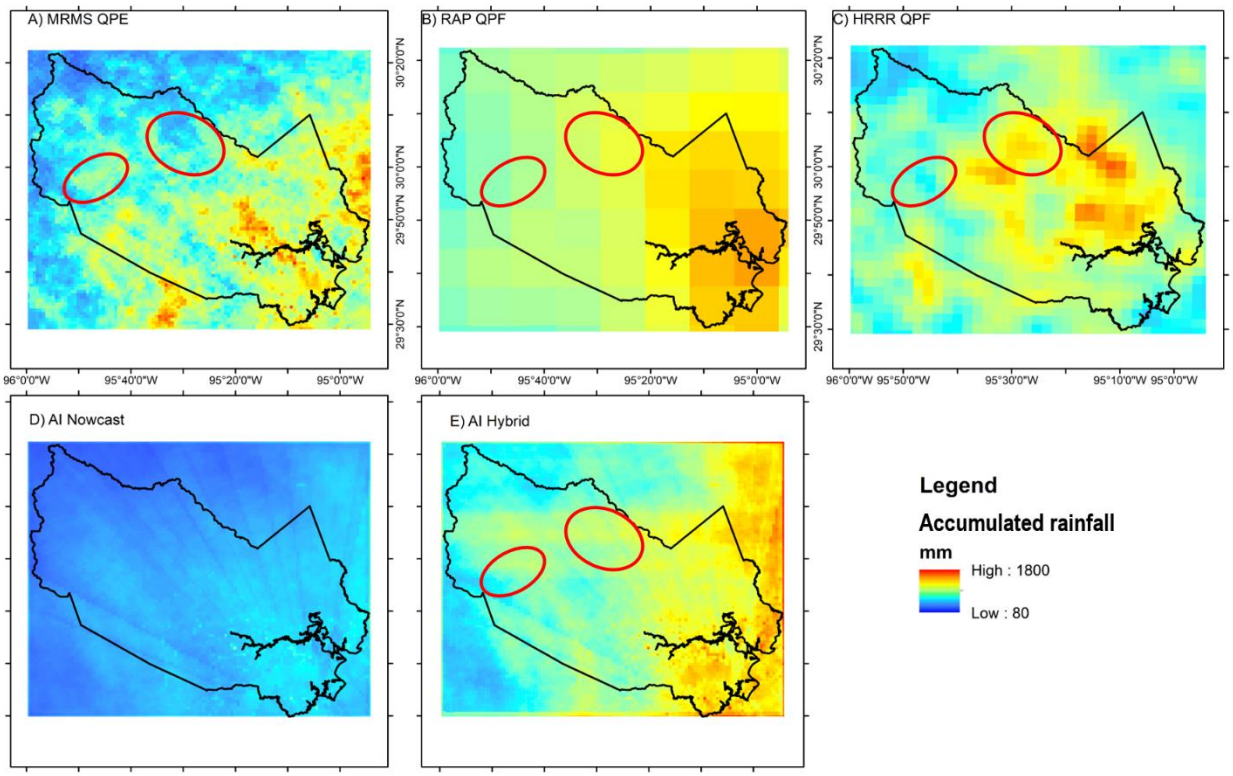


Figure 4.2. The accumulative precipitation during Hurricane Harvey of A) MRMS QPE, B) RAP QPF, C) HRRR QPF, D) AI nowcast, and E) AI hybrid, the red ovals circle the two examples of displacements

The first clear observation appears to be the spatial resolution differences, as RAP and HRRR show coarser resolution grid cells across the map compared to MRMS and the AI nowcasts. Due to the low spatial resolution, RAP can only provide a basic depiction of the precipitation core and has a smooth transition from high precipitation (southeast) to lower amounts (northwest). As listed in Table 4.1, RAP's highest mean precipitation rate (7.45 mm/hr), a mean accumulated precipitation (1,081 mm), but low maximum precipitation rate (49.83

mm/hr) means that it does not simulate extremes like MRMS. The second clear observation is the underprediction of the AI nowcasts. The average precipitation rate of AI nowcast during Hurricane Harvey is 3.18 mm/hr and its mean accumulated precipitation is only 190 mm, which is much less than the values of MRMS QPE, which are 6.16 mm/hr and 830 mm, respectively. Even though Google shows the U-Net framework providing well-performed nowcasting precipitation prediction from July 2017 to July 2019 (Agrawal et al., 2019), but for an extreme event like Hurricane Harvey, the result indicates there are difficulties for machine learning methods to predict the magnitude. Magnitude-wise, HRRR QPF is comparable with MRMS QPE with the mean precipitation rate of 6.65 and 6.16 mm/hr respectively (Table 4.1). However, the spatial displacement of large precipitation amounts from the HRRR is visually clear, where the red circles provide simple examples where MRMS has low accumulated precipitation but the HRRR has a large amount of total precipitation forecast and vice versa. The AI hybrid method shows improved precipitation magnitude from 3.18 mm/hr to 7.10 mm/hr and the mean accumulated precipitation increases from 190 to 930 mm. The CDF-matching method to hybrid machine learning and numerical modeling causes multiple problems: first, the high values of accumulated precipitation appeared to be concentrated at the east and northeast pixels at the edge; second, the horizontal, belt-shaped high-precipitation artifacts appear at the northern part of Harris County.

To quantify the difference between the benchmark precipitation and the precipitation forecasts, the first level statistics are applied to the county-wide averages as well as to each pixel from 25/08/2017 to 31/08/2017. The statistical results are listed in Table 4.4.

Table 4.4. The first level statistical results between precipitation forecasts and the benchmark precipitation estimation

Name	County averaged statistic			Pixel statistic		
	CC	RB	RMSE, mm/hr	CC	RB	RMSE, mm/hr
RAP QPF	0.76	255.46	5.41	0.31	129.60	15.72
HRRR QPF	0.55	364.42	7.31	0.13	144.75	20.06
AI nowcast	0.98	5.18	5.36	0.80	0.91	11.87
AI hybrid	0.62	446.81	6.51	0.66	75.40	13.16

Both county-averaged and pixel-level statistics agree that the AI nowcast precipitation forecast has the highest correlation coefficient (0.98 and 0.8), the lowest relative bias (5.18 and 0.91), and the least RMSE (5.18 mm/hr and 11.87 mm/hr). Among all precipitation forecasts, AI nowcast can capture the temporal pattern very well, despite that there is a large deficit in precipitation magnitude. HRRR QPFs show a very low correlation with benchmark precipitation both at county average (0.55) and pixel-level (0.13). The HRRR QPFs also have the largest RMSE among all other precipitation forecasts. RAP QPFs have a good correlation with the benchmark when averaged over Harris County (0.76) but poor correlation at each pixel (0.31), which might be caused by its coarse spatial resolution. The AI hybrid method is proved to be less correlated with the benchmark compared to AI nowcasts and has higher RMSE. The attempt to combine AI nowcast and HRRR provides worse statistical results, which can be caused by the artifacts found in Figure 4.2. To visualize the differences, the county-averaged precipitation estimates and forecasts are plotted in Figure 4.3.

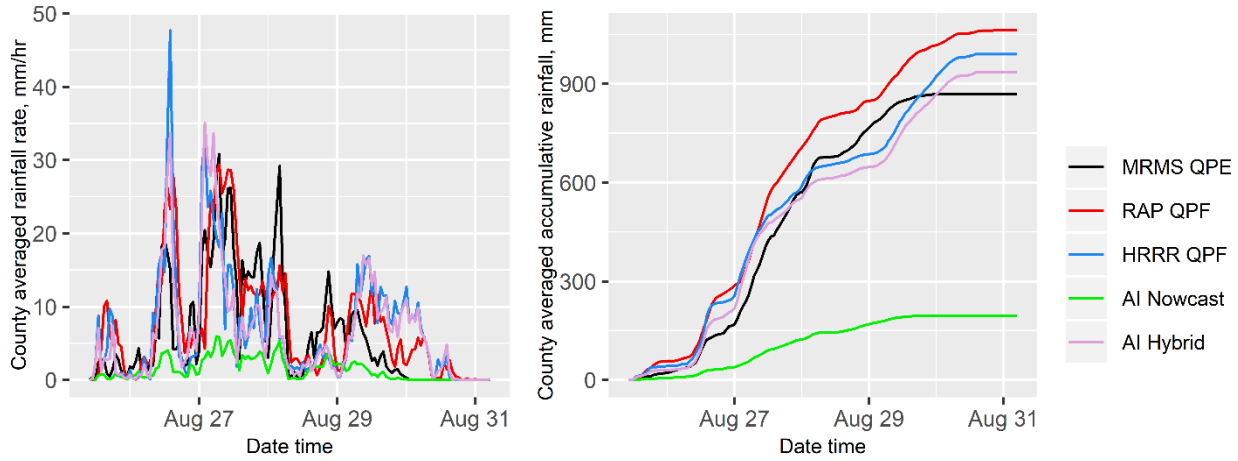


Figure 4.3. The county averaged precipitation rate (left) and the county averaged accumulative precipitation (right)

The AI nowcast (green) appears to follow all the peaks as the benchmark (black) and match well when the precipitation is light. However the magnitude of the peaks are only equivalent to a fraction of the benchmark data, and the accumulated rainfall amount indicates a severe underprediction. All other precipitation forecasts (RAP, HRRR, and AI hybrid) overpredict most of the precipitation peaks before 27/08/2017 and after 29/08/2017, while underpredicting during 28/08/2019, which explains the poor statistical results displayed in Table 4.3. As shown in Figure 3, the HRRR QPF (blue) and AI hybrid (plum) forecasts are overlapped in most days, which indicates that the CDF-matching method not only increases the magnitude of AI nowcasts but also picks up a lot of information from the HRRR. Since the HRRR has the spatial displacement problem, this property also passes along to machine learning causing the reduced predictability.

The pixel-level statistical results are plotted in Figure 4.4. As shown in the plot and Table 4.4, RAP and HRRR have a very low correlation with benchmark data throughout the area, and AI nowcasts and AI hybrid predictions have a higher correlation. Deep learning predictions appear to create minor artifacts in multiple radial circle patterns on the temporal correlation with

the benchmark data. For Relative Bias (RB), the numerical modeling forecasts (RAP and HRRR) have most of the high RB concentrated at the southeastern corner, which is towards the storm core. This indicates overprediction of precipitation at the highest intensity cores for the numerical modeling forecasts. The artifact of the AI hybrid found in Figure 4.2 is magnified in the bias analysis, where the horizontal stripes of high RB values concentrate at the northern half of the map. In this study, RAP has lower RMSE than HRRR, despite the fact that the HRRR is a newer and higher-resolution version of RAP, and the large RMSE values are concentrated towards the storm core.

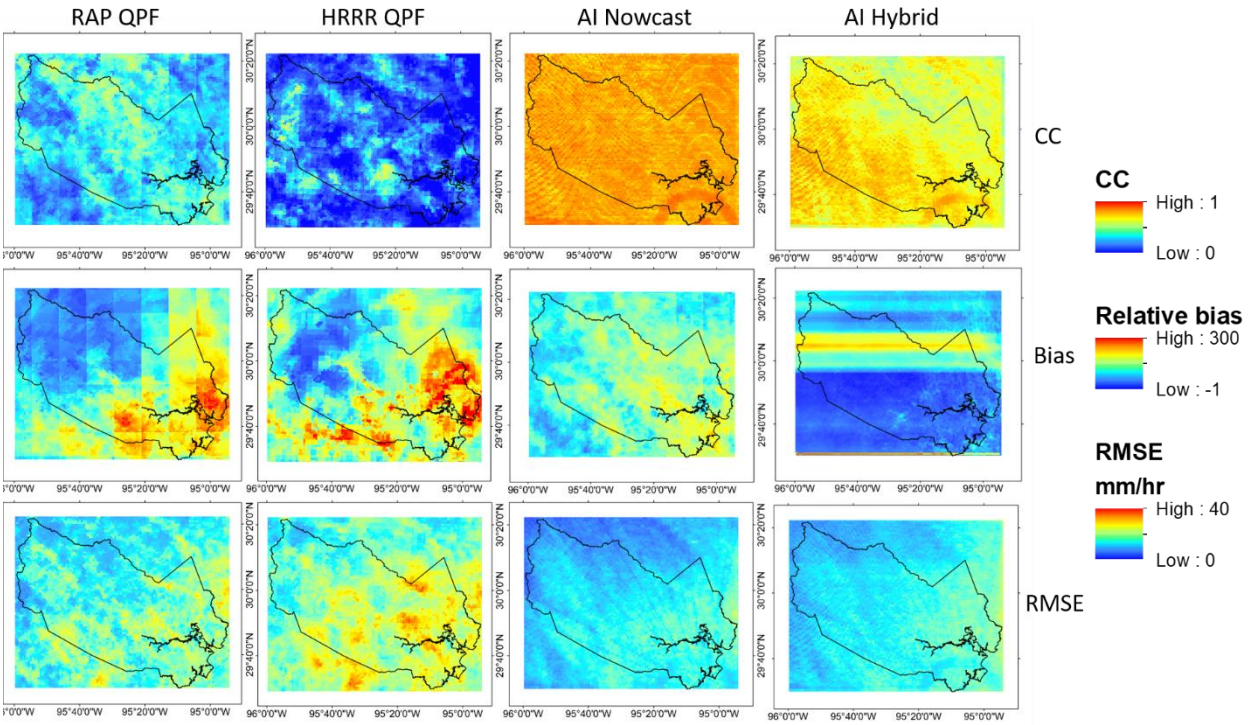


Figure 4.4. The statistic results (CC, RB, and RMSE) at each pixel for RAP QPF, HRRR QPF, AI nowcast, and AI hybrid, compared to the benchmark precipitation (MRMS QPE)

Based on the statistical analysis, the results indicate that there are significant differences between the precipitation forecasts and the benchmark, which should be carried through the CREST-iMAP simulations. The AI nowcasting method statistically matches the benchmark

except the magnitude, and the AI hybrid matches the magnitude but also carries the undesirable differences from the HRRR.

4.3.2 Hydrological analysis

The CREST-iMAP yields simulated hydrographs using forcing from the MRMS benchmark and all the precipitation forecasts from 15/08/2017 to 03/09/2017 at two USGS gauge locations at Cypress Creek and Spring Creek. Spring Creek is located at the northern part of the Spring basin, which includes an underdeveloped area and did not show too much overbank flow during Hurricane Harvey; while Cypress Creek is located at the southern part of the basin, across multiple developed urban areas, and had obvious overbank flow during Hurricane Harvey (Chen et al., 2020). The simulation results are plotted and listed in in Figure 4.5 and Table 4.5.

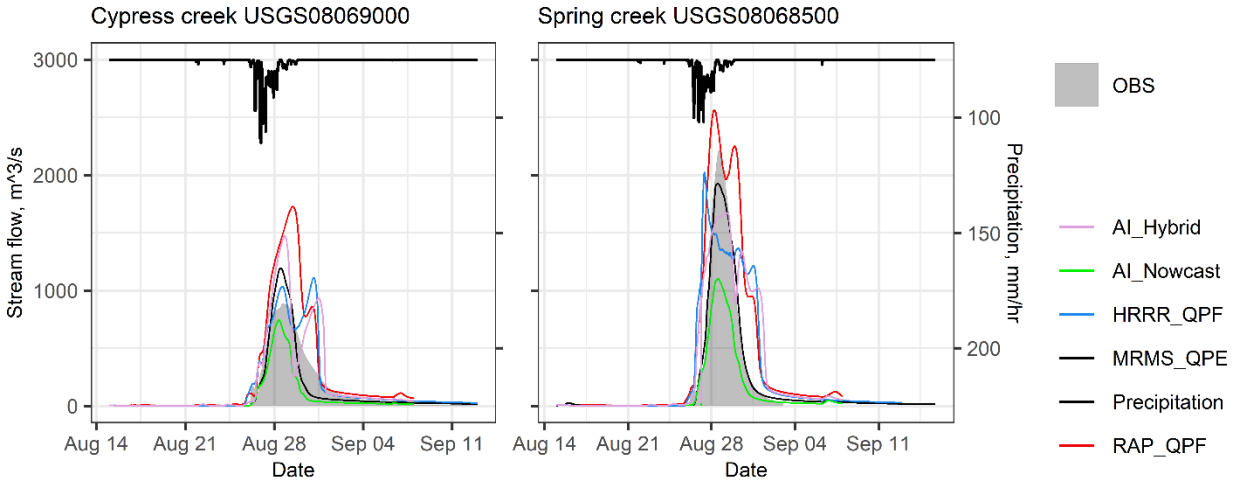


Figure 4.5. The flowrate simulation during Hurricane Harvey at the Cypress Creek (southern stream) and Spring Creek (northern stream)

In Figure 4.5, with the exception of the AI nowcasts, all precipitation forecasts generate a second flood peak at both Cypress and Spring creek, which was not observed by USGS stream gauges. As AI nowcasts underpredict the precipitation amounts and rates, they generate a smaller flood peak consequentially, but it has a NSCE (0.71) and CC (>0.9) with the USGS gauge data as shown in Table 4. The AI hybrid provides an undesirable flood forecast with NSCE less than

0.7 and positive PTE, which means the forecasted flood peak is later than the observed flood peak. The AI hybrid (plum line in Figure 4.5) carries some undesirable features from the HRRR, which causes a non-existent second flood peak. The NWP modeled QPFs produce less accurate flood prediction comparing to the deep-learning methods, in general. Particularly, RAP produces a negative NSCE at the Cypress Creek due to large overprediction. Overall, the precipitation forecasts cannot match the performance of the benchmark, which has the NSCE close to 0.9, CC close to 0.95, RMSE less than 100 m^3/s, and small but negative PTE. The benchmark precipitation data (MRMS) is utilized in the FLASH project as the forcing data for the CREST-EF5 model to produce the near-real-time streamflow (Flamig et al., 2020; Gourley et al., 2017). The FLASH project has shown its benefit as a supporting tool during Hurricane Harvey according to the NWS service assessment report (Murphy, 2018), therefore, with the consideration of the error propagated from the precipitation forecasts, the 1-hour lead-time flood prediction can be a very useful tool. Especially, the AI nowcast can prediction the flood timing accurately, so if the first responders know exactly when the flood is coming with hours of lead-time, it is a successful step for flood prediction.

As shown in the hydrological analysis, the current structure of using precipitation forecasts to drive CREST-iMAP has considerable value for the automated real-time flood prediction operation. Even though, no forecast data can provide a comparable flood prediction as MRMS benchmark observation, but the errors from the precipitation forecasts are expected and the prediction results do provide flood information with much longer lead-time. There is additional value from the AI nowcast method since it provides acceptable statistics despite the significant underprediction on magnitude.

Table 4.56. The hydrological analysis of precipitation forecasts and benchmark compared to the USGS stream gauge data

	Cypress Creek 08069000				
	MRMS	RAP	HRRR	AI Nowcast	AI Hybrid
NSCE	0.85	-0.05	0.66	0.71	0.44
RB (%)	71.96	219.36	141.16	3.51	84.75
CC	0.94	0.97	0.91	0.92	0.90
RMSE (m^3/s)	90	255	135	135	187
Peak Error (m^3/s)	301	834	219	-149	579
Peak Time Error (hour)	-6	16.75	56.5	-9.25	2
	Spring Creek 08068500				
NSCE	0.98	0.50	0.48	0.71	0.67
RB (%)	146.76	413.92	117.23	80.62	192.55
CC	0.99	0.92	0.77	0.99	0.85
RMSE (m^3/s)	59	588	627	441	367
Peak Error (m^3/s)	-298	340	-198	-1118	-544
Peak Time Error (hour)	-3.25	-11.25	-30.5	-2.75	11.5

4.3.3 Flood extent analysis

The CREST-iMAP can not only simulate the streamflow but also the 2D flood extent and flood depth. We take the maximum flood depth at every pixel through all the time steps to construct the maximum flood extent map for all precipitation forecasts and benchmark-forced simulations, which is shown in Figure 6. The binary statistical results are listed in in Table 4.6.

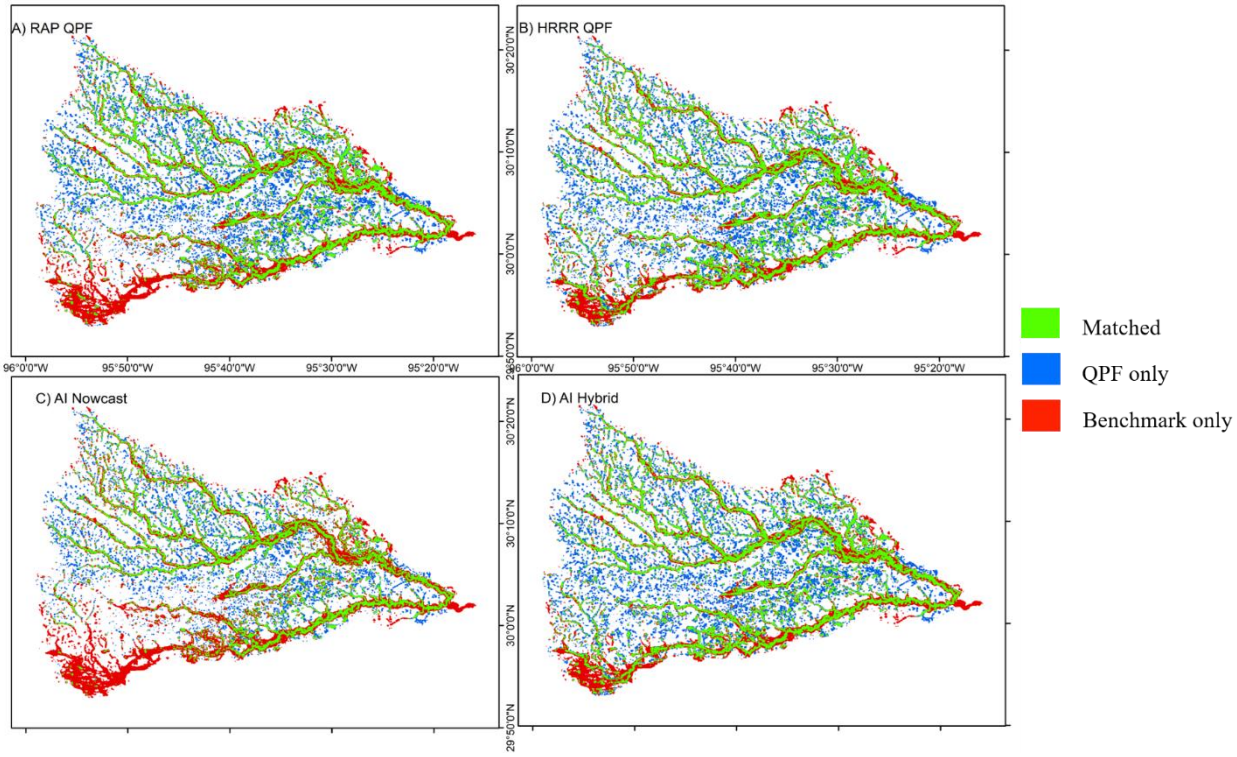


Figure 4.6. The 2D flood extents maps display the intersection of A) RAP QPF, B) HRRR QPF, C) AI nowcast, D) AI hybrid predictions with those from the benchmark flood map

Table 4.6. The flood extent binary statistics and FEMA flood claim coverage results

	POD	FAR	CSI	Covered FEMA claims	Percentage, %
RAP QPF	0.57	0.46	0.38	3157	30.18%
HRRR QPF	0.62	0.47	0.40	3327	31.81%
AI Nowcast	0.48	0.43	0.35	2581	24.68%
AI Hybrid	0.63	0.47	0.40	3292	31.48%
MRMS	1.00	0.00	1.00	9085	86.86%

Although previous statistical and hydrological analysis shows drastic differences between the different precipitation forecasts, the maximum flood extent maps are very similar, and the binary statistics in Table 5 are very comparable, too. The AI nowcast is the only one that has a lower POD (0.48) and a lower CSI (0.35), which is due to its underestimation of precipitation amount

as shown in Table 1 and Figure 3. The CREST-iMAP can neutralize most of the differences between different precipitation forecasts to produce similar maximum flood maps since the dominant factor for the maximum flood extent is the total precipitation during the event. However, there are no successful predictions for the upper Cypress Creek (southwestern part of the map). All flood predictions other than the AI nowcasts can achieve a CSI of about 0.4 and POD about 0.6, which can still provide limited information and guidelines for future flooding. However, the over 45% false alarming could mislead preparatory action to the flood.

There are 10,459 flood insurance claims within the Spring Basin, which the data included water damage, flood damage, water-related electrical damage, and property rupture or cracks due to water. Although not all of those claim locations were likely caused by being immersed in the accumulated stormwater, this FEMA flood insurance claims data can be a good indicator telling where the flood is located. By finding the number of claim cases that land within the maximum flood extent, the results simulated by different precipitation forecasts show around 30% successful prediction. Comparing to the benchmark flood extent (MRMS), the claim coverage reaches about 87%, so the flood extent predictions are still underperforming during Hurricane Harvey.

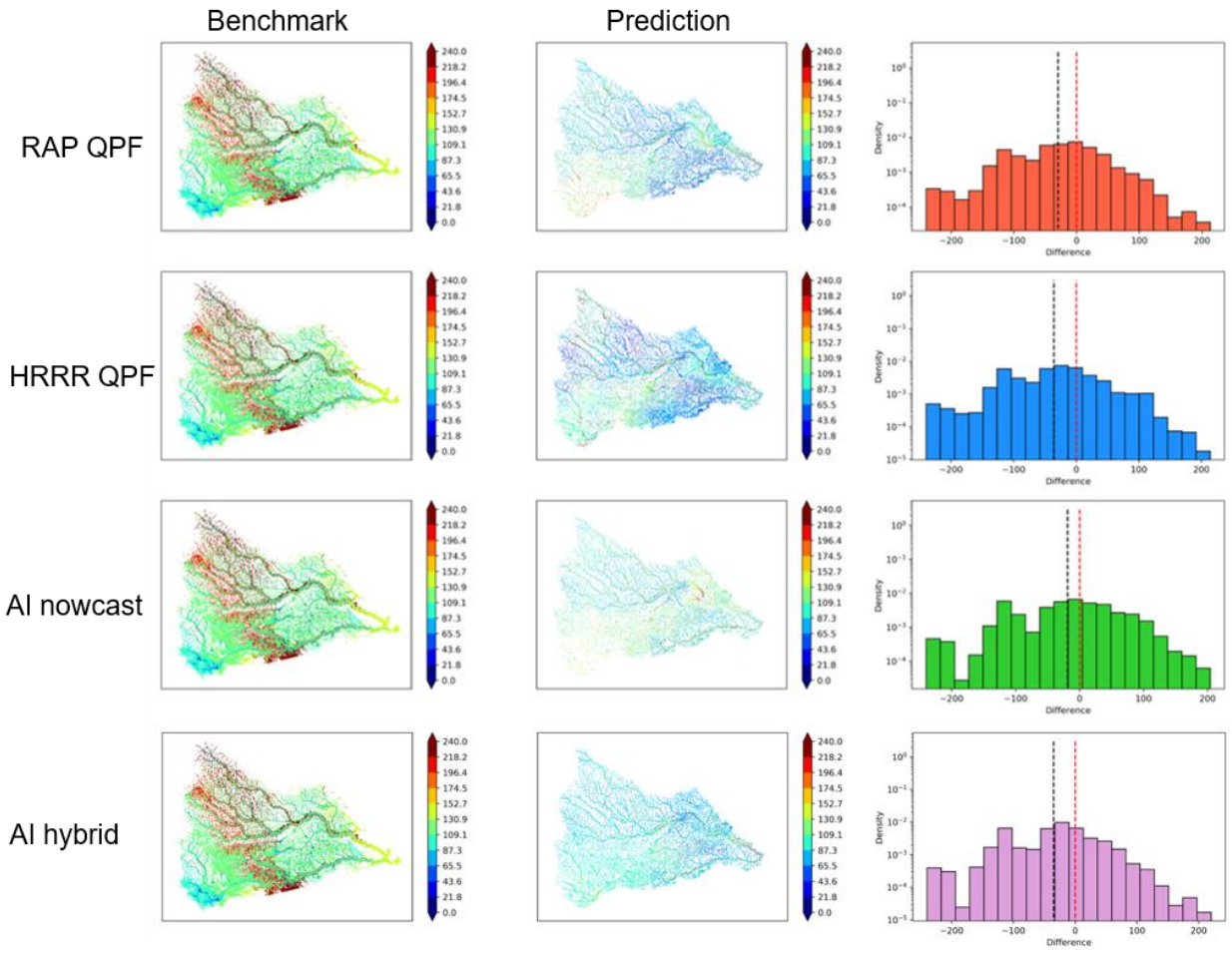


Figure 4.7. The half-hourly flood inundation time maps between the benchmark flood map (left) and the predictions (middle), as well as the time differences distribution (right)

The analysis of the initial inundation (water depth > 5 cm) time is listed in Figure 4.7, where the color ramp represents the unified time steps, which is equivalent to 30 mins. The dark blue color represents the first day of the study period (25/08/2017 00:00) and the dark red color represents the end of the last day of the study period (29/08/2017 24:00). For example, in the figure, all maps show dark blue color as the center of the main streamlines, which represent the river channels that have water running at the beginning of the period. The bar plot on the right column of Figure 7 represents the distribution of the flood time difference between the flood prediction and the benchmark flood extent, where the red dashed line is the 0-difference line and

the black dashed line represents the mean differences. Negative values mean the predicted flood comes earlier than the benchmark and the positive value means the predicted flood comes later than the benchmark.

For the benchmark flood map, most of the inundation occurs around 88 ~ 150 time-steps, which are late in the 26th to the morning of the 28th of August 2017. This result corresponds with the county-average precipitation results in Figure 3, where the benchmark (black line) shows the first peak of precipitation about 20 mm/hr around 1200 UTC on the 26th and two large rainfall peaks (30 mm/hr) on the 27th and early on the 28th of August. The flow dynamic analysis also picks up the overbank flow of the rivers where the color ramp changes from the center outward, from dark blue (flooded on 25th) to light green (flooded on 27th) then to dark red (flooded on 29th), which represents the river experiencing multiple overbank flows throughout the study period. The flood prediction from the RAP and HRRR QPF show much more blue (earlier inundation) compared to the benchmark, which corresponds to the time when the RAP (red) and HRRR (blue) have their largest peak precipitation rates (46 mm/hr and 34 mm/hr respectively) on the 26th of August (Figure 4.3). These extreme precipitation rates can cause flood inundation immediately, which leads to an earlier inundation than the benchmark. Since the HRRR has higher spatial resolution and greater peak precipitation rates than the RAP presented in Table 1, the initial inundation time map of the HRRR has more late (red) inundation along the river channels than RAP predicted inundation time, which indicates that there is more overbank flood with reasonable time lags represented in the HRRR than thje RAP. The AI nowcast-predicted inundation time map shows more light blue (32 hr) and green (48 hr) compared to other predictions, which is caused by the low precipitation intensity that caused the delayed water accumulation on the surface. The AI hybrid prediction shows earlier inundation than the

benchmark, which is very similar to the HRRR prediction. Overall, AI hybrid and HRRR predictions have over 70% of wet pixels (74% and 72% respectively) that are flooded earlier than the benchmark, and RAP prediction has 66% of wet pixels flooded earlier than the benchmark. Even with the significant underprediction of precipitation rate from AI nowcast, its flood prediction has about 50% of wet pixels flooded earlier than the benchmark.

The predicted flood extent provides slightly more useful information than the hydrograph predictions. Since the maximum flood extent reflects the total precipitation, the flood map predicted by the HRRR and RAP QPFs can provide information about the location of possible fluvial and pluvial floods, however, the flood prediction times are not reliable. Machine learning prediction methods show no advantage in the flood extent analysis, as much of the differences of precipitation forecasts are smoothed by the CREST-iMAP model.

4.3.4 Flood depth analysis

50 WHM sites are located in the basin, where the water depth values are extracted from all flood maps including the benchmark and predicted flood maps. In Figure 8, the scatter plot of the benchmark flood depth and predicted flood depth against the WHM data is presented. As the previous study indicates the error of simulated Harvey flood depth is up to 1 meter (Wing et al., 2019), it is no surprise that the benchmark (black dots) cannot align perfectly along the diagonal, and the predicted flood depths only perform worse. As shown in Figure 8, almost all predicted flood depths are below the diagonal except one location that is predicted by HRRR QPF with CREST-iMAP (blue dots). Since AI nowcast underpredicts the precipitation and associated flood, there are 38 out of 50 WHM sites that are predicted to be not inundated under this scenario. Other above-the-diagonal dots are the ones that WHM data shows 0 and the flood

predictions indicate positive values. Overall, all the flood predictions underpredict the inundation depth and provide little value for flooding emergency responses.

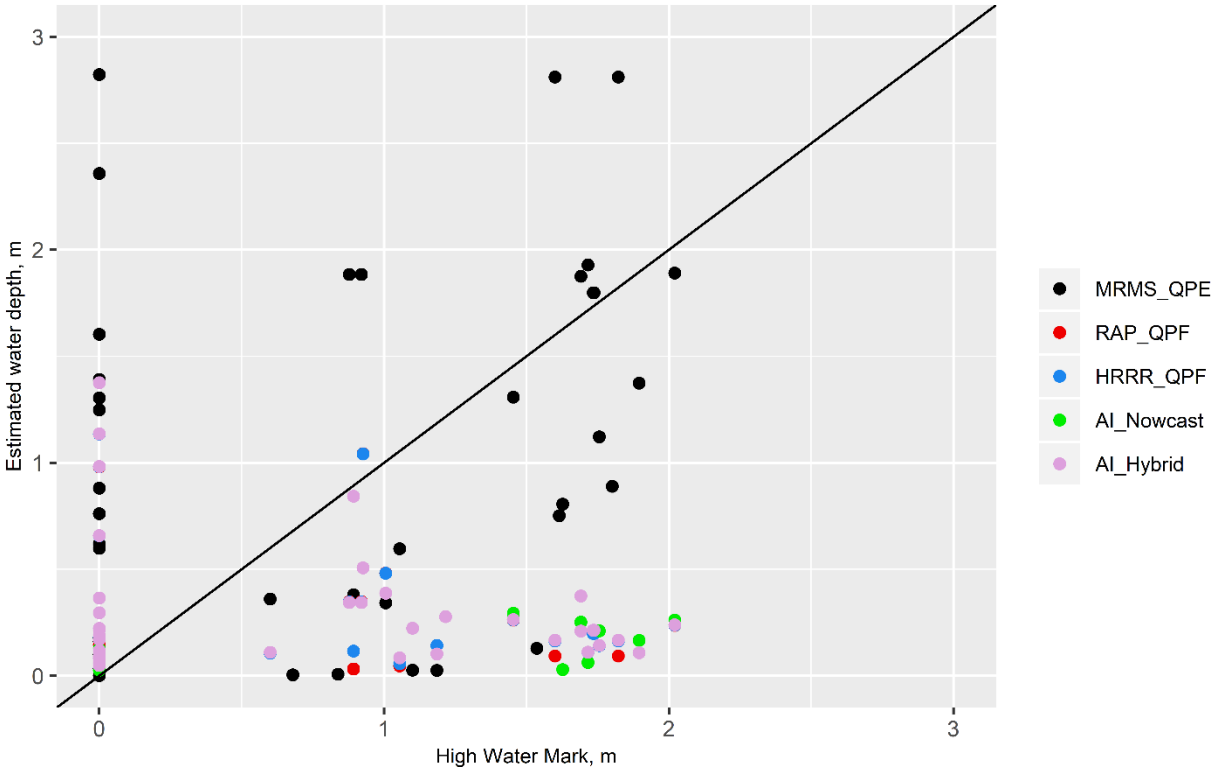


Figure 4.8. the scatter plot of flood inundation depths of benchmark flood map and predicted flood map against USGS WHM data

The error distributions of inundation depths are plotted in Figure 9. The plots show that no flood prediction error is as centered to 0 as the benchmark estimation (MRMS). Second to the benchmark, the RAP QPF can lead to a predicted flood depth that is slightly better than HRRR QPFs and the AI hybrid method-predicted flood depths. This result, again, demonstrates the similarity of HRRR QPF and AI hybrid forecasting, where their flood depth prediction errors are almost identical. It raises the speculation that using CDF-matching technique to hybrid machine learning and numerical modeling might pick up too much unnecessary information from the target dataset (HRRR in this study).

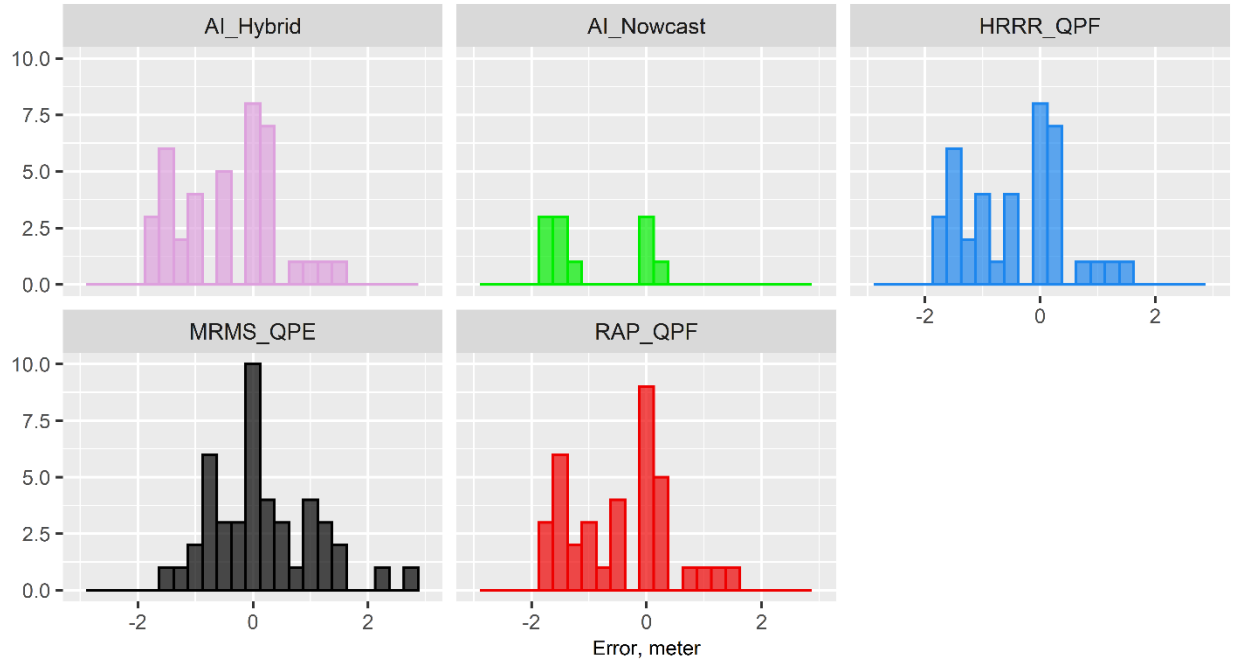


Figure 4.9. the flood depth error distribution between USGS WHMs and benchmark flood map (MRMS QPE) and predicted flood maps of RAP QPF, HRRR QPF, AI nowcast, and AI hybrid

In general, using RAP QPF, HRRR QPF, AI nowcast, and AI hybrid forecast can only provide limited information about the upcoming flood inundation depths for an extreme event. However, in the results, the AI nowcast precipitation forecast shows acceptable statistics compared with MRMS QPE except the rain-rate magnitude, so this method has the potential to yield good predictions if the extreme precipitation intensities could be properly forecasted. Even though the AI hybrid method in this study fails, other exploration can be done to improve the AI nowcasting products for extreme events.

4.4 Discussion

The NWP-based QPFs show poor predictability in precipitation spatial and temporal distribution, even though their predicted total quantity is relatively accurate. One possible cause could be the selection of the forecasted rain fields, which we select the +1-hour lead time forecasts while RAP and HRRR can forecast up to +18 hours. The study of Seo et al. (2018) suggests that the RAP and HRRR QPFs perform better with +4 ~ +6 hour lead time, indicating that the +1-hour forecasts are not necessarily the most reliable QPF product. However, due to the limited computational resources, the requirement of hyperresolution precipitation for hydrological forecast, and the heavy computational requirements of AI machine learning, the AI methods will not likely forecast more than 3 hours of lead time (only limited to 1-hour lead time in this study). Therefore, the underperformance of numerical modeling forecasts can be excused, and the predictability study over different lead-times will be conducted in the future. Unfortunately, the well-known QPF issues still exist, such as the underforecasting the convective system rainfall (RAP) and spatial displacement errors (HRRR), which are likely to be present with the other forecast periods.

The AI nowcast method lacks the ability to forecast the extreme precipitation event, such as Hurricane Harvey. The AI model is trained by all precipitation events in the Houston area from 2015 to 2019 that are listed in NOAA storm report and caused urban flooding, except Hurricane Harvey, which means it is trained specifically for high-intensity precipitation scenarios. But the results indicate the method cannot reproduce the precipitation intensity of Hurricane Harvey and provides no useful information for local flood prediction. Based on the study done by Google Research, the U-Net machine learning method outperforms HRRR QPF and the Optical Flow method using precision-recall curve analysis (Agrawal et al., 2019). This statement still holds in

this case study, however, as a binary analysis, the lack of predicted precipitation intensity is not revealed in the precision-recall curve analysis. Although the precision-recall curve is a common and powerful scoring tool for machine learning, this test is not sufficient for hydrological and hydrometeorological studies, since the precipitation intensity is as important, if not more, as its spatial-temporal placement. Therefore, deep learning methods for hydrological prediction still require much effort from the scientific community to improve the technology.

As a preliminary attempt to combine the results of AI nowcasts and numerical model outputs, the CDF-matching technique is used in this study to produce the AI hybrid forecasts, but failed to perform well in this case study. It has been shown in the previous section that the technique creates artifacts and border effects, but it also shows that the technique alters the spatial-temporal distribution of AI nowcasts towards HRRR, which leads to the similarities between HRRR and AI hybrid in section 3.3 and 3.4. As the AI nowcast captures the spatial-temporal pattern of MRMS estimation very well, increasing the precipitation intensity but not changing its pattern could be a future direction for machine-learning nowcasting approaches.

Lastly, the maximum flood map analysis using CREST-iMAP neutralizes or smooths much of the differences between precipitation forecasts as the accumulated precipitation dictates the maximum flood extent. Since the RAP and HRRR QPFs have a reasonable forecast of total precipitation amount over the region, the flood extent analysis results for numerical modeling are not completely unacceptable. Therefore, even with certain missing inundation scenarios, the QPF+CREAT-iMAP could still potentially provide preliminary information about the final outcome of a future flood event.

4.5 Conclusion

The ability of precipitation forecasts to predict flood discharges, inundation extent, and flood depth has been tested using numerically modeled QPFs (RAP and HRRR) and deep learning nowcasts (AI nowcast and AI hybrid) with +1-hour ahead lead time. None of the precipitation forecasts can provide comparable flood information as the radar-based benchmark (MRMS). However, as the total precipitation amount dictates the maximum flood extend, numerical weather prediction model QPFs can provide general information about inundation outcomes, where the HRRR slightly outperforms the RAP. The AI nowcast is incapable to capture the precipitation intensity of Hurricane Harvey, which indicates the potential inability of the method, as well as the inability of common machine learning performance tests to reveal such information.

The AI nowcast method can forecast the spatial-temporal pattern and extent of the precipitation; it will be the future objective to combine the AI nowcasts and QPFs using different methods, as the CDF-matching technique fails to improve the performance of the hybrid approach evaluated in this study. Since the NWP spatial displacement errors underly most of the QPF total error, taking advantage of spatial accuracy of AI nowcasts and the intensity accuracy of numerical modeling can be a promising research path. Further, flood predictability analysis with longer lead-time is needed to study the possible best combination of QPF+CREST-iMAP to provide the best possible flood forecast for local emergency response agencies.

Reference

Agrawal, S., Barrington, L., Bromberg, C., Burge, J., Gazen, C., Hickey, J., 2019. Machine Learning for Precipitation Nowcasting from Radar Images. arXiv:1912.12132 [cs, stat].

Arctur, D., Boghichi, E., Tarboton, D., Maidment, D., Bales, J., Idaszak, R., Seul, M., Castronova, A.M., 2018. Hurricane Harvey 2017 Collection. https://doi.org/10.4211/hs.2836494ee75e43a9bfb647b37260e461

Ashley, S.T., Ashley, W.S., 2008. Flood Fatalities in the United States. Journal of Applied Meteorology and Climatology 47, 805–818. https://doi.org/10.1175/2007JAMC1611.1

Barredo, J.I., 2007. Major flood disasters in Europe: 1950–2005. Nat Hazards 42, 125–148. https://doi.org/10.1007/s11069-006-9065-2

Bates, P.D., De Roo, A.P.J., 2000. A simple raster-based model for flood inundation simulation. Journal of Hydrology 236, 54–77. https://doi.org/10.1016/S0022-1694(00)00278-X

Bates, P.D., Horritt, M.S., Fewtrell, T.J., 2010. A simple inertial formulation of the shallow water equations for efficient two-dimensional flood inundation modelling. Journal of Hydrology 387, 33–45. https://doi.org/10.1016/j.jhydrol.2010.03.027

Bedient, P.B., Huber, W.C., Vieux, B.E., 1988. Hydrology and floodplain analysis, 5th ed. ed. Prentice Hall, Upper Saddle River, NJ.

Benito, G., Lang, M., Mariano Barriendos, M.C.L., Frances, F., Ouarda, T., Thorndycraft, V.R., 2004. Use of systematic, palaeoflood and historical data for the improvement of flood risk estimation. Review of scientific methods." Natural Hazards 31, 623–643.

Benjamin, S.G., Devenyi, D., Weygandt, S.S., Brundage, K.J., Brown, J.M., Grell, G.A., Kim, D., Schwartz, B.E., Smirnova, T.G., Smith, T.L., 2004. An hourly assimilation–forecast cycle: The RUC. Monthly Weather Review 132, 495–518. https://doi.org/10.1175/1520-0493(2004)132,0495:AHACTR.2.0.CO;2

Benjamin, S.G., Weygandt, S.S., Brown, J.M., Hu, M., Alexander, C.R., Smirnova, T.G., Olson, J.B., James, E.P., Dowell, D.C., Grell, G.A., Lin, H., Peckham, S.E., Smith, T.L., Moninger, W.R., Kenyon, J.S., Manikin, G.S., 2016. A North American Hourly Assimilation and Model Forecast Cycle: The Rapid Refresh. Monthly Weather Review 144, 1669–1694. https://doi.org/10.1175/MWR-D-15-0242.1

Brauer, N.S., Basara, J.B., Homeyer, C.R., McFarquhar, G.M., Kirstetter, P.E., 2020. Quantifying Precipitation Efficiency and Drivers of Excessive Precipitation in Post-Landfall Hurricane Harvey. Journal of Hydrometeorology 21, 433–452. https://doi.org/10.1175/JHM-D-19-0192.1

Chen, M., Nabih, S., Brauer, N.S., Gao, S., Gourley, J.J., Hong, Z., Kolar, R.L., Hong, Y., 2020. Can Remote Sensing Technologies Capture the Extreme Precipitation Event and Its Cascading Hydrological Response? A Case Study of Hurricane Harvey Using EF5 Modeling Framework. Remote Sensing 12, 445. https://doi.org/10.3390/rs12030445

Clark, R.A., Flamig, Z.L., Vergara, H., Hong, Y., Gourley, J.J., Mandl, D.J., Frye, S., Handy, M., Patterson, M., 2017. Hydrological Modeling and Capacity Building in the Republic of Namibia. Bulletin of the American Meteorological Society 98, 1697–1715. https://doi.org/10.1175/BAMS-D-15-00130.1

Cohen, S., Praskievicz, S., Maidment, D.R., 2018. Featured Collection Introduction: National Water Model. J Am Water Resour Assoc 54, 767–769. https://doi.org/10.1111/1752-1688.12664

Cuo, L., Pagano, T.C., Wang, Q.J., 2011. A Review of Quantitative Precipitation Forecasts and Their Use in Short- to Medium-Range Streamflow Forecasting. Journal of Hydrometeorology 12, 713–728. https://doi.org/10.1175/2011JHM1347.1

de Almeida, G.A.M., Bates, P., 2013. Applicability of the local inertial approximation of the shallow water equations to flood modeling: APPLICABILITY LOCAL INERTIAL. Water Resour. Res. 49, 4833–4844. https://doi.org/10.1002/wrcr.20366

Ebert, E.E., McBride, J.L., 2000. Verification of precipitation in weather systems: determination of systematic errors. Journal of Hydrology 239, 179–202. https://doi.org/10.1016/S0022-1694(00)00343-7

Feaster, T.D., Koenig, T.A., 2017. Field manual for identifying and preserving High-Water Mark data (Open-File Report No. 67). U.S. Geological Survey, Reston, Virginia.

Flamig, Z.L., Vergara, H., Gourley, J.J., 2020. The Ensemble Framework For Flash Flood Forecasting (EF5) v1.2: Description and Case Study (preprint). Hydrology. https://doi.org/10.5194/gmd-2020-46

Franch, G., Maggio, V., Coviello, L., Pendesini, M., Jurman, G., Furlanello, C., 2020. TAASRAD19, a high-resolution weather radar reflectivity dataset for precipitation nowcasting. Sci Data 7, 234. https://doi.org/10.1038/s41597-020-0574-8

Georgakakos, K.P., Hudlow, M.D., 1984. Quantitative Precipitation Forecast Techniques for Use in Hydrologic Forecasting. Bulletin of the American Meteorological Society 65, 1186–1200. https://doi.org/10.1175/1520-0477(1984)065<1186:QPFTFU>2.0.CO;2

Gobiet, A., Kotlarski, S., Beniston, M., Heinrich, G., Rajczak, J., Stoffel, M., 2014. 21st century climate change in the European Alps—A review. Science of The Total Environment 493, 1138–1151. https://doi.org/10.1016/j.scitotenv.2013.07.050

Gochis, D.J., Dugger, A.L., Yu, W., Yates, D.N., Sampson, K., Barlage, M., Pan, L., Zhang, Y., McCreight, J.L., RafieeiNasab, A., Karsten, L., Read, L., Gaydos, A., McAllister, M., Mills, J., Towler, E., Grim, J., FitzGerald, K., 2017. The NOAA National Water Model: Research to Operations to Research (Presentation). CUAHSI, Boulder, CO.

Golding, B.W., 2000. Quantitative precipitation forecasting in the UK. Journal of Hydrology 239, 286–305. https://doi.org/10.1016/S0022-1694(00)00354-1

Gourley, J.J., Flamig, Z.L., Vergara, H., Kirstetter, P.-E., Clark, R.A., Argyle, E., Arthur, A., Martinaitis, S., Terti, G., Erlingis, J.M., Hong, Y., Howard, K.W., 2017. The FLASH Project: Improving the Tools for Flash Flood Monitoring and Prediction across the United States. Bulletin of the American Meteorological Society 98, 361–372. https://doi.org/10.1175/BAMS-D-15-00247.1

Hapuarachchi, H.A.P., Wang, Q.J., Pagano, T.C., 2011. A review of advances in flash flood forecasting. Hydrol. Process. 25, 2771–2784. https://doi.org/10.1002/hyp.8040

Kain, J.S., Xue, M., Coniglio, M.C., Weiss, S.J., Kong, F., Jensen, T.L., Brown, B.G., Gao, J., Brewster, K., Thomas, K.W., Wang, Y., Schwartz, C.S., Levit, J.J., 2010. Assessing Advances in the Assimilation of Radar Data and Other Mesoscale Observations within a Collaborative Forecasting–Research Environment. Weather and Forecasting 25, 1510–1521. https://doi.org/10.1175/2010WAF2222405.1

Kingma, D.P., Ba, J., 2017. Adam: A Method for Stochastic Optimization. arXiv:1412.6980 [cs].

Kumar, A., Islam, T., Sekimoto, Y., Mattmann, C., Wilson, B., 2020. Convcast: An embedded convolutional LSTM based architecture for precipitation nowcasting using satellite data. PLoS ONE 15, e0230114. https://doi.org/10.1371/journal.pone.0230114

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521, 436–444. https://doi.org/10.1038/nature14539

Lee, T.R., Buban, M., Turner, D.D., Meyers, T.P., Baker, C.B., 2019. Evaluation of the High-Resolution Rapid Refresh (HRRR) Model Using Near-Surface Meteorological and Flux Observations from Northern Alabama. Weather and Forecasting 34, 635–663. https://doi.org/10.1175/WAF-D-18-0184.1

Li, Z., Chen, M., Gao, S., Hong, Z., Tang, G., Wen, Y., Gourley, J.J., Hong, Y., 2020a. Cross-Examination of Similarity, Difference and Deficiency of Gauge, Radar and Satellite Precipitation Measuring Uncertainties for Extreme Events Using Conventional Metrics and Multiplicative Triple Collocation. Remote Sensing 12, 1258. https://doi.org/10.3390/rs12081258

Li, Z., Wen, Y., Schreier, M., Behrangi, A., Hong, Y., Lambrigtsen, B., 2020b. Advancing Satellite Precipitation Retrievals with Data Driven Approaches: Is black box model explainable? Earth and Space Science. https://doi.org/10.1029/2020EA001423

Ligda, M.G.H., 1953. Horizontal Motion of Small Precipitation Areas as Observed by Radar. (PhD Thesis). MASSACHUSETTS INSTITUTE OF TECHNOLOGY.

Liu, Z., Merwade, V., Jafarzadegan, K., 2019. Investigating the role of model structure and surface roughness in generating flood inundation extents using one- and two-dimensional hydraulic models. J Flood Risk Management 12, e12347. https://doi.org/10.1111/jfr3.12347

Loken, E.D., Clark, A.J., Xue, M., Kong, F., 2017. Comparison of Next-Day Probabilistic Severe Weather Forecasts from Coarse- and Fine-Resolution CAMs and a Convection-Allowing Ensemble. Weather and Forecasting 32, 1403–1421. https://doi.org/10.1175/WAF-D-16-0200.1

McCuen, R.H., 2005. Hydrologic analysis and design, 3rd ed. ed. Pearson Prentice Hall, Upper Saddle River, N.J.

Meehl, G.A., Zwiers, F., Evans, J., Knutson, T., Mearns, L., Whetton, P., 2000. Trends in Extreme Weather and Climate Events: Issues Related to Modeling Extremes in Projections of Future Climate Change. Bulletin of the American Meteorological Society 81, 427–436. https://doi.org/10.1175/1520-0477(2000)081<0427:TIEWAC>2.3.CO;2

Murphy, J.D., 2018. Service assessment August–September 2017 Hurricane Harvey (US DOC). NOAA National Weather Service, Silver Spring, Maryland.

Newell, J.E., Deaver, D.G., 1981. The LFM-II model--1980 (Technical memo No. 66). NOAA, Springfield, VA.

Sadeghi, M., Nguyen, P., Hsu, K., Sorooshian, S., 2020. Improving near real-time precipitation estimation using a U-Net convolutional neural network and geographical information. Environmental Modelling & Software 134, 104856. https://doi.org/10.1016/j.envsoft.2020.104856

Sampson, C.C., Smith, A.M., Bates, P.D., Neal, J.C., Alfieri, L., Freer, J.E., 2015. A high-resolution global flood hazard model: A HIGH-RESOLUTION GLOBAL FLOOD HAZARD MODEL. Water Resour. Res. 51, 7358–7381. https://doi.org/10.1002/2015WR016954

Seo, B.-C., Quintero, F., Krajewski, W.F., 2018. High-Resolution QPF Uncertainty and Its Implications for Flood Prediction: A Case Study for the Eastern Iowa Flood of 2016. Journal of Hydrometeorology 19, 1289–1304. https://doi.org/10.1175/JHM-D-18-0046.1

Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W., Woo, W., 2015. Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting. arXiv:1506.04214 [cs].

Smith, K., Ward, R., 1998. Floods: Physical processes and human impact. John Wiley, Chichester.

Sønderby, C.K., Espeholt, L., Heek, J., Dehghani, M., Oliver, A., Salimans, T., Agrawal, S., Hickey, J., Kalchbrenner, N., 2020. MetNet: A Neural Weather Model for Precipitation Forecasting. arXiv:2003.12140 [physics, stat].

Srinivasan, R., Arnold, J.G., 1994. INTEGRATION OF A BASIN-SCALE WATER QUALITY MODEL WITH GIS. J Am Water Resources Assoc 30, 453–462. https://doi.org/10.1111/j.1752-1688.1994.tb03304.x

Sun, J., Xue, M., Wilson, J.W., Zawadzki, I., Ballard, S.P., Onvlee-Hooimeyer, J., Joe, P., Barker, D.M., Li, P.-W., Golding, B., Xu, M., Pinto, J., 2014. Use of NWP for Nowcasting Convective Precipitation: Recent Progress and Challenges. Bulletin of the American Meteorological Society 95, 409–426. https://doi.org/10.1175/BAMS-D-11-00263.1

Trenberth, K.E. (Ed.), 2010. Climate system modeling. Cambridge University Press, Cambridge.

USGCRP, 2017. Climate Science Special Report: Fourth National Climate Assessment, Volume I. U.S. Global Change Research Program, Washington, DC, USA. https://doi.org/10.7930/J0J964J6

van Oldenborgh, G.J., van der Wiel, K., Sebastian, A., Singh, R., Arrighi, J., Otto, F., Haustein, K., Li, S., Vecchi, G., Cullen, H., 2018. Corrigendum: Attribution of extreme rainfall from Hurricane Harvey, August 2017 (2017 *Environ. Res. Lett.* **12** 124009). Environ. Res. Lett. 13, 019501. https://doi.org/10.1088/1748-9326/aaa343

Vrugt, J.A., 2016. Markov chain Monte Carlo simulation using the DREAM software package: Theory, concepts, and MATLAB implementation. Environmental Modelling & Software 75, 273–316. https://doi.org/10.1016/j.envsoft.2015.08.013

Vrugt, J.A., Braak, C.J.F., Gupta, H.V., Robinson, B.A., 2009. Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling? Stochastic Environmental Research and Risk Assessment 23, 1101–1026. https://doi.org/10.1007/s00477-008-0274-y.

Wang, J., Hong, Y., Li, L., Gourley, J.J., Khan, S.I., Yilmaz, K.K., Adler, R.F., Policelli, F.S., Habib, S., Irwn, D., Limaye, A.S., Korme, T., Okello, L., 2011. The coupled routing and excess storage (CREST) distributed hydrological model. Hydrological Sciences Journal 56, 84–98. https://doi.org/10.1080/02626667.2010.543087

Werner, M., Cranston, M., 2009. Understanding the Value of Radar Rainfall Nowcasts in Flood Forecasting and Warning in Flashy Catchments. Met. Apps 16, 41–55. https://doi.org/10.1002/met.125

Wing, O.E.J., Bates, P.D., Sampson, C.C., Smith, A.M., Johnson, K.A., Erickson, T.A., 2017. Validation of a 30 m resolution flood hazard model of the conterminous United States: 30 m RESOLUTION FLOOD MODEL OF CONUS. Water Resour. Res. 53, 7968–7986. https://doi.org/10.1002/2017WR020917

Wing, O.E.J., Sampson, C.C., Bates, P.D., Quinn, N., Smith, A.M., Neal, J.C., 2019. A flood inundation forecast of Hurricane Harvey using a continental-scale 2D hydrodynamic model. Journal of Hydrology X 4, 100039. https://doi.org/10.1016/j.hydroa.2019.100039

Wood, E.F., Lettenmaier, D.P., Zartarian, V.G., 1992. A land-surface hydrology parameterization with subgrid variability for general circulation models. J. Geophys. Res. 97, 2717. https://doi.org/10.1029/91JD01786

Xue, M., Kong, F., Thomas, K.W., Gao, J., Wang, Y., Brewster, K., Droegemeier, K.K., 2013. Prediction of Convective Storms at Convection-Resolving 1 km Resolution over Continental United States with Radar Data Assimilation: An Example Case of 26 May 2008 and Precipitation Forecasts from Spring 2009. Advances in Meteorology 2013, 1–9. https://doi.org/10.1155/2013/259052

Xue, X., Zhang, K., Hong, Y., Gourley, J.J., Kellogg, W., McPherson, R.A., Wan, Z., Austin, B.N., 2016. New multisite cascading calibration approach for hydrological models: case study in the Red River Basin using the VIC model. Journal of Hydrologic Engineering 21. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001282.

Chapter 5. Overall Conclusion

Climate change is anticipated to bring more frequent and severe flooding around the globe. Therefore, tools are necessary to provide accurate and reliable flood prediction with enough lead-time for responses, which is, and will be the constant request from the society to protect public properties and human lives. Therefore, consistent effort is needed from past decades and continuously to the future in real time operational flood monitoring and prediction. The recent improvement of remote-sensing technology not only improves the coverage of flood monitoring, but also increases the accuracy of flood simulation. Moreover, improvement of precipitation forecasting technologies has been significant in recent years, which has higher accuracy, higher spatial resolution, and temporal resolution that is possible for hydrological models to predict flooding with longer lead time. In this work, different remote sensing precipitation estimates and different precipitation forecasting products are evaluated to find the best practice for the potential real-time operational flood predictions for an extreme event. Meanwhile, to expand the efficacy of the flood prediction modeling, this work develops a comprehensive, hydrologic and hydraulic coupled flood prediction model that can provide multi-level information including but not limited to the streamflow rate, flood extent, overland flow speed and direction, and flood depth.

The hypothesis A is confirmed by the study in Chapter 2 regarding the advanced remote sensing technologies can indeed provide accurate precipitation observation to generate very accurate flood simulations using CREST-EF5 during the extreme event, which can even compensate the cases when the measuring instruments are damaged or overbank flood occurs. Chapter 2 also shows that the MRMS QPE can provide high accuracy on flood simulation, and the GPM IMERG provides acceptable flood simulation, too, while the uncalibrated products performed better than calibrated products for the unprecedented event. The CREST-EF5 model

is once-again proven to be an effective hydrological model that is capable for potential applications beyond CONUS.

Built on the knowledge from Chapter 2, the CREST model is coupled with ANUGA model to simulate the flood during Hurricane Harvey, and the simulation results are compared with three other operational flood-mapping systems in the world. The hypothesis B is confirmed by the study in Chapter 3, where the CREST-iMAP model is developed and tested. The results not only show the success of simulating the flood caused by Hurricane Harvey that is comparable to or better than three other prestigious real-time flood monitoring systems in the world, but also takes the advantage of the fully solved St. Venant equation, which allows it to bypass certain data requirements that are commonly needed for 2D flood simulation, such as channel geometry. Therefore, the CREST-iMAP can potentially be applied to many locations by utilizing the remote sensing data and existing CREST-EF5 parameters, regardless of the limitation of local data supply. The CREST-iMAP has the potential to be a nested system of CREST-EF5 for flood monitoring when the flood warning alerts, which can be implemented globally to provide comprehensive flood predictions and early warning.

The proposed hypothesis C is confirmed by the study described in Chapter 4. First, the CREST-iMAP model is proven to be able to neutralize a portion of the uncertainties from the precipitation forecasts. The NWP products with CREST-iMAP can predict about 60% of maximum flood extent (POD) of the benchmark with 1-hour lead-time, which can provide some information about the upcoming flood for the local emergency responders and planners. Moreover, the deep learning precipitation nowcast shows its advantage to eliminate any well-known displacement error from NWP models, which leads to a possible future study to combine the advantages of deep learning precipitation nowcasting and NWP forecast.

In the future, the CREST-iMAP simulation can be extended to the entire Hurricane Harvey impact area and utilize the economic analysis to further test the scaling-up capability. It can also be implemented as a nested model in the FLASH project, which is designed to be triggered when the flood warning is issued at concerning locations. It is also a promising research track in the future to improve the NWP and deep learning forecast, which can potentially produce more accurate flood prediction by utilize the forecast success of the accurate precipitation intensity from NWP model and the lack of displacement error from deep learning model.