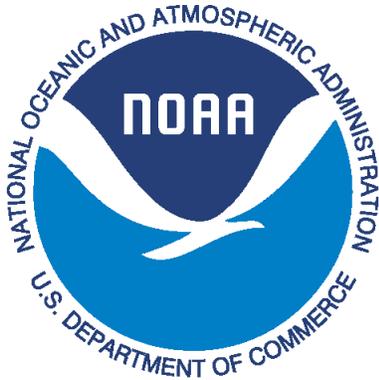


TECHNICAL REPORT 15
September 20, 2018



OWP | OFFICE OF
WATER
PREDICTION



CUAHSI
universities allied for water research

National Water Center
Innovators Program
Summer Institute Report 2018

National Water Center Innovators Program Summer Institute Report 2018

Editors:

Fernando Aristizabal

Lauren E. Grimley

Jerad Bales

Danielle Tijerina

Trey Flowers

Edward P. Clark

Prepared in cooperation with the Consortium of Universities for the Advancement of Hydrologic Science, Inc. and the National Water Center

CUAHSI Technical Report No. 15

DOI: <http://doi.org/10.4211/hs.c7ee150767114d3a927b67a94f0edc04>

Version 1.07

September 20th, 2018

Suggested Citation:

Aristizabal F, Grimley L E, et al. 2018. National Water Centers Innovators Program Summer Institute Report. Consortium of Universities for the Advancement of Hydrologic Science, Inc. Technical Report No 15.

Contents

Preface	4
Project Summary	7
Chapter 1: Groundwater-Surface Water Exchange and Streamflow Prediction using the National Water Model in the Northern High Plains Aquifer region, USA.....	9
Chapter 2: New Algorithms for Groundwater Discharge Estimation for National Water Model Streamflow Forecasts	25
Chapter 3: Effects of Spatial Resolution on a Distributed Hydrologic Model through Dynamical Forcings: Flood Extent and Depth in Low Gradient Watersheds	40
Chapter 4: Sensitivity of Urban Flooding to Presence of Subsurface Storm Drainage Systems in Hydrologic Models for Low-Gradient Watersheds.....	53
Chapter 5: Using Dimensionless Scaling Parameters as Decision Metrics in a Heterogeneous Hydraulic Routing Scheme	71
Chapter 6: An Approach for Incorporating Realistic Channel Geometry into Continental-Scale Hydrological Modeling	87
Chapter 7: Applications of Citizen Science Data for National Water Model Streamflow Forecasts.....	101
Appendix	136

Preface

The Consortium of Universities for the Advancement of Hydrologic Science, Inc. (CUAHSI) is a non-profit organization of more than 130 member institutions and provides programs and services which the advancement of interdisciplinary water science. The Office of Water Prediction of the NOAA National Weather Service (NWS) established the National Water Center (NWC) on the Tuscaloosa campus of the University of Alabama to serve as the U.S. center for water forecast services, in partnership with other federal agencies. The NWS, in partnership with CUAHSI, established in 2015 the National Water Center Innovators Program to engage the academic community in research to advance the mission of the National Water Center.

The key activity of the Innovators Program is a seven-week Summer Institute. The Summer Institute is a residential program at the National Water Center which brings graduate students and senior academic faculty or federal scientists (collectively called theme leaders) together with NWC staff and other senior scientists to conduct group projects that involve rapid prototyping of new ideas focused on the National Model, which became operational in August 2017. The intent is to create an innovation incubator where students from many universities can exchange ideas and advance concepts that, although developed over a short timeframe and study areas, are illustrative of issues that affect the functioning of the National Water Model. This year's Summer Institute was held June 11-July 27, 2018 and involved 23 graduate students from 18 universities (**Figure 1**).



Figure 1. The affiliations of the students, theme leaders, and course coordinators who participated in the National Water Center Summer Innovators Program Summer Institute 2018.

The first week of the Summer Institute focused on introducing the students to the 2018 Summer Institute themes and to the National Water Model. The students and theme leaders participated in team formation activities, project brainstorming sessions, and data management training sessions. Additional exposure to the nation's water challenges were provided through presentations from theme leaders on topics including flooding in the Lower Rio Grande Valley, the importance of groundwater-surface water exchanges in the Northern High Plains, flood mapping techniques used after Hurricane Harvey, channel routing, crowdsourcing data, data assimilation, real-time flood forecasting, and computational advancements. The theme leaders had a wide range of expertise to broaden the student's knowledge of water resource challenges.

The 2018 Summer Institute themes and theme leaders included:

- The ground and surface water interaction theme was led by Joseph Hughes of U.S. Geological Survey (USGS) and Dave Steward of North Dakota State University. Additional technical support for this theme was supported by Paul Barlow, Bill Cunningham, Wes Zell, and Ward Sanford of the USGS.
- The hyper-resolution modeling theme was led by Fred Ogden of the National Water Center and the University of Wyoming, Jude Benavides of the University of Texas Rio Grande Valley, and Sarah Praskievicz of the University of Alabama.
- The computational aspects of the National Water Model (NWM) and citizen science data themes were led by Ehab Meselhe of The Water Institute of the Gulf and Tulane University, Chris Lowry of The State University of New York at Buffalo, Kyle Mandli of Columbia University, and Sagy Cohen of the University of Alabama.

Two Student Coordinators helped with organization and execution of the projects: Lauren Grimley from the University of Iowa, and Fernando Aristizabal from the University of Florida. Danielle Tijerina of CUAHSI assisted the course coordinators in the organization of the Summer Institute. Several National Water Center staff provided guidance for projects including Ed Clark, Trey Flowers, Fernando Salas, and Nels Frazier.

At the end of the first week of the Summer Institute, the students were led by Martin Briggs of the USGS in fieldwork focused on geophysical methods to investigate groundwater-surface water exchanges. In addition, Sarah Praskievicz demonstrated operation of a high precision survey-grade GPS and Neil Moss from the Alabama Geological Survey discussed groundwater flows specific to Alabama. Whitney Henson of the National Water Center gave a presentation to inform the students of how data can inform flood emergency response units. Dave Gochis and others at National Center of Atmospheric Research (NCAR) provided guidance and information relevant to the core structure of the National Water Model. Steve Turnbull and Nawa Pradhan of the U.S. Army Corps of Engineers, Engineer and Research and Development Center led a two-day GSSHA modeling workshop. Bob Steinke of the University of Wyoming led a two-day ADHydro modeling workshop and continued to provide technical support to the hyper-resolution modeling theme. Additional support to the theme of Citizen Science data was provided by Ben Ruddell of Northern Arizona University. Training in data management and data analysis was provided by Dave Blodgett of USGS and Tony Castronova of CUAHSI.

It can be appreciated that an activity of this magnitude involves a great deal of organization. Jerad Bales, Emily Clark, Ainsley Brown of CUAHSI, and Pamela Harvey of the University of Alabama, were the main people who helped with the institutional arrangements and with travel, housing, and living arrangements in Tuscaloosa. University of Alabama Students Dinuke Munasinghe and Austin Raney assisted with field activities and student projects. The contribution from all Univ. of Alabama support is greatly appreciated.

A key to the success of the National Water Center Innovators Program is the support it receives through the voluntary collaboration of the academic community, along with commercial and government partners. Altogether, over the four Summer Institutes held since the inaugural event in 2015, more than one hundred graduate students have had the experience of working together at the National Water Center in group research projects. Aside from the technical progress that they make, equally important are the friendships formed and professional networks established among the Summer Institute participants that they carry with them into the future. This is a unique and valuable professional experience, and we express our appreciation to the NOAA National Weather Service for

hosting and supporting this innovative activity and this opportunity to contribute to the enhancement of water prediction for our nation.

Lauren E. Grimley

Student Program Coordinator, NWC Innovators Program Summer Institute 2018,
Department of Civil & Environmental Engineering, University of Iowa,
IIHR-Hydrosience & Engineering, Iowa Flood Center

Project Summary

In August of 2016, the NOAA National Water Model, a hydrologic model simulating observed and simulated streamflow over the continental United States (CONUS), became operational marking an unprecedented effort that rightfully challenged the way research is produced and the way hydro-intelligence can benefit society. This modeling framework has also introduced ways in which interdisciplinary research surrounding water resources can be achieved. By providing river forecasts for 2.7 million reaches within CONUS, the NWM provides a nerve center in which, two previously distinct groups - 'data generators' and 'data users' - may unite to answer new questions in a time where population pressures and a changing climate make water management increasingly paramount.

'Data generators' have traditionally included the modeling community and field scientists, who are interested in ensuring the data are as accurate as possible in as many places as possible, and 'data users' are those who use this data to communicate, plan, and study emerging properties surrounding water resources. With the NWM serving as common platform to both generate and provide data, these groups can interact in ways that result in more accurate models and more robust applications. Capitalizing on this opportunity, the fourth National Water Center Innovators Program - Summer Institute brought together 23 students from 18 universities ranging from Oregon to Puerto Rico with backgrounds across a breadth of fields. In an intensive seven-week sprint, these students, with the help of theme leaders, industry professionals, and government researchers, demonstrated how data generators and users can not only come together under the framework of the NWM, but how their knowledge and experience can integrate, supplement, and improve each other's work. In total, seven projects were executed that can thematically be grouped under the domains of hyper-resolution modeling, groundwater, and computational aspects of NWM & citizen science data. Combined, they explore questions relating to the when a hyper resolution model should be implemented, how to represent groundwater/surface water interactions in the NWM, how channel routing and geometry can affect NWM forecasts, and how citizen science data can be processed for NWM implementation.

The first two chapters of this report focus on evaluating and improving the methods used to model groundwater-surface water interactions in the context of the NWM. The NWM only considers ground to stream water fluxes which can be problematic in losing streams where the neighboring groundwater level is lower than the stream stage due to natural topography, drought, or groundwater withdrawals. The Northern High Plains Aquifer (NHPA) was used a case study for this series of studies. The authors of **Chapter 1** chose to analyze the effect of the losing stream mechanisms on the NWM's streamflow prediction abilities in hydrologic extremes. The project utilized USGS gage data, NWM output, groundwater availability models, and Normalize Difference Vegetation Indices (NDVI) from remote sensing to analyze how the NWM's current groundwater/surface water interaction scheme affects flow during floods and droughts. **Chapter 2** involves an investigation into how the NWM estimates baseflow and its performance on two basins with different soil compositions in the NHPA using Rorabaugh and SWAT baseflow formulations as well as USGS gage data. Alternative formulations and methodologies to account for subsurface interflow were explored and demonstrated reductions in error metrics.

The next two chapters were concerned with the implementation of hyper-resolution models in the Lower Rio Grande (LRG) river near Brownsville, TX. This region provided an excellent case-study due to its challenging hydrologic features including clay soils, rapid urban development, extensive system of secondary channels known by the Spanish word *resacas*, and a diverse, transboundary watershed. Both groups utilized the 2-D, physics-based Gridded Surface/Subsurface Hydrologic Analysis model (GSSHA) developed and maintained by the U.S. Army Corps of Engineers (USACE).

Chapter 3 sought to investigate the sensitivity of hyper-resolution models to varying rain event intensities and model spatial resolutions with the goal of providing insights to when and how these models should be implemented. Model outputs of inundated areas were compared relative to each other providing insights to how these factors could help guide decisions about hyper-resolution computations. **Chapter 4** discusses another study that also utilized GSSHA in the Lower Rio Grande River for hyper-resolution modeling. The authors considered how the intricate network of storm sewer pipes affects flood inundation in the Brownsville, TX and what rainfall events they become important.

The last three chapters discuss the computational aspects of the NWM including channel routing schemes, channel geometry methods, and the assimilation of crowd sourced data. In **Chapter 5**, the authors tackled the problem associated with homogenous routing schemes which often results in a trade-off of computational cost for computational accuracy. By associating dimensionless scaling parameters to terms of the Saint-Venant momentum equation, the project shed light on how a heterogeneous routing scheme could be better optimized for computing cost and accuracy metrics. **Chapter 6** includes an assessment of the NWM trapezoidal channel geometries compared to actual channel geometry utilizing the HEC-RAS model as a catalyst. Improvements were suggested for better representing channel geometry while not sacrificing parsimony. **Chapter 7** is concerned with how stage data collected from citizen scientists can be used to enhance the NWM. The group utilized a decision tree model to turn raw data into corrected data by factoring in several variables that could indicate the reliability of a crowd sourced data point.

The students of the 2018 Summer Institute have been an absolute pleasure to work with and all successes are a testament to their work ethic, ability to collaborate, and dedication to spend their summer away from friends, co-workers, and family. The breadth of their work and respective backgrounds highlights the impact a community model can have on bringing scientists together. They have all contributed towards a goal of changing how water resources research is carried out and how citizens can engage with their water circumstance. More so, this shared experience has brought this academic community closer by fostering strong personal and professional relationships that will last for the length of their careers.

The reports published here represent the culmination of seven weeks of research. Whether ‘data-generators’ or ‘data-users,’ this work presents a platform not only for these students’ continued collaboration with their advisors, each other, and the NWC, but also enables the rest of the community to become more engaged in the development of the NWM and the overall advancement of hydrologic science.

Fernando Aristizábal

Student Program Coordinator, NWC Innovators Program Summer Institute 2018,
Department of Agricultural & Biological Engineering,
Center for Remote Sensing, University of Florida

Chapter 1

Groundwater-Surface Water Exchange and Streamflow Prediction using the National Water Model in the Northern High Plains Aquifer region, USA

Elizabeth Jachens¹, Holly Hutcheson², and Matt Thomas³

¹ Department of Biological and Ecological Engineering, Oregon State University; jachense@oregonstate.edu

² Department of Geology, The University of Georgia; hollyhutcheson@uga.edu

³ Department of Geology, The University of Georgia; matthew.thomas@uga.edu

Academic Advisors: John Selker, Oregon State University; Adam Milewski, University of Georgia

Summer Institute Theme Advisors: David Steward, North Dakota State University, david.steward@ndsu.edu; Joseph Hughes, USGS, jdbughes@usgs.gov

Abstract: Currently, the National Water Model (NWM) only considers a one-way flux between streams and underlying aquifer systems, where groundwater enters streams but cannot return to the aquifer. Consequently, streamflow only gains discharge from groundwater in the downstream direction. Losing stream conditions occur where the stream stage is higher than the surrounding groundwater table as a result of natural topography, drought, or groundwater pumping. The purpose of this study is to evaluate how the inability to simulate losing streams in the NWM impacts streamflow prediction during hydrologic extremes. This study identifies possible losing stream reaches using USGS gage data, an existing USGS Groundwater Availability Model for the Northern High Plains Aquifer, and Normalized Difference Vegetation Index (NDVI) from remote sensing techniques in the Northern High Plains Aquifer, USA. While losing streams are traditionally identified using field techniques like seepage meters and soil temperature, this novel integrated approach utilized identification techniques with high spatial and temporal resolution to provide a preliminary assessment of losing reaches in the area. Hydrographs comparing the NWM and USGS gage data for losing and gaining stream reaches were created across floods and drought. For flood events, the lack of a losing stream mechanism resulted in modeled flood response characterized by an earlier peak discharge and an overestimate of the observed flood volume. For intermittent streams and in drought conditions, the NWM was unable to capture streams that go dry because there is no mechanism to lose water from the streams. On average, the USGS Groundwater Availability Model identified a losing mechanism contributing to streamflow loss of 0.1% per mile. The NWM has been shown to overestimate streamflow in both losing and gaining streams during drought conditions in the Northern High Plains aquifer. By incorporating a streamflow loss correction factor, the NWM could potentially improve predictions of the magnitude of streamflow prediction during hydrologic extremes.

1. Motivation

Extreme weather events are the single highest risk event in terms of impact and likelihood according to the World Economic Forum Global Risks Perception Survey in 2016 [1]. With extreme weather events predicted to increase frequency and intensity, forecasting and warning of these events will play

a significant role is water management in the future [2]. The National Water Model (NWM) currently provides Contiguous United States (CONUS) scale water forecasting capability at one-kilometer grid resolution and model improvements in hydrologic extreme forecasting are currently in development. Improved flood hydrology for the NWM is slated for completion in fiscal year 2020. Additionally, the integration of the NWM with groundwater and transport models to for low flow predictions will be incorporated to provide a complete integrated risk evaluation.

Streamflow predictions for hydrologic extremes will require a comprehensive model that captures surface water and groundwater exchange processes. Despite acknowledgment of the need to manage groundwater and surface water as a single resource, in its current form the NWM provides a weak groundwater representation and a limited handling on groundwater-surface water exchange [3, 4, 5]. Currently the NWM only considers a one-way flux between the stream and the aquifer, where water can enter the stream from the groundwater (hereafter referred to as a gaining stream) but water cannot leave the stream. Consequently, streamflow does not have a mechanism for losing water in the downstream direction (referred to as a losing stream). A losing stream can occur where the stream water surface is higher than the surrounding aquifer water table and the stream bed permeability permits flow such that the hydraulic gradient indicates that flow will flow from the stream to the groundwater [5]. As a result, streamflow in the NWM has the potential to systematically overestimate streamflow in regions where the losing condition indicates that water would otherwise be leaving the stream system, especially in instances of extreme hydrologic events.

2. Background

In its current version, the NWM only includes mechanisms for streams to gain flow from the non-linear reservoir representation of groundwater to surface water in streams, ignoring the effects of losing streams as a source of aquifer recharge. The assumption that streams do not lose water to the groundwater may not hold in areas where the water table is below the streambed due to natural variations in topography, prolonged drought, or anthropogenic causes such as agricultural pumping [6, 7]. One such location where the water table can be significantly below the stream level, thus invalidating the gaining stream assumption, is the Northern High Plains Aquifer (NHPA) with depth to aquifer varying up to several hundred feet below ground surface [8].

The NHPA has been a region of interest considering the interaction between groundwater and surface water since the 1970's with one of the first papers describing the transmission losses of streamflow to the groundwater for various streams in the region [9]. The USGS High Plains Regional Aquifer-System Analysis was released in 1988, highlighting the declines in groundwater levels that put the aquifer and basin on the map for future studies [19]. This report emphasized the complexity of the groundwater system in the region and the unique issues associated with agricultural pumping and groundwater levels that are still major areas of research. In the NHPA, the low precipitation months coincide with the irrigation pumping season; both factors contribute to a lowering of the groundwater table associated with losing stream conditions. Agricultural irrigation in the NHPA is primarily sustained by groundwater pumping, placing a large demand on the aquifer [10]. The large number of agricultural pumping wells and the large discharge volumes contribute to both spatial and temporal changes in losing stream conditions based on the pumping cycles and locations of the well. Recently, interest in the spatiotemporal distribution of groundwater recharge and transmission losses has grown in the scientific literature [20, 21, 22, 23]. With varying geology and precipitation amounts across the region, capturing the transient nature of groundwater-surface water fluxes needs to be considered when addressing losing streams.

As a result of the historically documented losing stream conditions in the region, the analysis in this

paper focuses on case studies of the NHPA because the lack of two-way stream-aquifer coupling in the NWM may result in uncertainty in streamflow prediction ability. For regions with significant losing streams like the NHPA, the NWM has the potential to systematically overestimate streamflow in regions where the condition indicates that water would otherwise be leaving the stream system. For flood events, we hypothesize that the NWM will result in a more intense modeled flood response with an earlier flood peak and overestimated discharge [11]. For drought conditions, the omissions of a mechanism for the water to exit the stream prevents the stream from going dry even if water table and channel conditions drive losing stream conditions. We hypothesize that the lack of a losing stream mechanism in the NWM will result in some stream reaches having modeled non-zero flows during no-flow conditions.

3. Methodology

The following section breaks down the methodology into two sections: (1) the identification of possible gaining and losing streams and (2) the analysis of NWM performance in extreme hydrologic events on these reaches. Identification of potential losing and gaining stream reaches was completed using USGS streamgage data, a USGS Groundwater Availability Model for the NHPA, and remote sensing data. The analysis consists of constructing longitudinal discharge profiles along some stream reaches and testing for differences between modeled and observed discharges in identified reaches during extreme events.

3.1. Identification Methods for Surface Flux to Groundwater

While the primary purpose of this study was to evaluate the NWM's ability to predict streamflow in hydrologic extremes for losing stream conditions, first losing streams needed to be identified. In the past, stream reaches in the NHPA have been identified as losing or gaining by field measurements including seepage meters, sediment cores for hydraulic conductivity with depth, sediment temperature, as well as groundwater modeling [12, 13, 14]. Many of these methods for determining groundwater-surface water flux involve time and monetarily expensive field measurements and equipment which would not be feasible to perform for all stream reaches within the NHPA. Additionally, these field measurements are only representative of static locations for a given time and thus do not capture the spatiotemporal variability of groundwater-surface water flux. Consequently, these field based techniques for identifying losing stream reaches have limited applicability for generalization of such processes and trends for the larger region. In order to capture the spatial and temporal variability of losing stream reaches, we used identification methods that do not require significant field work, have good spatial coverage, and are transient by nature. The three data sources that were selected to identify losing streams in the NHPA were USGS streamgages, an existing USGS Groundwater Availability Model, and remote sensing products.

Losing streams were first identified by streamflow differencing using USGS discharge data. This method was selected because the gages are widespread and have accessible data at high temporal resolution. Since stream gages only report streamflow, identifying losing stream reaches using this method only accounts for reach average changes in total discharge and does not capture finer scale changes between gages. Losing stream reaches were identified as the reach upstream of a gage with a lower mean annual discharge than the next gage upstream.

Secondly, losing streams were identified using the USGS Groundwater Availability Model. This transient groundwater flow model was constructed to evaluate the water resources within the NHPA and quantify the groundwater-surface water interactions with the inclusion of the effects of groundwater pumping wells. The USGS Groundwater Availability Model was chosen because the

existing model provides a detailed biannual groundwater-surface water flux for 1km grid cells along the stream for the NHPA from 1940-2008. Limitations of the model in identifying losing stream reaches include no modeled surface runoff and constant environmental forcings.

The last method for determining the location of losing streams utilized remote sensing data. Landsat 5 and 7 images with less than 10% cloud cover were used from 1993-2009. The Normalized Difference Vegetation Index (NDVI) and the Normalized Difference Water Index (NDWI) were calculated for 1km grids along the streams. NDVI was used to proxy tree cover and vegetation including cottonwood trees. With relatively shallow root zones, cottonwood tree growth has been associated the groundwater water table with slower tree growth rates associated with losing stream reaches compared to the faster growth rates within gaining reaches [15]. Thus, areas without substantial vegetation are likely in areas with deeper groundwater table levels and may be indicative of losing stream conditions. Grids with low NDVI values (0-0.2) (indicating barren and low shrub lands) and do not have water confirmed by the NDWI were identified as possible losing reach locations for this method [24]. Using remote sensing provides easy access to high spatiotemporal resolution imagery; however, this method uses vegetation as a proxy for streamflow and does not consider other environmental factors that might affect the relationship between streamflow and vegetation.

Each of the three identification techniques produced a different map of likely losing and gaining streams (**Figure 1**). The three identification methods each produce one metric of a losing stream: decreased streamflow in the downstream direction, groundwater model indicating a hydraulic gradient with flux from the stream to the aquifer, and vegetation. However, no method directly measures or simulated the groundwater-surface water flux and thus no single method was an absolute metric of a losing stream's flux from streamflow to groundwater. Instead, each method provides a different way to identify losing stream conditions and thus a combination of the methods should be used for final selection of reaches to study. An overlay of losing and gaining stream reaches was used to identify stream locations as being likely to be a losing or gaining stream when flagged by at least two of the three identification methods. This allows for reducing the total streams of interest to only streams that are susceptible to losing conditions.

3.2. Evaluation Methods

In order to evaluate the performance of the NWM, longitudinal profiles and hydrographs for NWM streamflow were compared with those for corresponding USGS streamgages. The spatial extent of discharge in the longitudinal profile is one way to visualize and compare the data, with the advantage of geologic references and downstream trends.

The differences in mean overprediction of streamflow per event and flood peak timing between the NWM and corresponding streamgages was the primary metric for determining NWM streamflow prediction ability. Overpredictions were determined for rising, peak, and falling regions of the hydrograph for flood events while overpredictions for the entire event were used for droughts. Averages are reported per event, per location and summarized over the entire period of record for losing and gaining streams. The differences in mean streamflow overprediction in each event

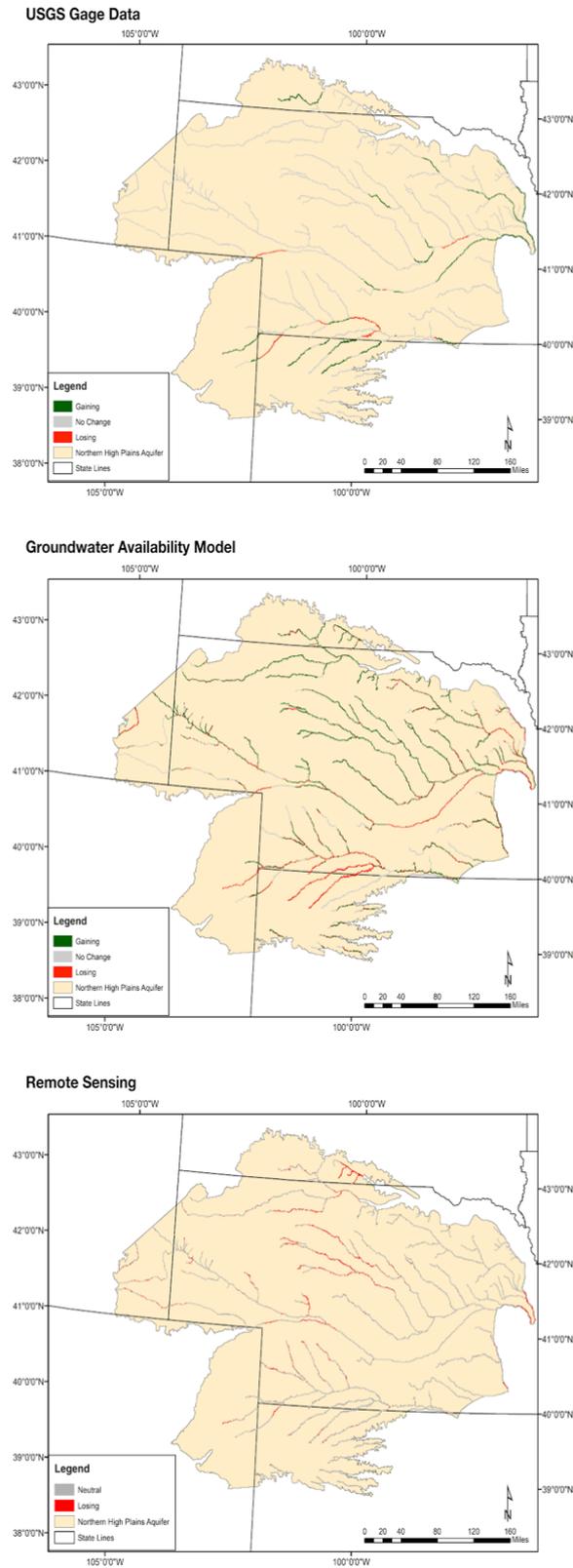


Figure 1. Maps of losing stream reaches in the Northern High Plains Aquifer in 2003 produced by USGS Gage Data, the USGS Groundwater Availability Model and Remote Sensing.

between gaining and losing reaches were analyzed in order to determine whether the difference in mean flow overestimate by the NWM between losing and gaining reaches was significant.

Extreme hydrologic events included in this study are floods and droughts based on water years 1993-2008 and 2000-2008, respectively. Flood events were chosen by identifying where the USGS discharge at the gage was above the bankfull discharge using the Groundwater Toolbox [16]. For gages that did not have a bankfull discharge available, a return period of 5 years was used to estimate flooding events from streamgage data. Drought events are selected using the US Drought Monitor records for events from D1-D4 severity [17]. T-tests were then used to investigate the potential difference between the NWM estimation in flow in gaining and losing reaches for both different severity levels and stream orders.

4. Results and Discussion

The streamflow prediction ability of the National Water Model (NWM) for losing and gaining streams was tested using the case study region of the Northern High Plains Aquifer (NHPA) between 1993 and 2009.

4.1. Losing Stream Reach Identification

Probable losing stream reaches to be compared against the NWM were identified when at least two of the three independent identification techniques flagged a reach as

losing. **Figure 2** shows an overlay map of possible losing stream locations for each of the three

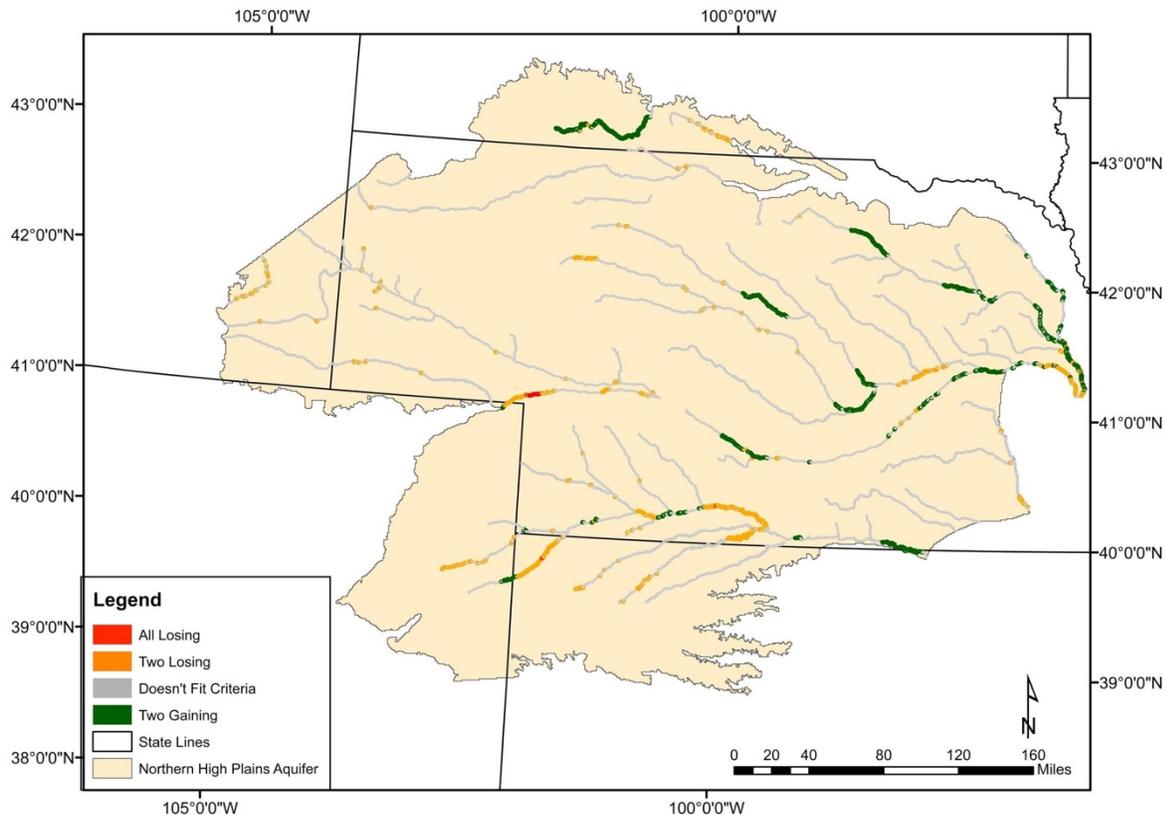


Figure 2. Probable losing and gaining stream reaches for the Northern High Plains Aquifer in 2003.

identification methods for a representative pumping season of May-September, 2006. The northern basins in the NHPA are predominantly characterized by the Sand Hills region. This area has a large saturated thickness of 1000 feet or more, minimal surface or groundwater irrigation, and significant annual recharge from precipitation; by our analysis streams tend to be gaining [8]. The shallow depth to groundwater suggests that losing stream conditions are not likely. In contrast, the Platte and Republican River Basins are characterized by conditions where the water table is likely below the stream water level, and by our analysis have losing reaches. Both of the major basins in southern NHPA are characterized by Ogallala formations with Quaternary valley fill deposits with soil classes of sandy clay loam and aquifer saturated thickness ranging between 0-500 feet with large depths to aquifer, paired with anthropogenic influences to the natural streamflow including groundwater pumping and surface water diversion [8].

The Republican River was selected as a representative river for analysis because it has both losing and gaining reaches according to each of the three identification techniques. **Figure 3** shows the spatial trends of the average stream discharge from May-September 2016 from upstream to downstream for the NWM, USGS stream gages, and the USGS Groundwater Availability Model for the Republican River. The longitudinal profile shows the spatial extent of the discharge, with

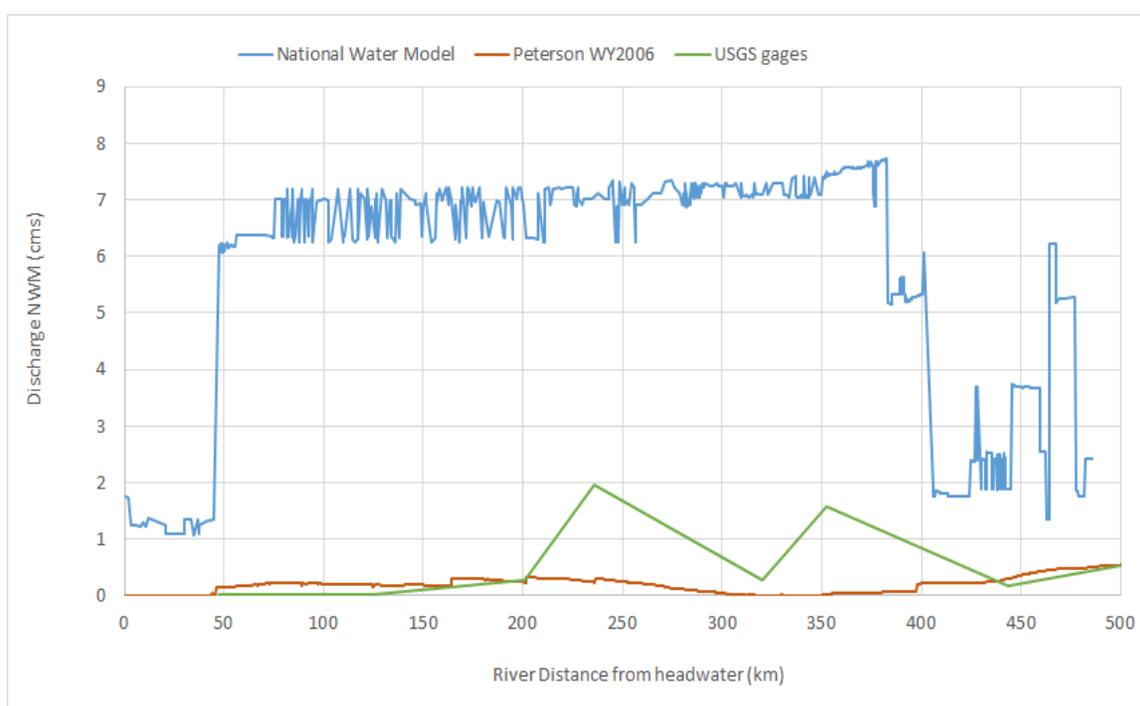


Figure 3. Longitudinal profile of the Republican River for average discharge for May-Sept 2006.

decreases in streamflow associated with stream diversions or losing stream conditions and increases in streamflow associated with incoming tributaries or springs [18]. Significant geologic features that contribute to changes in streamflow include a major tributary approximately 50 km downstream, a high density of pumping wells between 250 and 350 km downstream, and the Harlan Reservoir at 400 km downstream. However, not all changes in flow can be immediately explained by geologic features. The NWM does not show a general increase in streamflow in the downstream direction,

despite numerous tributaries that contribute flow to the mainstem of the Republican River. The noise in the NWM is likely a product of the independent calculation of discharge in stream segments by the NWM. Additionally, the NWM discharge is notably higher than the discharge measured by the USGS Gages. Linear interpolation is applied in between the eight USGS Gages. The USGS Groundwater Availability Model captures the decrease in streamflow associated with the agricultural pumping, but does not include runoff which could underestimate flow. All things considered, the NWM discharge is systematically higher than the USGS Gage data and does not show a change in flow in the downstream direction.

4.2. Losing Stream Distribution

The groundwater-surface water flux was quantified using the USGS Groundwater Availability Model. The distribution of groundwater-surface water flux of stream segments is centered around 0, with 20% of stream cells with no groundwater-surface water flux (**Figure 4**). With 25% of the stream segments designated as losing reaches, the arithmetic average of stream loss is 2% of streamflow while the weighted average normalized by portion of streamflow in the basin is 0.1% (**Figure 5**). When compared to all stream segments, streamflow loss accounts for a cell-average of 10% and a streamflow weighted average of 0.5% of flow in the stream segment (**Figure 6**).

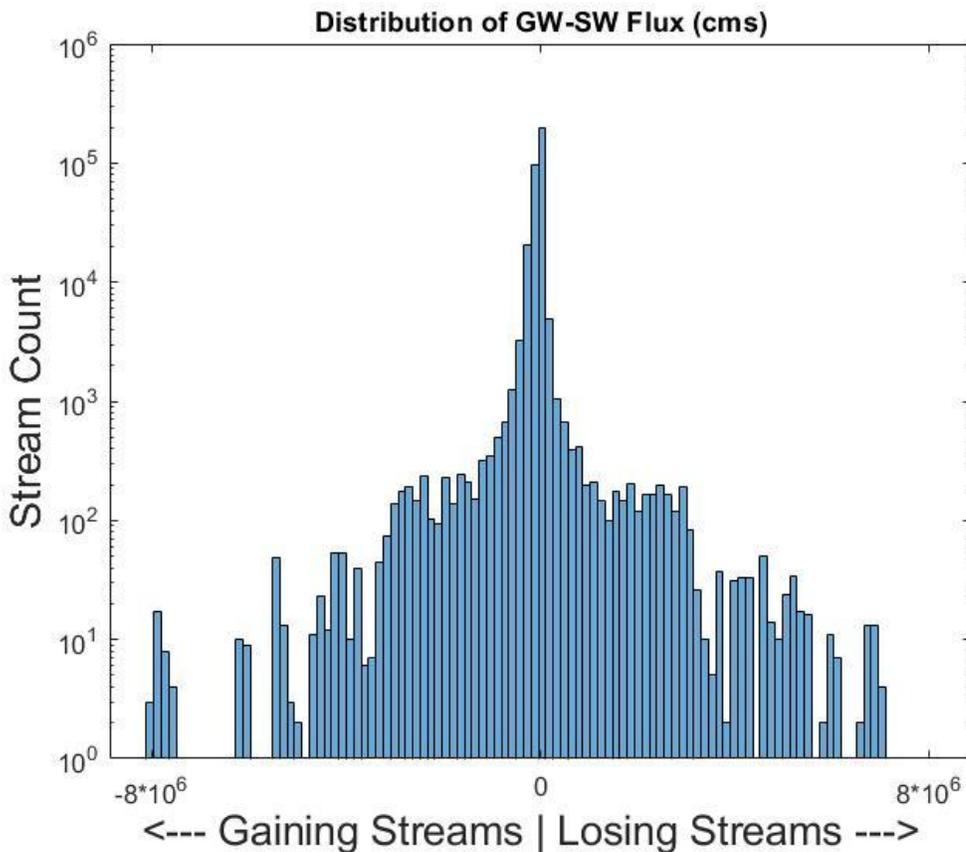


Figure 4. Distribution of GW-SW flux for stream segments in the NHPA.

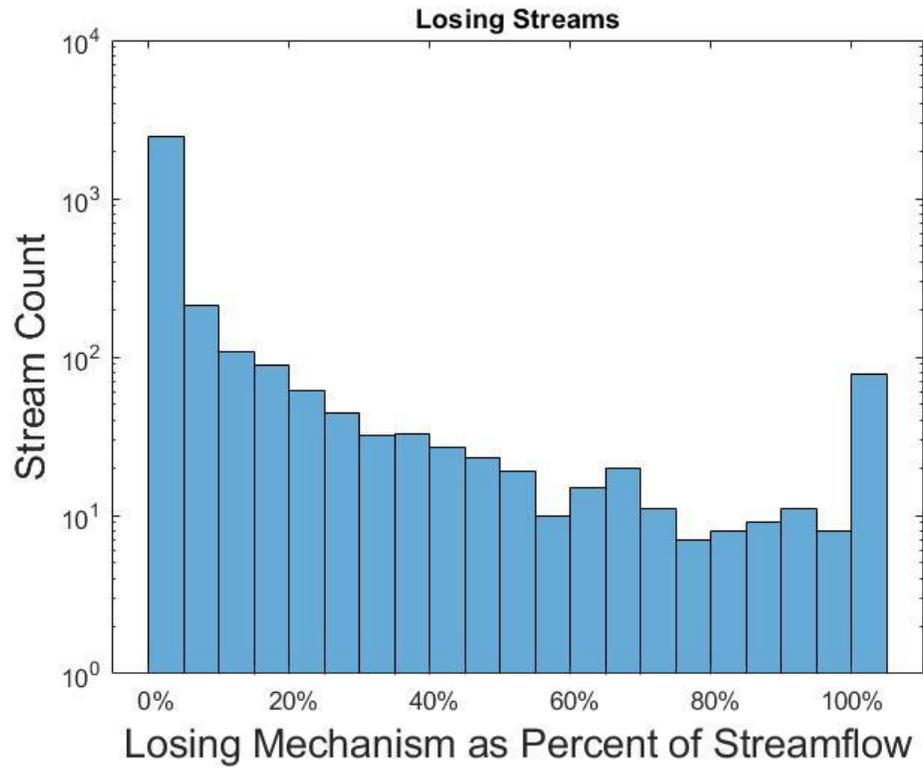


Figure 5. Distribution of loss as a percent of streamflow for designated losing reaches.

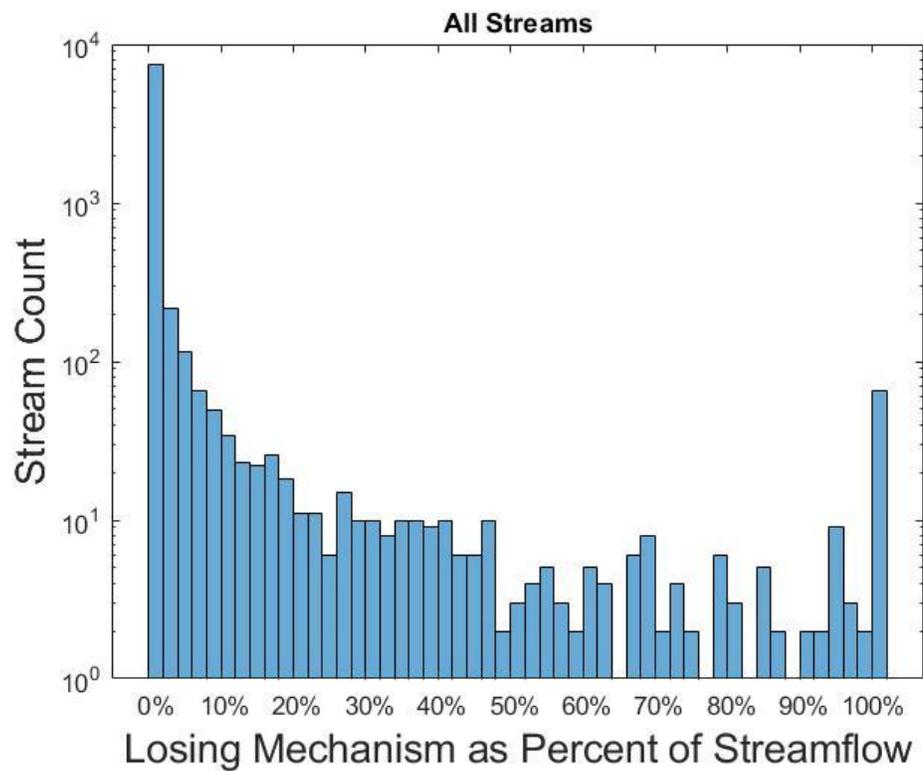


Figure 6. Distribution of loss as a percent of streamflow for total streamflow in all streams.

4.3. Floods

Due to small sample size, no statistical tests were performed for flood events losing and gaining reaches over the time period 2003-2008. However, using observations of the hydrographs, there are several differences between the NWM and USGS streamgauge hydrographs for a section of the Republican River. NWM hydrographs for flood events tend to have much larger predicted peak discharges as well as earlier peaks compared to USGS streamgages in losing streams (**Figure 7.a**). While the peak discharge in the NWM is still considerably higher than USGS streamgauge data in gaining streams, the difference is less than with losing streams (**Figure 7.b**). The NWM predicts peak

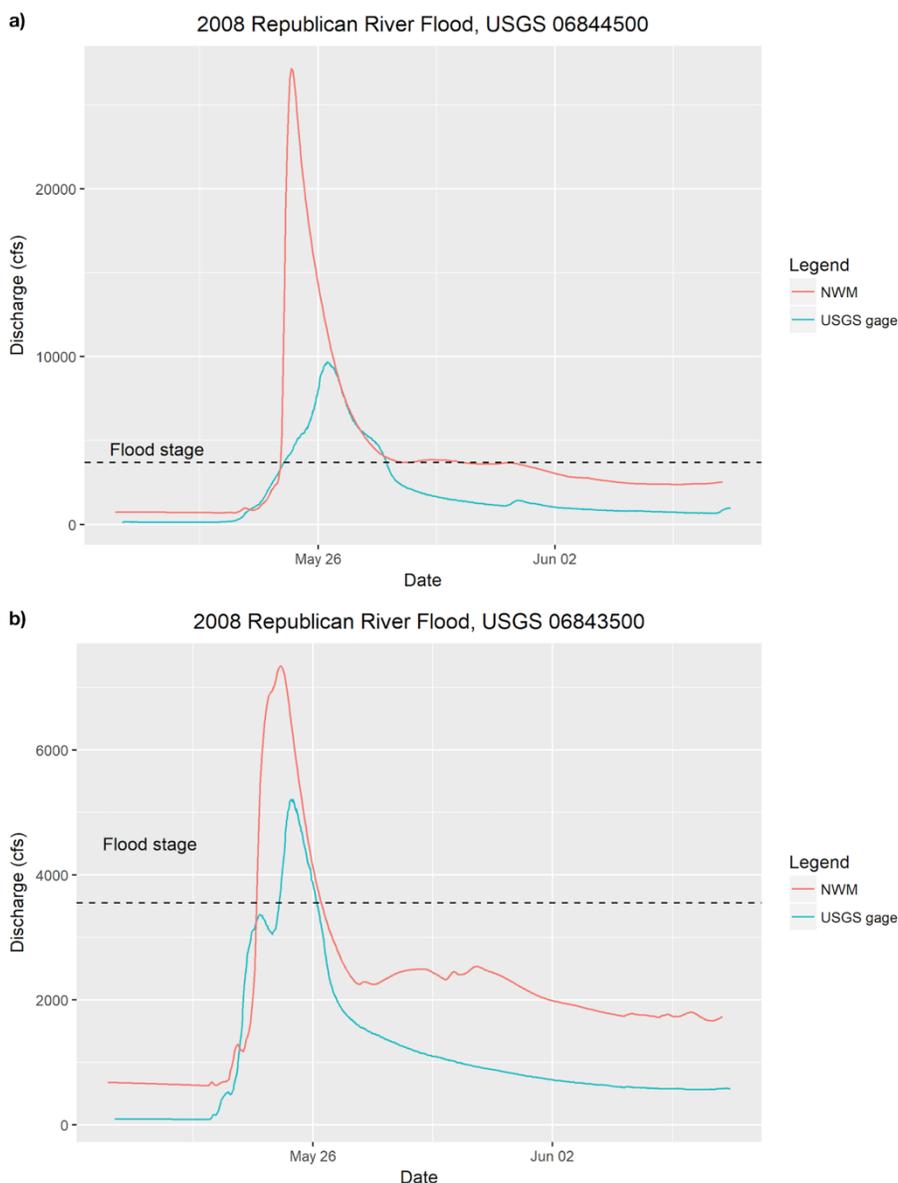


Figure 7. (a) Hydrograph for May 26, 2008 flood event on the Republican River on a losing reach. The NWM predicts a 17000 cfs higher peak discharge occurring 8 hours earlier than the observed peak; (b) Hydrograph for the same flood event on an upstream gaining reach. The NWM predicts a 2100 cfs higher peak discharge occurring 8 hours earlier than the observed peak.

discharges and peak timing that are larger in magnitude and earlier for losing streams compared to gaining streams (**Table 1**). While the NWM streamflow prediction skill tends to be better in gaining reaches than in losing reaches, the NWM systematically overestimates the peak discharge with an early peak timing.

Table 1. Summary of Differences in Flood Prediction in Gaining vs Losing Streams for the example in Figure 4.

	Losing	Gaining
Peak Discharge (cfs)	17516	2156
Peak Timing (hrs)	26	8
Total Volume (ft³)	2.2 (10) ⁹	1.4 (10) ⁹

Another factor that could be attributed to the higher peak discharge of NWM besides a lack of a losing stream mechanism like bank storage is the singular continuous trapezoidal channel geometry used in the NWM. By modeling a channel as having infinitely deep sides, there is no floodplain for the water to overflow onto and as a result that water stays in the channel producing a higher peak flow. Additional research on more representative stream cross sections to include a floodplain should be pursued in combination with a losing stream mechanism in order to improve flood modeling in the NWM.

4.4. Droughts

Statistics of identified losing streams compared to gaining streams for droughts are summarized in **Table 2**. Hydrographs for drought events in losing reaches and gaining reaches are not statistically different ($t = -0.507$, $df = 366$, $p\text{-value} = 0.612$). However, the NWM hydrographs for drought tend to have much larger predicted discharges compared to the USGS Gages in both losing and gaining streams (**Figure 8.a-b**). In both cases, the NWM fails to model a dry stream when the USGS Gage indicates that the stream is not flowing.

Table 2. Summary of differences in drought prediction in gaining vs losing streams for the example in Figure 5.

	Losing	Gaining
Mean Discharge (cfs)	259	250
Flow Volume (ft³)	3.2 (10) ⁹	3.3 (10) ⁹

In addition to looking at overall losing and gaining stream differences, differences in the overestimation of losing and gaining streams were investigated by drought severity level. While the p -values suggest that the difference is not significant between losing and gaining streams, overall the gaining streams have a higher overestimation (2589.09 cfs) than the losing streams (2396.71 cfs) (**Table 3**).

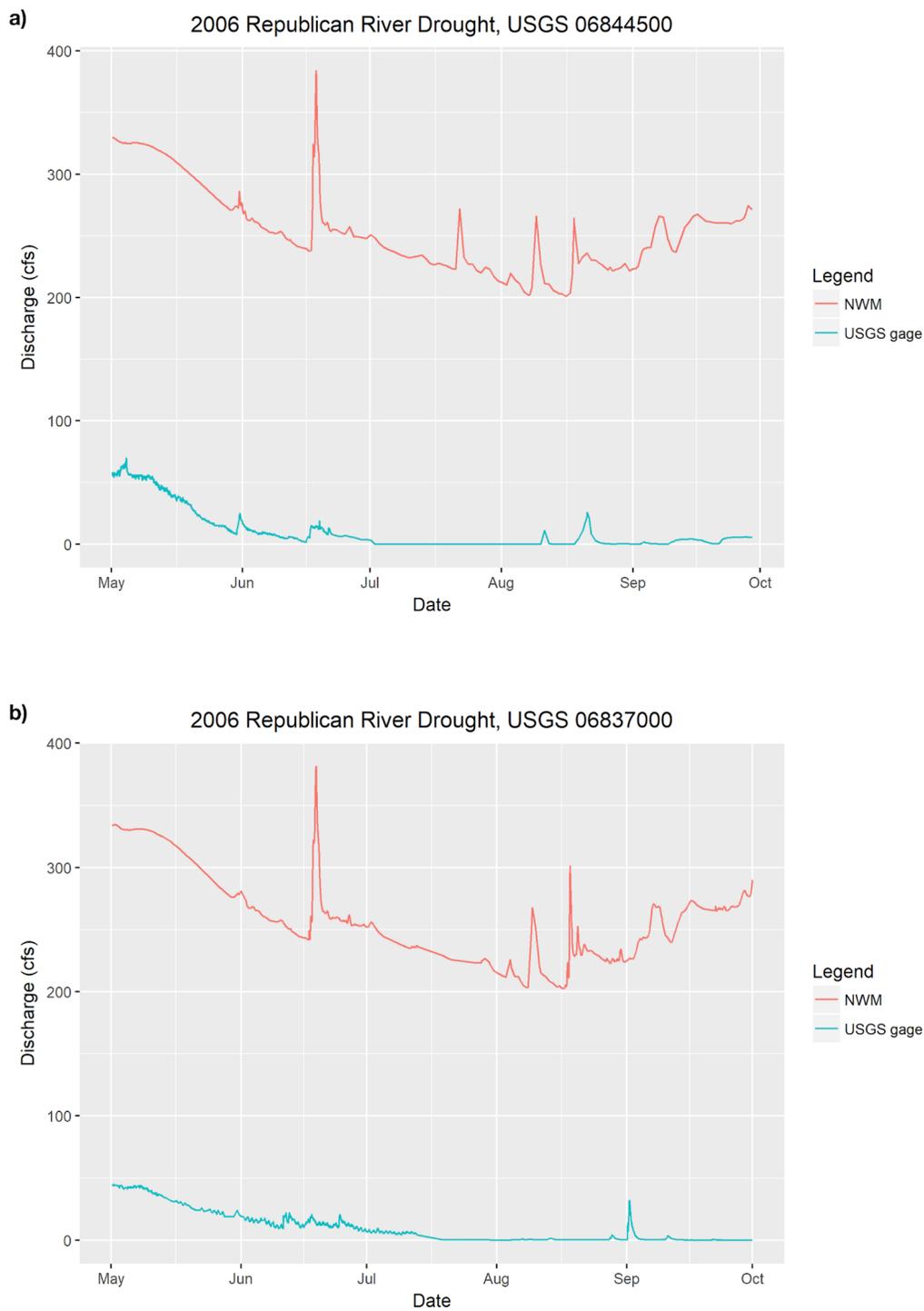


Figure 8. (a) Hydrograph for a 2006 drought event on a losing reach of the Republican River. The mean overestimate of discharge by the NWM is 259 cfs; (b) Hydrograph for the same drought event on an upstream gaining reach of the Republican River. The mean overestimate of discharge by the NWM is 250 cfs.

Table 3. *Summary of the Statistic Differences in the drought severities for losing and gaining streams.*

Drought Level	t-score	Degrees of Freedom	p-values	Mean losing overestimate (cfs)	Mean gaining overestimate (cfs)
All	-0.51	366.14	0.61	2396.71	2589.09
D1	0.12	159.36	0.90	2958.35	2889.53
D2	-1.25	148.58	0.22	1922.53	2587.57
D3	0	60	1	2414.61	2414.61
D1-2	-0.68	306.20	0.50	2476.40	2745.85

Stream orders were also considered when looking at the overestimation of the NWM. According to the t-tests run on different stream orders, the NWM overestimates more in gaining 5th order streams and in losing 6th and 7th order streams. The 4th and 8th order streams did not have a significant difference in overestimates (**Table 4**).

Table 4. *Summary of the Statistic Differences in the stream orders for losing and gaining streams.*

Stream order	t-score	Degrees of Freedom	p-values	Mean losing overestimate (cfs)	Mean gaining overestimate (cfs)
All	-0.51	366.14	0.61	2396.71	2589.09
4	-1.32	49.72	0.19	22.38	40.53
5	-7.11	60.16	1.61 (10) ⁻⁹	3.23	182.88
6	3.36	57.75	1.40 (10) ⁻³	902.54	428.54
7	2.74	21.34	0.01	2079.58	596.37
8	1.41	57.02	0.16	8765.86	7636.76

5. Conclusion

The National Water Model (NWM) does not include two-way surface water and groundwater coupling- however both components should be included to effectively model or manage water as a single resource. Losing streams can significantly change the streamflow pattern because water can exit the stream and recharge the groundwater. Losing streams should be considered particularly in areas where the groundwater is significantly below the surface water. With the natural variations in topography leading to large depths to groundwater and contribution of agricultural pumping, the Northern High Plains Aquifer is one location where losing stream conditions are met and have been confirmed in the past.

This study uses the Northern High Plains Aquifer as a case study for determining the NWM prediction skill for losing versus gaining streams during hydrologic extreme events, namely droughts and floods. Overall, the NWM systematically overestimates streamflow which could be a result of excluding a losing stream mechanism. Although no results were made with the overall floods in the study area,

the NWM over-predicted the peak discharge as well as predicted the peak to be over a day earlier in some cases with losing streams performing worse than gaining streams in a flood example on the Republican River. Future research should have a larger sample size of flood events so that statistical tests can be run. Overall there was not a statistical difference in streamflow prediction skill for drought events for losing versus gaining streams. For drought events, losing and gaining reaches weren't statistically different in NWM overprediction of streamflow. Similarly, no difference was found between losing and gaining reaches when drought events were analyzed by severity. NWM overprediction was significantly higher in losing reaches on sixth and seventh order streams; analysis of fifth order streams showed that overpredictions were instead significantly higher in gaining streams. Many streams are seasonally dry in the Northern High Plains Aquifer, which the NWM was not able to capture at all. This research would benefit from a larger sample size for analysis accomplished by extending the time period in order to include more hydrologic extreme events to better describe the system as a whole.

While a comprehensive groundwater model being incorporated into the NWM would be computationally expensive and require dedication of resources and is therefore not currently feasible, the NWM could make steps towards creating a more accurate representation of groundwater and stream water flux by incorporating a losing stream term into the water balance. This study proposes using an average stream loss term for a region to represent a losing stream reach with water exiting the stream. The total water lost from the stream that went into the aquifer is 10% in losing stream reaches, or alternatively 1% of the total streamflow in the entire region of the Northern High Plains Aquifer using the USGS Groundwater Availability Model. Thus, the NWM could incorporate a uniform losing term of 1% as a losing stream mechanism in order to improve the NWM streamflow prediction. However, this method of applying a blanket percent loss term will vary across regions and thus would need to be calculated before being applied to the NWM. Additional work on quantifying the losing stream potential and the corresponding changes in streamflow should be pursued to improve the NWM.

Acknowledgements: We wanted to thank all of the following people who helped with guidance, data and resources, technical assistance, and support: David Steward, Joseph Hughes, Dave Blodgett, Wes Zell, Paul Barlow, Bill Cunningham, Ward Sandford, Bradford Bates, and Fernando Aristizabal.

References

1. US Drought Monitor. (2018). Retrieved from <https://www.climate.gov/maps-data/data-snapshots/data-source-drought-monitor>
2. *Attribution of Extreme Weather Events in the Context of Climate Change*. (2016). Washington, D.C.: The National Academies Press. <https://doi.org/https://doi.org/10.17226/21852>
3. Sophocleous, M. (2002). Interactions between groundwater and surface water: The state of the science. *Hydrogeology Journal*, 10(1), 52–67. <https://doi.org/10.1007/s10040-001-0170-8>
4. Woessner, W. W. (2000). Stream and fluvial plain ground water interactions: Rescaling hydrogeologic thought. *Ground Water*. <https://doi.org/10.1111/j.1745-6584.2000.tb00228.x>
5. Winter, T. C., Harvey, J. W., Franke, O. L., & Alley, W. M. (1999). *Ground water and surface water: a single resource*. U.S. Geological Survey Circular 1139. Retrieved from <http://pubs.usgs.gov/circ/1998/1139/report.pdf>
6. Hiscock, K. M., & Grischek, T. (2002). Attenuation of groundwater pollution by bank filtration. *Journal of Hydrology*, 266(3–4), 139–144. [https://doi.org/10.1016/S0022-1694\(02\)00158-0](https://doi.org/10.1016/S0022-1694(02)00158-0)

7. Barlow, P. M., & Leake, S. A. (2012). *Streamflow depletion by wells-Understanding and managing the effects of groundwater pumping on streamflow*. U.S. Geological Survey Circular 1376.
8. Peterson, S. M., Flynn, A. T., & Traylor, J. P. (2016). Groundwater-Flow Model of the Northern High Plains Aquifer in Colorado, Kansas, Nebraska, South Dakota, and Wyoming. U.S. Geological Survey, (Scientific Investigations Report 2016-5153), 88. <https://doi.org/10.3133/sir20165153>
9. Jordan, P.R. (1977). Streamflow transmission losses in western Kansas. ASCE 103 1977; HY8: 905-19
10. Steward, D. R., Yang, X., Lauwo, S. Y., Staggenborg, S. A., MacPherson, G. L., & Welch, S. M. (2011). From precipitation to groundwater baseflow in a native prairie ecosystem: A regional study of the Konza LTER in the Flint Hills of Kansas, USA. *Hydrology and Earth System Sciences*. <https://doi.org/10.5194/hess-15-3181-2011>
11. Pacheco-Guerrero, A., Goodrich, D. C., González-Trinidad, J., JÚnez-Ferreira, H. E., & Bautista-Capetillo, C. F. (2017). Flooding in ephemeral streams: Incorporating transmission losses. *Journal of Maps*, 13(2), 350–357. <https://doi.org/10.1080/17445647.2017.1305303>
12. Chen, X., Dong, W., Ou, G., Wang, Z., & Liu, C. (2013). Gaining and losing stream reaches have opposite hydraulic conductivity distribution patterns. *Hydrology and Earth System Sciences*, 17(7), 2569–2579. <https://doi.org/10.5194/hess-17-2569-2013>
13. Rosenberry, D. O., & Pitlick, J. (2009). Local-scale variability of seepage and hydraulic conductivity in a shallow gravel-bed river. *Hydrological Processes*. <https://doi.org/10.1002/hyp.7433>
14. Dong, W., Chen, X., Wang, Z., Ou, G., & Liu, C. (2012). Comparison of vertical hydraulic conductivity in a streambed-point bar system of a gaining stream. *Journal of Hydrology*, 450–451, 9–16. <https://doi.org/10.1016/j.jhydrol.2012.05.037>
15. Harner, M. J., & Stanford, J. A. (2018). Differences in Cottonwood Growth between a Losing and a Gaining Reach of an Alluvial Floodplain Author (s): Mary J . Harner and Jack A . Stanford Published by : Wiley on behalf of the Ecological Society of America Stable URL : <http://www.jstor.org/stabl>, 84(6), 1453–1458.
16. Barlow, P.M., Cunningham, W.L., Zhai, Tong, and Gray, M. (2017). U.S. Geological Survey Groundwater Toolbox version 1.3.1, a graphical and mapping interface for analysis of hydrologic data: U.S. Geological Survey Software Release. Retrieved from <http://dx.doi.org/10.5066/F7R78C9G>
17. US Drought Monitor. (2018). Retrieved from <https://www.climate.gov/maps-data/data-snapshots/data-source-drought-monitor>
18. Larned, S. T., Arscott, D. B., Schmidt, J., & Diettrich, J. C. (2010). A framework for analyzing longitudinal and temporal variation in river flow and developing flow-ecology relationships. *Journal of the American Water Resources Association*. <https://doi.org/10.1111/j.1752-1688.2010.00433.x>
19. Weeks, J. B. (1988). Summary of the High Plains regional aquifer-system analysis in parts of Colorado, Kansas, Nebraska, New Mexico, Oklahoma, South Dakota, Texas, and Wyoming [1988]. Retrieved from <http://agris.fao.org/agris-search/search.do?recordID=US8921597>
20. Lauffenburger, Z. H., Gurdak, J. J., Hobza, C., & Woodward, D. (2018). Irrigated Agriculture and future climate change effects on groundwater recharge, norther High. *Agricultural Water Management*.
21. Scanlon, B. R., Healy, R. W., & Cook, P. G. (2002). Choosing appropriate technique for quantifying groundwater recharge. *Hydrogeology Journal*, 10, 18–39. <https://doi.org/10.1007/s10040-0010176-2>

22. Sjodin, A., Lewis, W. M., & Saunders, J. F. (2001). Analysis of groundwater exchange for a large plains river in Colorado (USA). *Hydrological Processes*, 15(4), 609–620. <https://doi.org/10.1002/hyp.173>
23. Zhang, J., Felzer, B. S., & Troy, T. J. (2016). Extreme precipitation drives groundwater recharge: the Northern High Plains Aquifer, central United States, 1950–2010. *Hydrological Processes*, 30(14), 2533–2545. <https://doi.org/10.1002/hyp.10809>
24. Gandhi, G. M., Parthiban, S., Thummalu, N., & Christy, A. (2015). Ndvi: Vegetation Change Detection Using Remote Sensing and Gis – A Case Study of Vellore District. *Procedia Computer Science*, 57, 1199–1210. <https://doi.org/10.1016/J.PROCS.2015.07.415>

Chapter 2

New Algorithms for Groundwater Discharge Estimation for National Water Model Streamflow Forecasts

Minki Hong¹, Ritesh Karki², Joseph Krienert³, and Sama S. Memari⁴

¹ Texas A&M University; mkhong@tamu.edu

² Auburn University; rzk0045@auburn.edu

³ Southern Illinois University, Carbondale; jkrienert@siu.edu

⁴ University of Alabama; ssheikhmemari@crimson.ua.edu

Academic Advisors: Binayak P. Mohanty, Texas A&M University; Puneet Srivastava, Auburn University; Jonathan Remo & Liliana Lefticariu, Southern Illinois University, Carbondale; Prabhakar Clement, University of Alabama

Summer Institute Theme Advisors: Joseph Hughes, USGS, jdbughes@usgs.gov; David Steward, NDSU, david.steward@ndsu.edu

Abstract: Freshwater is essential to human civilization, and groundwater amounts to approximately 30% of the overall freshwater available on earth [1, 2]. The National Center for Atmospheric Research (NCAR) along with National Weather Information Service (NWIS) and National Oceanic and Atmospheric Administration (NOAA) released version 1.2.0 of the National Water Model (NWM) in October 2017. The NWM utilizes a conceptual (not physically-explicit) model for estimating groundwater discharge (baseflow) to streams, and this non-linear method only expresses a part of the interaction between groundwater and surficial hydrology. This research evaluates the current representation of groundwater in the NWM with a case study of two watersheds located within the Northern High Plains Aquifer (NHPA) region. A comparison between USGS observed streamflow and baseflow, and the NWM output showed that the NWM does a much better job of predicting streamflow and baseflow in the clayey catchment than in the sandy catchment. Based on the results of this analysis, Rorabaugh-Rutledge' and SWAT baseflow alternative functions were analyzed for representing the interactions between reservoir storage and baseflow. By comparing baseflow hydrographs derived from each alternative formulation, potential improvements to the NWM baseflow estimation were investigated. The results show that the magnitude and duration of NWM baseflow can be better correlated with observed baseflow with the use of the suggested alternative solutions.

1. Motivation

The National Water Model (NWM) is a hydrologic model that simulates and forecasts streamflow in 2.7 million reaches across the Continental United States (CONUS). However, the current version of the NWM (v1.2.0) represents baseflow discharge to streams using a conceptual bucket model characterized with a non-linear exponential function to determine groundwater storage and discharge. As there is no explicit representation of aquifer systems in the NWM, it is important to evaluate the NWM's ability to estimate groundwater discharge in different hydrogeological conditions. Baseflow

can be an important component of streamflow especially in reaches with strong stream-aquifer interaction. Utilization of a formulation that allows for only a one-way interaction between the subsurface reservoir storage and stream across the CONUS could also make it harder for the NWM to provide accurate baseflow estimations. It is important to understand the limitations of the simplified groundwater system used in the NWM to improve baseflow estimations and increase the reliability of the NWM forecasting in flood and drought scenarios.

2. Objectives and Scope

The objective of this research can be divided into two parts: 1) Evaluating the performance of the NWM in simulating baseflow and streamflow, and 2) Improving forecasting of the NWM using alternative formulations. Five study catchments were selected in the Northern High Plains Aquifer (NHPA) region to evaluate baseflow estimations. Simulated baseflow and streamflow from the NWM in the five catchments were compared with USGS observation data to evaluate the NWM results. Time series analysis of the NWM retrospective simulation output with the USGS observation dataset allowed for analyzing the weaknesses and strengths of the groundwater estimation scheme in the NWM. In addition, baseflow volume contribution from the groundwater reservoir of the NWM to total streamflow was also conducted to assess what portion of total streamflow is represented by baseflow. These investigations were used as a baseline for improving the NWM forecasting.

Improvements to baseflow estimations were based on adapting two alternative baseflow recession formulations that address the relationship between subsurface storage and groundwater discharge. To make it easy to represent the alternative formulations, a python module was created which can exactly represent the baseflow estimation scheme in the NWM. The alternative formulations have been represented in the python module and the performance of each formulations baseflow estimation at the watershed-scale were assessed. As part of the adaptation of new groundwater solution functions, a new partitioning methodology, which reduces the magnitude errors in deep percolation loss in the NWM output, was also included. By using alternative formulations, partitioning, and calibration, the possibility of forecasting improvement to the NWM groundwater module was evaluated.

3. Previous Studies

This is the first research evaluating the representation of groundwater baseflow in the NWM. Groundwater represents approximately 30% of the available freshwater on Earth [1, 2], and changes in groundwater reservoirs have a critical impact on the surficial water budget [3]. These aspects emphasize the importance of accurately representing groundwater systems in the NWM, so as to address the NWM goal of providing reliable short and long term hydrologic forecasts [4].

Groundwater baseflow is difficult to measure in the field, and for this reason many efforts have proposed a variety of methods to mathematically estimate baseflow, such as the historical analysis of a peak streamflow events' recession slope, or the explicit investigation into a groundwater reservoirs' physical parameters [5, 6]. Various approaches of characterizing groundwater baseflow were investigated and adapted into this research. Although these methods idealize the natural environment, they provide a conceptual basis for model representation and parameter adjustment, in order to best represent a natural groundwater-surface water flow system [7].

4. Methodology

4.1. Study Area

Five separate catchments located within the NHPA region with USGS gaging stations were selected to evaluate the current groundwater component of the NWM (**Figure 1**). Study catchments were selected within the NHPA region because of the strong interaction between the streams and aquifer which is important for evaluating the current and proposed formulations of the groundwater component in the NWM. Baseflow is a major component of the streamflow constituting up to 90 percent in North-Central Nebraska [8]. Precipitation in the NHPA ranges from about 406 mm in the west to about 787 mm in the east [9]. Grasslands and agriculture are the two major land use types covering about 56% and 36% of the region respectively. A contiguous area of sand deposits known as the Nebraska Sand Hills located in the central NHPA region is an important surficial feature for aquifer recharge. The eastern region of the NHPA is less permeable and has limited stream-aquifer interaction than in the central and the western regions due to the glacial deposits overlying the NHPA. The five catchments were selected such that it allowed for the groundwater component to be evaluated in regions of different soil and land use types and aquifer thickness. Further information about the five catchments are provided in **Table 1**.

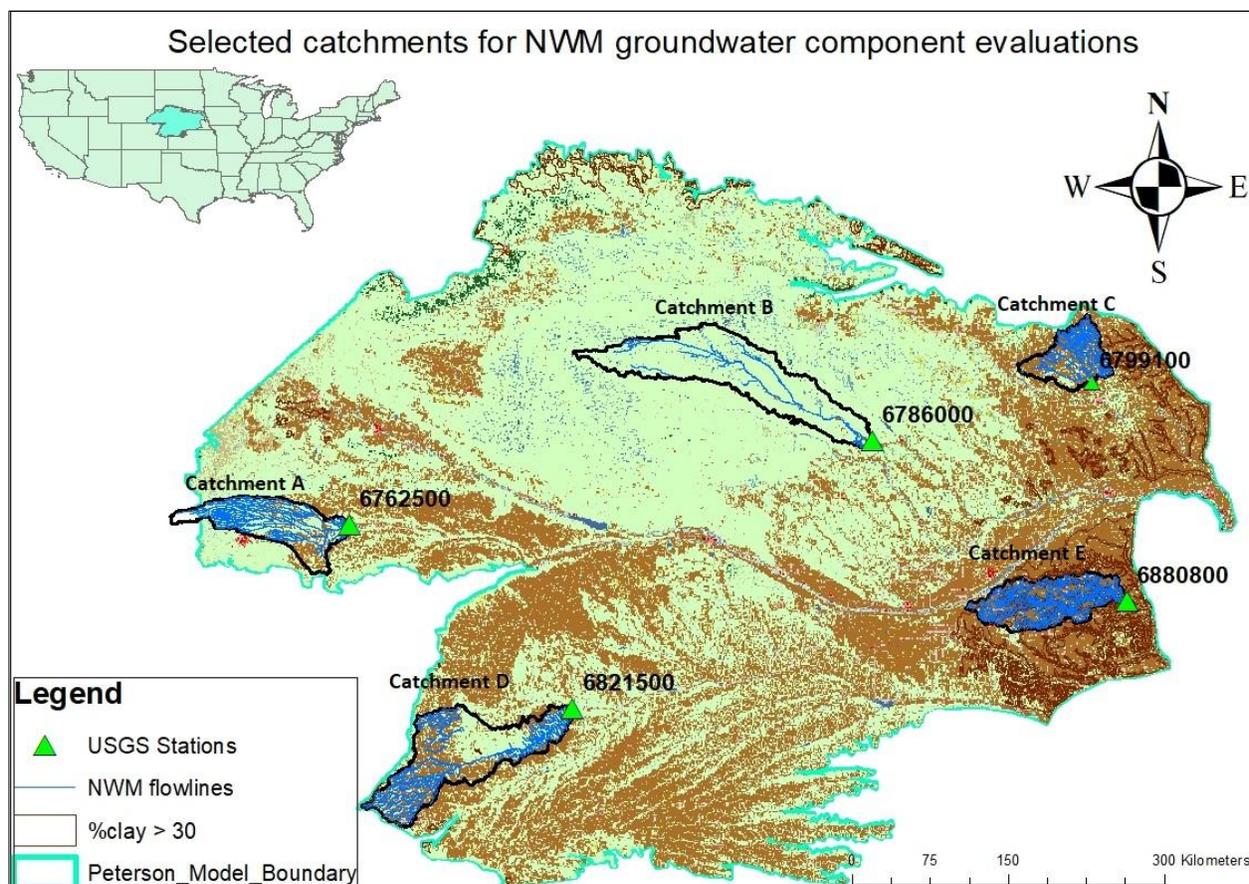


Figure 1. Study catchments with their corresponding USGS gaging station IDs in the Northern High Plains Aquifer (NHPA) region.

Table 1. Soil type, land-use, NWM feature ID, % clay, and land-use information of the five study catchments in the Northern High Plains Aquifer (NHPA) region.

Catchment	Area (km ²)	NWM Feature ID	USGS ID	%Clay	Terrain	Land use
A	3157	17457349	6762500	< 30%	High Plains	Pasture + Ag
B	6096	17321395	6786000	< 30%	Sand Hills	Pasture
C	1839	17274026	6799100	< 30%	Loess Hills	Ag
D	5636	11721499	6821500	< 30%	High Plains	Pasture + Ag
E	3125	18799565	6880800	> 30%	Loess & Glacial Drift	Ag

4.2. Groundwater in the National Water Model

The NWM implements a simplified groundwater representation using a conceptual groundwater reservoir “bucket” hereinafter referred to as Non-Linear Reservoir (NLR). The NLR transfers water to surficial hydrology based on the following exponential storage-discharge function [16]:

$$Q_{out} = C * \left(e^{exponent * \frac{z}{z_{max}}} - 1 \right) \quad (1)$$

Where, Q_{out} is the groundwater discharge for each time step, C is the unit coefficient, exponent is a calibration parameter, z is the water table in the bucket, and z_{max} is the NLR depth (bucket depth).

The current formulation used to solve water flux from the NLR has only a one-way coupling with the streams in that the NLR can lose water to the stream but cannot gain water from the streams through bed leakage. The only recharge to the NLR is through deep percolation from soil column, and the only outflow is through discharge at the headwater end of each stream reach. It is also important to note that the maximum NLR depth in the five study catchments is significantly less than the observed aquifer thickness throughout the study region (**Figure 2**).

4.3. Data Retrieval of USGS Observations and NWM Input/Outputs

Total stream flow data recorded at the outlet point for each of the five study areas was used in tandem with a suite of other environmental characterization datasets. These products were acquired and processed using the USGS Groundwater Toolbox (GWT). This graphical mapping interface includes several built-in hydrological analysis tools, including modules designed for separation of baseflow from observed streamflow [10]. One of the baseflow separation tools is the Base-Flow Index (BFI Standard and Modified) method. This method was determined the most efficient option based on available data in the five study areas of this research. The GWT software intuitively produces tabular results that were then translated into format for comparison with the simulated NWM datasets.

The NWM simulation outputs used for the study was from the 25-year retrospective simulation conducted for the CONUS. A significant portion of this research dealt with the acquisition, organization, and processing of NWM outputs inclusive to the spatial and temporal region

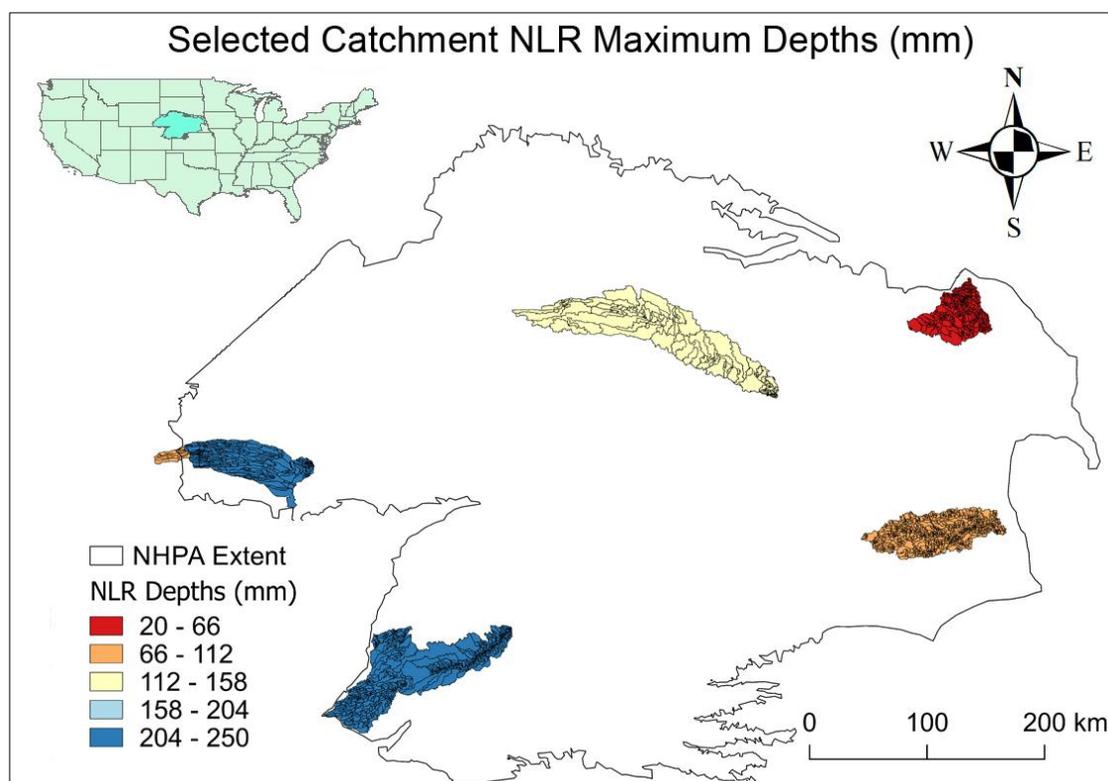


Figure 2. NLR depths for the stream reach catchments.

of interest within this research. The outputs were acquired in a NetCDF4 binary file format. The initial datasets were sourced from the NWM CHRTOUT module (channel routing output) from 2003 to 2009 [11, 12]. This seven-year period was selected for its 2007 to 2009 overlap with the simulated results of a publicly unreleased NWM version 1.2 subset region of the NHPA region [13]. In order to extract a small subset of data for the five study regions from these sizable data frames, approximately 3000 geospatially referenced stream reach polylines, and their respective catchment polygons provided by NCAR were used to define an aggregated search index based on featureID/ComID [4].

Both of these NWM datasets also match the spatial and temporal extent of a groundwater availability model (GWAM) produced for the NHPA using the USGS MODFLOW groundwater modeling software [9, 13]. The GWAM data utilized in this research helps validate the simulated NWM outputs, and observed USGS measurements, for a robust basis to develop alternate NLR functions.

4.4. Representing alternative formulations using Python module

The module that estimates baseflow in the NWM, which is named `module_GW_baseflow.F`, is a Fortran-based module. This Fortran-based module was translated to a Python-based module called `Baseflow_calc.py`. The purposes of developing this Python module are 1) to model the baseflow estimation scheme being used in the most recent version of NWM, and 2) to use it as a platform to evaluate the alternative formulations. The benefit of using the Python module is that it is possible to simulate baseflow without having to drive the entire NWM. Also, the selected equations, a modified Rorabaugh-Rutledge function, and an adaptation of the Soil and Water Assessment Tool's (SWAT) [5,15]. The Rorabaugh's and SWAT baseflow functions were easily represented in the Python module

and testing their effects on baseflow was simple and efficient. These formulations were used because they consider hydrogeological components of aquifer more explicitly than the NWM. The NWM dataset that includes deep percolation outputs from the soil columns was used as the input data for each alternative equation. It is important to note, the NWM is not currently accounting for an interflow component in subsurface routing processes. This suggests that the deep percolation, which is influx to the NLR, is usually overestimated. The apparent lack of interflow appears to cause baseflow magnitude discrepancies, as made evident by NWM total streamflow in sandy areas originating almost solely from baseflow. In order to consider the absent subsurface routed interflow reduction of deep percolation, the deep percolation was partitioned into two components: 1) inflow to the NLR calculated by multiplying deep percolation by BFI, and 2) interflow, calculated by multiplying deep percolation by an interflow index, or (1-BFI).

4.4.1. Rorabaugh-Rutledge

Rorabaugh formulated an equation to estimate groundwater reservoir baseflow into surface waters [5]. It is important to note that Rorabaugh's function assumes that there are only two components to the hydrograph that baseflow is separated from, groundwater discharge and surficial streamflow [7]. Rorabaugh's original function was further developed by Rutledge [6]. A variation of Rorabaugh and Rutledge's formulations was adapted for this work's NWM NLR evaluation, and is expressed by the governing Equation (2) below:

$$Q_{out} = \frac{1.866(\frac{Rt}{A})}{K} * e^{-\frac{0.933\pi^2 t}{4K}} \quad (2)$$

Where Q_{out} is groundwater discharge (m^3/s), R is recharge (m^3/s), t is time adjusted to the hourly datasets (3600 sec), A is catchment area (m^2), and K is as recession index of the time for groundwater flux to move through the conceptual NLR.

In order to accurately couple the value of K with the parameters for each NWM polyline stream reach available in the CHANPARAM.nc file, further development of this function is necessary. Currently our formulation does not include a dynamic K value per each stream reach. Instead an average K of 45 days (3.8e6 seconds) was sourced from the literature for groundwater environments similar to those expressed in the five study areas [3].

4.4.2. SWAT baseflow formulation

The Soil and Water Assessment Tool (SWAT) utilizes the following equation modified from Hooghoudt, Smedema and Rycroft to estimate groundwater flow into streams [14-16]:

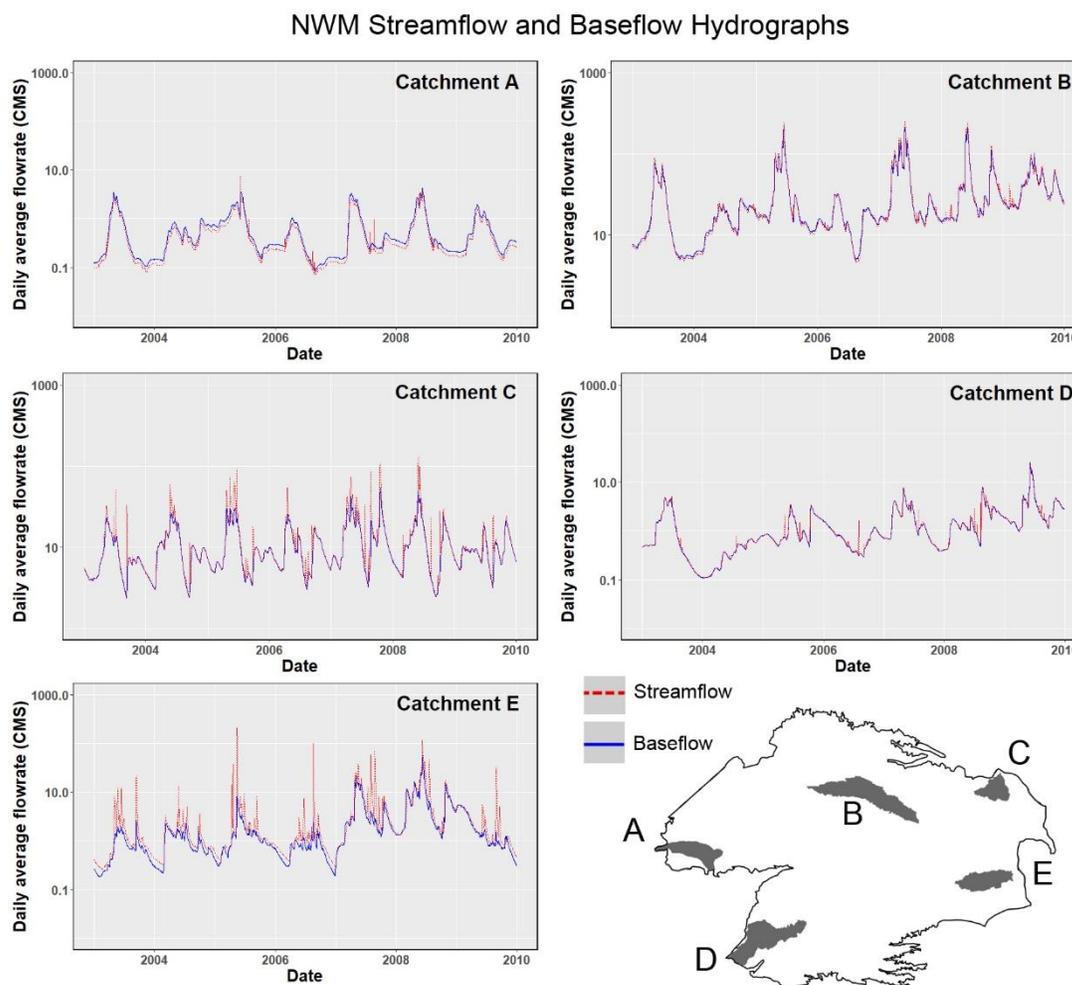
$$Q_{gw,i} = Q_{gw,i-1} * \exp^{-\alpha*\Delta t} + w_{rchrg,sh}(1 - \exp^{-\alpha*\Delta t}) \quad (3)$$

Where, $Q_{gw,i}$ is the groundwater flow into the channel on day i (mm), $Q_{gw,i-1}$ is the groundwater flow into the main channel on day $i-1$ (mm), α is the baseflow recession constant, Δt is the time step, $w_{rchrg,sh}$ is the amount of recharge entering the reservoir on day i (mm).

The baseflow recession constant, α , is a lumped parameter depending on the hydraulic conductivity, specific yield, and distance from the ridge or sub-basin divide for the groundwater system to the stream reach and an index for groundwater flow response to recharge [14, 16].

5. Results

Preliminary time series analysis of the NWM retrospective daily streamflow and baseflow outputs for each of the catchments showed two distinct trends. Catchments A, B, and D, which are the sandier catchments, had a similar trend in which baseflow contributed to almost all of the streamflow in these catchments (**Figure 3**). Catchment C and E, which had higher clay percent, followed a different trend and had overland flow as an important component to the total streamflow simulated. Hence, catchment B, which had the highest sand percent, and catchment E, which had the highest clay percent, were selected as representative catchments for the different soil types for detailed evaluation. It is also important to note that the NWM predicts streamflow from catchment A, yet the catchment USGS station has recorded zero streamflow for the last 2 years which shows the NWM's inability to simulate a losing stream.



5.1. Evaluation of NWM streamflow and baseflow estimations

Streamflow and baseflow plot of the NWM retrospective output for the sandy catchment B (**Figure 4a**) showed that baseflow accounted for about 97% of the total streamflow from 2003-2009. This indicates that almost all the water that reaches the land surface in the NWM percolates through the

homogeneous soil column of the land surface model (LSM) of the NWM and gets dumped to the stream as baseflow. It also demonstrates the distinct lack of overland flow generation by the NWM in the sandy soils of catchment B although overland/storm runoff was observed at the USGS station (Figure 4b). As a result, baseflow contribution to streamflow in the NWM is considerably higher than observed baseflow contribution (Figure 4b, 4d) with a R^2 of only 0.09 (Table 2). There is, however, almost no time lag in the water entering and leaving the groundwater bucket and almost all of the water entering the NLR leaves the NLR at the same time step (no storage) as demonstrated by the plot of water inflow and outflow for a reach in catchment B (Figure 5). As a result, the peak flows from the NWM during big storm runoff events is much higher than observed in the USGS station at catchment B (Figure 4c) resulting in a R^2 of 0.28 (Table 2). These results show that the model might not be able to do a good enough job of streamflow forecasting during drought periods because of the lack of storage and time lag in the groundwater component of the NWM.

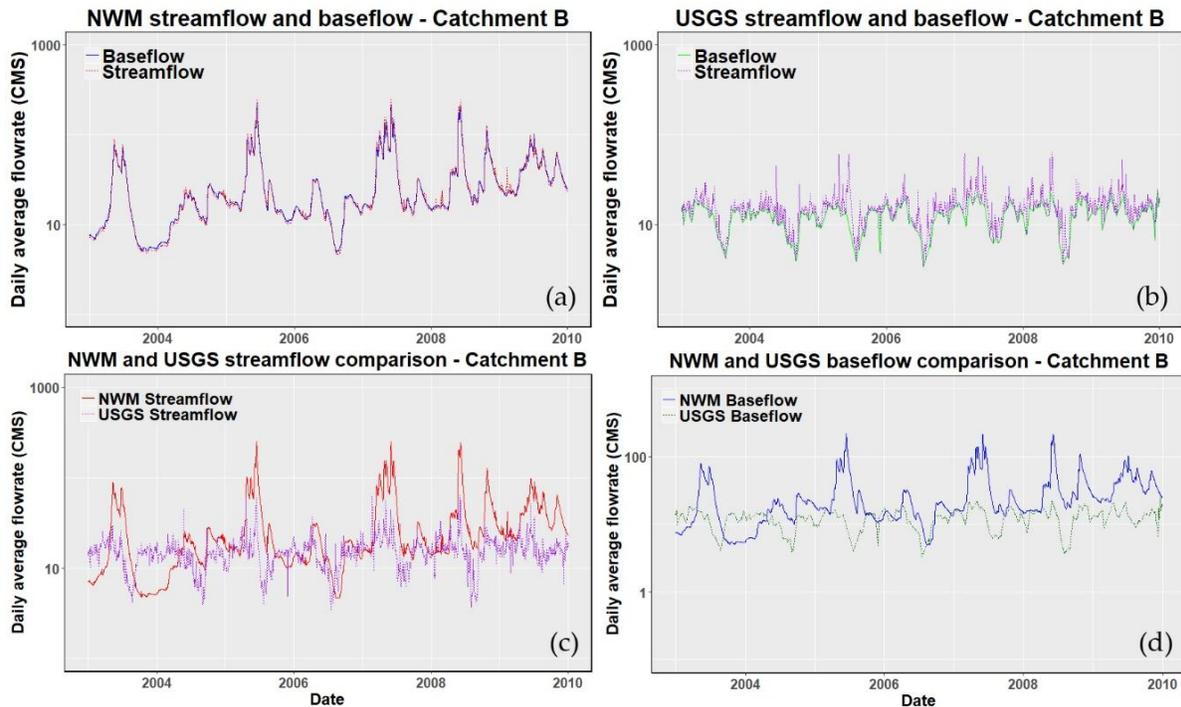


Figure 4. a) Streamflow and baseflow estimation by the NWM for catchment B from 2003 – 2009, b) Observed USGS streamflow and baseflow for catchment B for the same time period, c) Streamflow comparison between NWM and USGS, and d) baseflow comparison between NWM and USGS.

Table 2. Statistical results for streamflow and baseflow comparison for the evaluated catchments.

	Catchment B		Catchment E	
	Streamflow	Baseflow	Streamflow	Baseflow
R^2	0.28	0.09	0.43	0.41
RMSE (m^3/s)	33.07	32.71	7.81	3.73

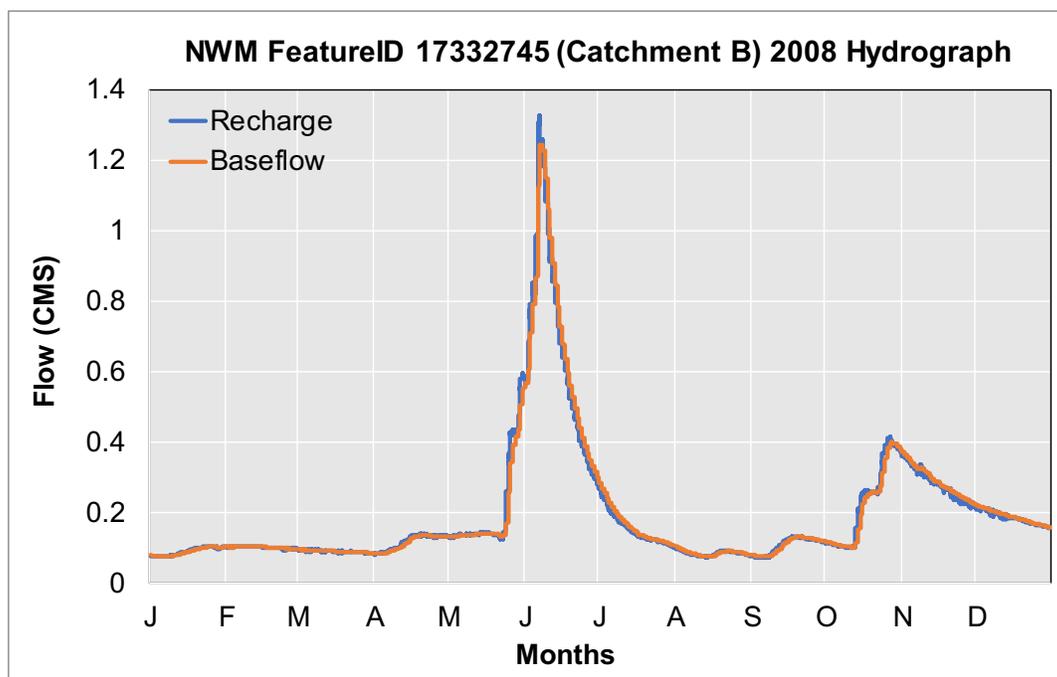


Figure 5. Comparison of groundwater recharge (bucket inflow) and groundwater discharge (bucket outflow) for each time step for reach “17332745” of catchment B.

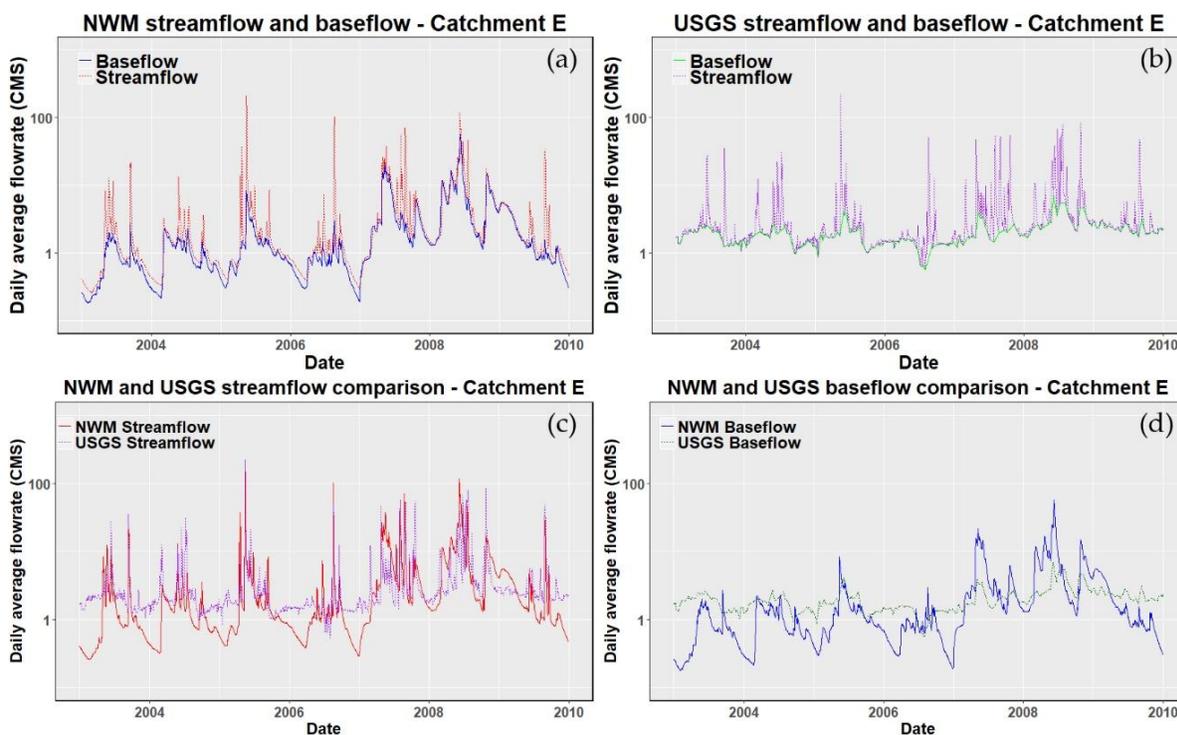


Figure 6. a) Streamflow and baseflow estimation by the NWM for catchment E from 2003 – 2009, b) Observed USGS streamflow and baseflow for catchment E for the same time period, c) Streamflow comparison between NWM output and USGS, and d) baseflow comparison between NWM output and USGS.

Streamflow and baseflow analysis of the NWM retrospective output for catchment E with USGS station observed data shows that the NWM does a much better job of simulating streamflow and baseflow (**Figure 6a, 6b**) in the clayey catchment with R^2 of 0.43 and 0.41 for streamflow and baseflow respectively (**Table 2**). Baseflow accounted for about 61% of the total streamflow in the NWM output from 2003-2009 closely matching the observed baseflow contribution of 49% to streamflow. Baseflow in catchment E, however, averaged less than $1 \text{ m}^3/\text{s}$ for most of the evaluation period and was considerably lower than in catchment B. The model was also able to simulate overland flow much better in catchment E when compared to the sandy catchment B, and, as a result, the NWM predicted streamflow was much more comparable to observed streamflow even during big storm events (**Figure 6c**). The model still greatly over predicted baseflow during big storm events (**Figure 6d**) which could again indicate to the lack of storage in the groundwater component of the NWM. The results show that the NWM does a good job of simulating streamflow and baseflow in clay regions where there is no prominent connection between the streams and aquifer and baseflow is not a major contributor to streamflow.

5.2. Evaluation of the NWM groundwater recharge

In the NWM, infiltrating water from precipitation that leaves the soil column is the only source of water entering the NLRs. This percolation of vertical drainage into the NLRs varies horizontally per each catchment area based on other simulated components of the NWM, such as climatic forcing, evapotranspiration, soil type, or land surface slope. The NWM recharge through time expresses interesting trends when comparing values from 2008 for all five study areas, with those derived from GWAM of the NHPA [9, 13]. **Figure 7 (a-e)** below express the two main findings, 1) the difference in mean peak recharge entering the NLRs within our study area exceed values from the GWAM of the NHPA as much as 3x (**Figure 7(c)**), suggesting that infiltration is not being accurately partitioned

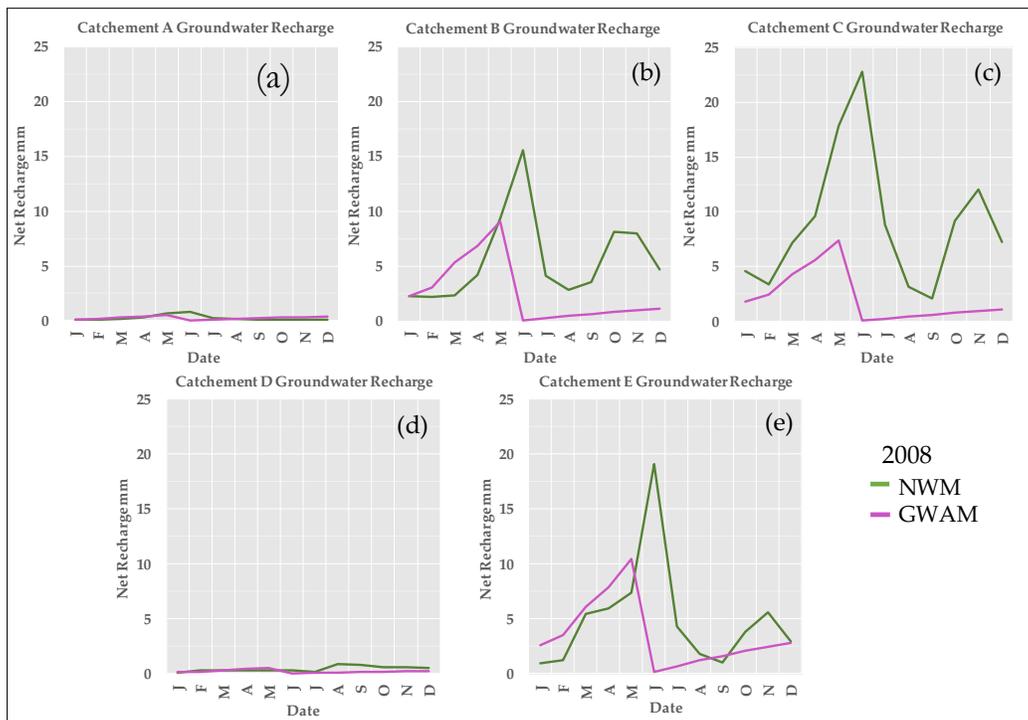


Figure 7. 2008 mean areal recharge comparisons between the NWM and the GWAM of the NHPA for five study areas.

between components of runoff/interflow and recharge within the soil columns. 2) The increasing trend in recharge magnitude from west to east across the study areas (**Figure 7 (a) -> (b) -> (c)**) suggests that the NWM distribution of precipitation generally agrees with the inputs used in the GWAM.

5.3. Current and alternate Non-Linear Reservoir discharge formulations

The effects of applying Partitioning method to deep percolation on baseflow hydrograph were evaluated graphically and statistically. Graphical evaluation showed that the errors in magnitude decreased for both catchments after the application of the method. Reduction of the maximum flux to the NLR for both catchments using the method resulted in better agreement with baseflow observations. The partitioning of the deep percolation, however, was not able to deal with the offset issue seen between simulated and observed baseflow as seen in the figure below (**Figure 8**). The effect of the partitioning method was also indicated in statistical parameters (**Table 3**). It was observed that the method can reduce the magnitude errors in baseflow hydrograph irrespective of soil type.

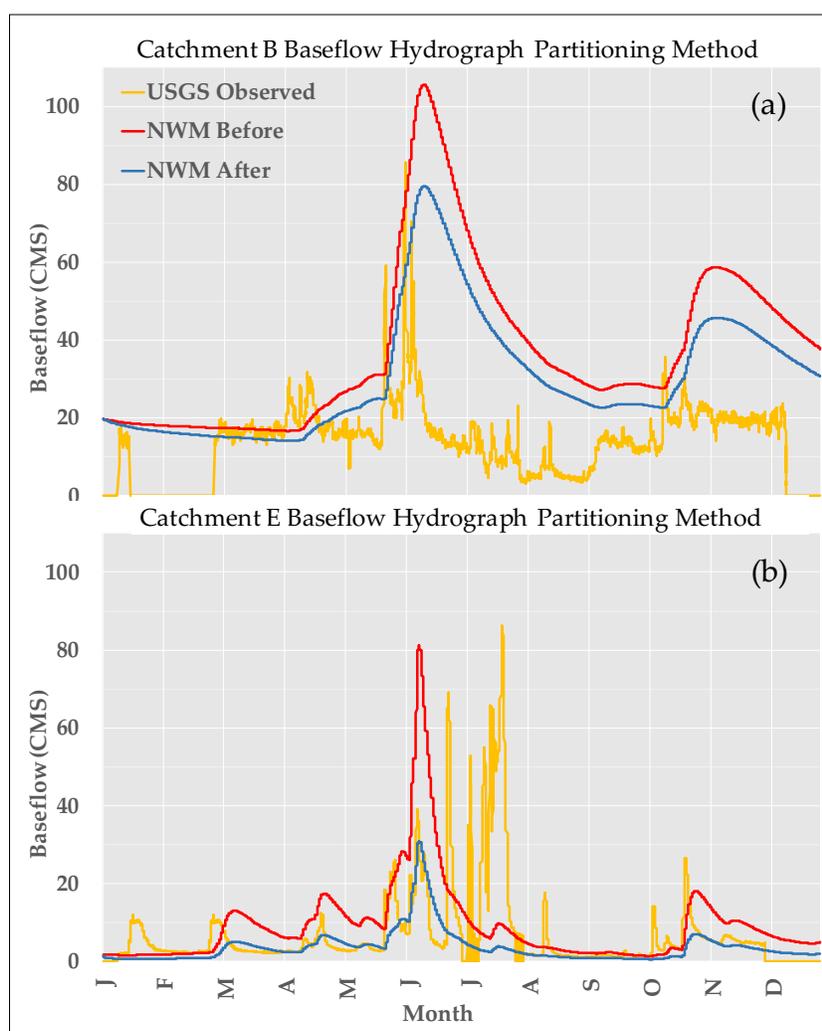


Figure 8. a) Improved baseflow hydrograph with partitioning method for the catchments B (BFI: 0.786) for the year 2008, b) Improved baseflow hydrograph with partitioning method for the catchments E (BFI: 0.391) for the year 2008.

Table 3. *Statistical analysis of baseflow hydrograph with the partitioning method.*

	Catchment B		Catchment E	
	Raw	Partitioned	Raw	Partitioned
R^2	0.14	0.12	0.11	0.11
RMSE (m^3/s)	31.53	23.12	13.44	11.45

Baseflow hydrographs derived from each alternative formulation were calibrated for the two catchments against the USGS hourly baseflow dataset. The calibrated baseflow hydrographs were evaluated based on statistical parameters. Baseflow hydrograph comparisons were conducted in reach-scale and catchment-scale. For catchment B, it was observed that the Rorabaugh's formulation performs much better than current NWM and alternate SWAT formulation. The RMSE value decreased by $15.72 m^3/s$ ($23.70m^3/s \rightarrow 7.98m^3/s$, **Table 4**) when the current NWM formulation was changed to Rorabaugh's. As more specific hydrogeological condition is considered, it was also observed that Rorabaugh's formulation was able to capture the peak timing well resulting in the improvement of R value ($0.12 \rightarrow 0.41$, **Table 5**). SWAT baseflow formulation seems to have good performance in lowering the peak magnitude of baseflow hydrograph but the time discrepancy between the observed and SWAT formulation simulated peak got worse when compared to the current and Rorabaugh's formulation. For catchment E, no noticeable differences were observed in the baseflow hydrographs between the current and alternate formulations which is also found in **Table 5**. The hydrographs derived from each of formulations appeared almost the same in terms of capturing peak timing and of adjusting magnitude.

Table 4. *Statistical evaluation of functions against baseflow separation data (catchment B).*

	NWM function	Rorabaugh's function	SWAT-modified function
R^2	0.12	0.41	0.07
RMSE (m^3/s)	23.70	7.98	22.02

Table 5. *Statistical evaluation of functions against baseflow separation data (catchment E).*

	NWM function	Rorabaugh's function	SWAT-modified function
R^2	0.09	0.12	0.10
RMSE (m^3/s)	12.21	11.96	12.20

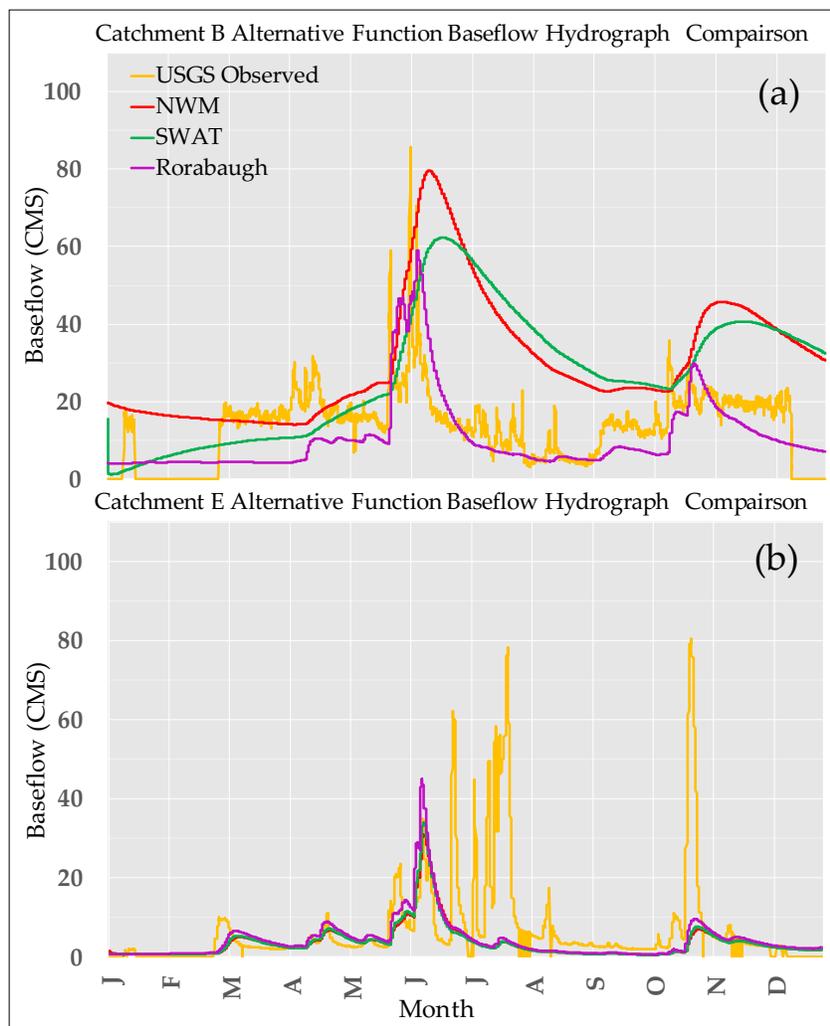


Figure 9. a) Comparison of baseflow hydrographs derived from each alternative formulation and baseflow observation separated from total streamflow using USGS groundwater toolbox (catchment B), b) Comparison of baseflow hydrographs derived from each of alternative formulations and baseflow observation separated from total streamflow using USGS groundwater toolbox (catchment E).

6. Conclusion

Preliminary results from the evaluation study shows that the NWM does a good job of estimating streamflow and baseflow in regions with clay soils where there is not a strong interaction between the stream and the underlying aquifer system. The model, however, struggles in predicting streamflow and baseflow in sandy regions, especially in big storm runoffs and drought conditions. This could be because of the model’s inability to generate overland flow in sandy soils resulting in all the precipitation that reaches the land surface to percolate to the NLR. The issue is compounded by the fact that the NLR, that is responsible for baseflow discharge, has near-zero storage or runoff delay because of which almost all of the water entering the NLR is discharge to the stream at the same time step. As a result, the model greatly overpredicts streamflow during big storm runoff events and possibly under predicts streamflow during drought conditions because of the lack of storage in the NWM. Further evaluation of the NWM streamflow prediction in drought conditions in regions where there is a strong connection between streamflow and the underlying aquifer is needed to generate a better understanding on NWM prediction in drought conditions. The NWM’s inability to generate adequate

overland flow in sandy regions was also demonstrated by the comparison of the NWM deep percolation loss to the GWAM for the NHPA which showed that the NWM greatly overpredicted groundwater recharge (deep percolation loss) in the central and eastern study catchments.

Application of the partitioning method to the deep percolation from the NWM to account for subsurface interflow provided a better fit between the simulated and observed baseflow datasets. Moreover, application of the Rorabaugh's function as an alternate formulation to estimate baseflow from the NLR provided a better fit when compared to the NWM current formulation and the SWAT formulation in catchment B. However, such effectiveness of using alternative formulations was not as definite in catchment E. This investigation into the groundwater component of the NWM shows that our approaches for improved hydrograph forecasting work better for sand vs clay soil types. Future work on alternative functions for the NLR could include similar evaluations as this work, in a variety of new physiographic provinces, and climatic trends than those observed in the NHPA study area.

Acknowledgements: We are grateful to the National Water Center and CUAHSI for providing us with this wonderful opportunity. Thank you to our theme leaders Joseph Hughes from USGS and David R. Steward from North Dakota State University for their time, suggestions, and guidance in this research. Support from Fernando Salas and David Blogett of the NWM was very valuable in the understanding and analysis of the NWM output files. We would also like to acknowledge the help of Paul Barlow, Ward Sanford, Bill Cuning, and Westley Zell from USGS for providing us with the trainings and suggestions that were critical to completion of this work. Finally, thank you to our course coordinators: Fernando Aristizabal, Lauren Grimley, and Danielle Tijerina for supporting us, being with us the whole way, and making this an unforgettable experience.

References

1. *Publications of the U.S. Geological Survey*; Washington, D.C., 1985; p 326.
2. Gleick, P. H., *Water in crisis: a guide to the worlds fresh water resources*. **1993**.
3. Brutsaert, W., Long-term groundwater storage trends estimated from streamflow records: Climatic perspective. *Water Resources Research* **2008**, *44* (2).
4. Gochis, D. J., M. Barlage, A. Dugger, K. FitzGerald, L. Karsten, M. McAllister, J. McCreight, J. Mills, A. RafieeiNasab, L. Read, K. Sampson, D. Yates, W. Yu, The NCAR WRF-Hydro Modeling System Technical Description. **2018**.
5. Rorabough, M., Estimating changes in bank storage and groundwater contribution to streamflow. *Int. Assoc. Sci. Hydro. Publ.* **1964**, *63*, 432-441.
6. Rutledge, A., *Model-estimated ground-water recharge and hydrograph of ground-water discharge to a stream*. US Department of the Interior, US Geological Survey: 1997.
7. Rutledge, A., The appropriate use of the Rorabaugh model to estimate ground water recharge. *Groundwater* **2005**, *43* (3), 292-293.
8. Jennifer S. Stanton, S. M. P., and Michael N. Fienen *Simulation of Groundwater Flow and Effects of Groundwater Irrigation on Stream Base Flow in the Elkhorn and Loup River Basins, Nebraska, 1895–2055—Phase Two*; 2010.
9. Steven M. Peterson, A. T. F., and Jonathan P. Traylor *Groundwater-Flow Model of the Northern High Plains Aquifer in Colorado, Kansas, Nebraska, South Dakota, and Wyoming*; 2016.
10. Barlow, P. M.; Cunningham, W. L.; Zhai, T.; Gray, M., *US Geological Survey groundwater toolbox, a graphical and mapping interface for analysis of hydrologic data (version 1.0): user guide for estimation of base flow, runoff, and groundwater recharge from streamflow data*. US Department of the Interior, US Geological Survey: 2015.

11. Repository of NWM 25 Year Retrospective Data. Amazon, E. C. Amazon web services (AWS);, 2018.
12. Blogett, D., Personal communication and transfer of NWM, and GWAM datasets. **2018**.
13. Hughes, J., Personal communication and transfer of NWM, and GWAM datasets. National Water Center, Tuscaloosa, AL.
14. Hooghoudt, S., General consideration of the problem of field drainage by parallel drains, ditches, watercourses, and channels. *Contribution to the Knowledge of Some Physical Parameters of the Soil* **1940**, 7.
15. S.L. Neitsch, J. G. A., J.R. Kiniry, J.R. Williams, Soil and Water Assessment Tool. **2009**.
16. Smedema, L. K. R., David W, *Land drainage: planning and design of agricultural drainage systems*. Cornell University Press, Ithaca, N.Y. 1983.

Chapter 3

Effects of Spatial Resolution on a Distributed Hydrologic Model through Dynamical Forcings: Flood Extent and Depth in Low Gradient Watersheds

Deepa Gurung¹, Andrew Goenner², Francisco Perez³, and Tasnuva Rouf⁴

¹ University of Alabama; dgurung1@crimson.ua.edu

² Iowa State University; agoenner@iastate.edu

³ Michigan Technological University; fperez@mtu.edu

⁴ George Mason University; trouf@maonline.gmu.edu

Academic Advisors: Dr. Sarah Praskievicz, University of North Carolina; Dr. Kristie Franz, Iowa State University; Dr. William Gallus, Iowa State University; Dr. John Gierke, Michigan Technological University; Viviana Maggioni, George Mason University

Summer Institute Theme Advisors: Dr. Jude Benavides, University of Texas Rio Grande Valley, jude.benavides@utrgv.edu; Dr. Fred Ogden, National Water Center, fred.ogden@noaa.gov; Dr. Sarah Praskievicz, University of North Carolina, sjpraski@uncg.edu

Abstract: Floods are the most frequent natural disasters that occur in the United States, annually costing approximately a hundred lives and billions of dollars [1]. NOAA's National Water Model is a hydrologic model that simulates real time and forecasted streamflow along with other hydrologic information for 2.7 million river reaches across the contiguous United States. This model has proven beneficial in forecasting when and where flooding can be expected as its terrain routing module operates at a spatial resolution of 250 m. However, in topographically complex regions and areas that experience precipitation induced flooding, a finer resolution is necessary to better predict flood extent and depth. This study aims to determine the difference between two different grid resolutions on the low gradient topography of Brownsville, Texas by using a 2D physics-based hydrologic model, Gridded Surface/Subsurface Hydrologic Analysis (GSSHA). For these comparisons, a 50 m grid was used as the hyper-resolution simulation and a 250 m grid was used to emulate the NWM. Multiple rainfall intensities were chosen as different atmospheric forcings. Results indicated that the 250 m resolution estimated larger inundation extent, whereas the 50 m estimated larger overland depth values and was able to better identify potentially inundated areas. Linear regression showed a statistically significant correlation between increasing flood threshold and specificity between the model. The increasing threshold with increasing between flood threshold during each rain events. Therefore, the estimates for inundation area and water depth in hyper-resolution modeling could be more accurate than those of the NWM during each rain events.

1. Motivation

“Floods are the most chronic and costly natural hazard in the United States, causing an average of 140 fatalities and \$5 billion damage each year” [2]. In an effort to better predict and mitigate damage caused by floods, hydrologic models have continuously been developed and enhanced. NOAA's National Water Model (NWM) simulates observed and forecast streamflow for 2.7 million river reaches across the continental U.S. at a grid resolution of 250 meters. While this is considered a high

spatial resolution, many rivers and other hydrologic features tend to be on the order of tens of meters. Additionally, areas that contain minimal topographical changes and/or experience precipitation induced flooding, are more complicated to accurately model at this resolution. As urbanization increases throughout the country, especially in coastal regions, more people and property will become vulnerable to the effects of localized flooding. Therefore, it is essential for existing models to be improved upon in order to properly predict and assess inundation extent and depth in low gradient areas. Unfortunately, increases in resolution are accompanied by an exponential growth in computing time and power. To provide sufficient lead time to residents with the most accurate data possible, this project aims to determine when and how to necessitate a hyper-resolution model (<100 m).

2. Objectives and Scope

2.1. Objectives

The intention behind this research is to discover if there is a significant difference in flood inundation extent and depth between model outputs with grid resolutions at 250 meters and 50 meters in topographically complex, low gradient watersheds. To do this comparison, a 250 m grid resolution was chosen to emulate the resolution employed by the NWM, and a 50 m resolution was determined to be sufficient for the “hyper-resolution” model. Variations in rainfall intensity were incorporated in the model by using the NOAA Atlas 14 precipitation frequency data. The Gridded Surface/Subsurface Hydrologic Analysis (GSSHA) model prepared in the Watershed Modeling System (WMS) application was used to analyze these objectives.

2.2. Case Study

Brownsville, Texas was chosen as the area of interest for this research. Brownsville is a city located on the border between southern Texas and Mexico that is characterized by flat slopes and clay-rich soils which make the region especially susceptible to flooding [3]. In addition to the effects of topography and soil type, urban development is rapidly increasing in the area therefore causing drainage density to increase while also lowering soil permeability [3]. Although the NWM NHDPlus stream network is able to delineate larger drainage networks in the area, such as the Rio Grande River (RGR), Brownsville is unique due to small scale, hydrologically significant features called *resacas*. *Resacas* are old, secondary river channels of the Rio Grande that once transported excess water away from the river during times of high flow but have since been disconnected from the RGR. As flow was reduced by the construction of upstream dams and reservoirs, the *resaca* features remained and now play a significant role in drainage and overland flow mechanisms during large precipitation events.

For this study, the watershed incorporating Town Resaca and Resaca de la Guerra was analyzed (**Figure 1**). While they are not directly connected to the Rio Grande River, they can still contain water due to rainfall events or pumping from the RGR. They can also be used to move floodwaters during storm induced flooding and/or major flooding on the RGR. These *resacas* are sensitive to precipitation events and can cause concerns about localized flooding. In addition, Brownsville is located near the Gulf of Mexico, and so elevational change over the watershed is minimal with a maximum elevation of 11 m in the Northwest corner of the watershed to sea level in the Northeast corner.

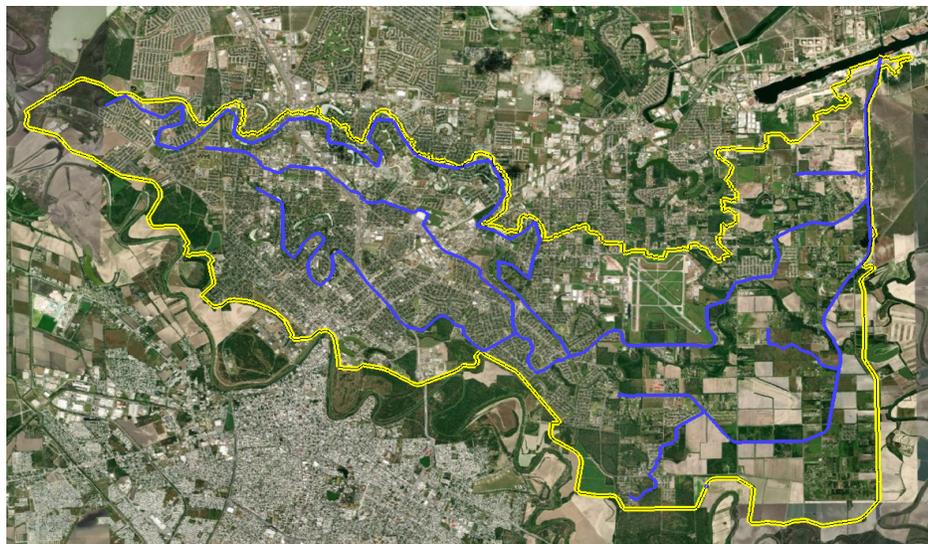


Figure 1. Watershed boundary (in yellow) of the Town Resaca and Resaca de la Guerra in Brownsville, TX. The western part of the watershed contains the more urbanized section of Brownsville, TX. The northeast corner contains the outlet of the resacas into the Brownsville shipping channel. Streamlines are designated in blue.

3. Previous Studies

There have been numerous reports and past studies evaluating the *resaca* networks and general nature of flooding in Brownsville. One study, conducted by Whitko, focused on the accuracy of flood modeling in the *resacas* utilizing high resolution LIDAR data and potential land use changes incorporated into a HEC-RES/HMS distributive model [4]. This study found that, in general, the simulated rainfall year floods generated a higher flood peak than expected, and that land use change through future urbanization would increase flood inundation. Another study, submitted to the city of Brownsville by Rust, Lichliter, and Jameson, analyzed the causes of flooding and reaction of the *resaca* and drainage structure networks to heavy rainfall events [5]. Several solutions were proposed that included increasing the size of the *resacas* through dredging and improving the hydraulic drainage structures.

The number of studies that examined how changes in model grid resolution and the temporal resolution of precipitation events affected flash flooding have been scarcer, especially on the hyper-resolution scale. Rafieeinasab et al., that compared the peak flow of a hydrograph in an urbanized landscape between resolutions [6]. The spatial resolutions compared in the study started at a resolution of 2 km and decreased to 250 m. The authors determined that a 500 m resolution was the best choice due to statistically insignificant increases in accuracy from 500 m to 250 m. Our study hopes to expand upon the aforementioned work in order to determine the effects of an even finer grid resolution on inundation extent prediction. It also has the ability to analyze factors such as time to flooding and flood duration times so that flood hazards can be better converted to flood risks.

4. Methodology

4.1. Terrain Pre-processing, TOPAZ, and Grid Creation

The primary input files for a GSSHA rainfall/runoff simulation are the 2-D finite difference grid and accompanying surface elevations. The grid is a rectangular area that covers the extents of the watershed. 1-meter LIDAR data were obtained from the International Boundary and Water

Commission (IBWC) in 2011. The various files were merged into a single digital elevation model (DEM) that was resampled to 4 m in order to increase computational efficiency. This DEM was initially run through the Topographic Parameterization Program (TOPAZ), contained in WMS, to delineate the watershed boundary and stream network. However, due to the study area's location, the topography is relatively flat. Additionally, several underground drainage structures exist that route water to an outlet point towards the Northeast. Predetermined flowlines incorporating proper drainage features from previous field studies were artificially burned into the 4 m DEM using the ArcHydro tool found in ArcMap 10.6. Stream burning is a common flow enforcement technique used to correct surface drainage patterns derived from digital elevation models (DEM). The technique involves adjusting the elevations of grid cells that are coincident with the features of a vector hydrography layer.

Once burned, the DEM was imported to WMS and resampled again from 4 m to 6 m as this resolution significantly lowered computational time and power when running TOPAZ. After the watershed and streams were delineated, 2 different grids were created. In order to represent the resolution employed by the NWM, a 250 m grid cell size was chosen. A 50 m cell size was utilized for the "hyper-resolution" aspect after several model runs indicated resolutions greater than 50 m significantly increased computational time and power, but did not result in much output variability compared to a 30 m resolution. Hydraulic structures were then added to the model. Due to time constraints, all 142 structures found in this watershed were unable to be incorporated; therefore, only weirs and culverts, structures that were deemed the most hydraulically important, were added to the model. **Figure 2** details the locations of important hydraulic structures and delineated hydrological features.

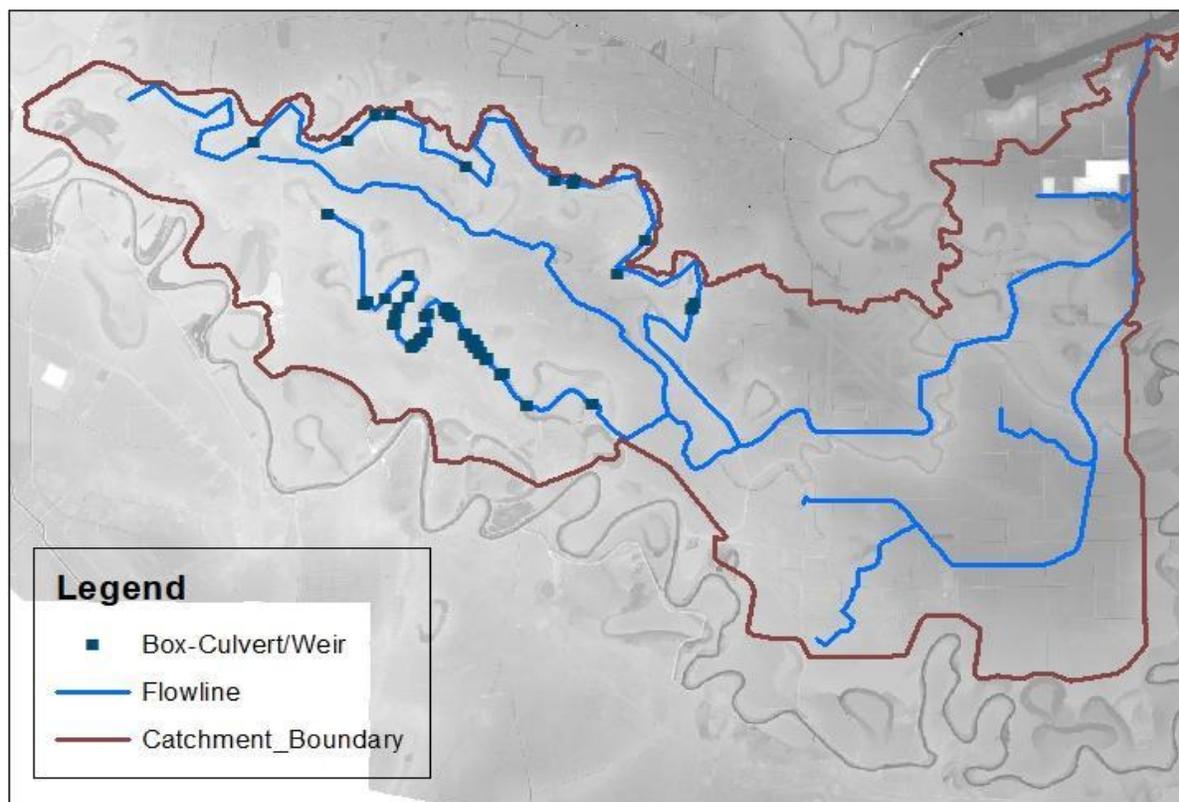


Figure 2. 4-m LIDAR based elevation map showing the location of box-culverts and weirs along the stream channels in the watershed

4.2. Atmospheric Forcing

To enable a more efficient model and minimize complications, the effects of evaporation and evapotranspiration were ignored as the average daily evapotranspiration rate for Brownsville was calculated to be 1.41 mm, inconsequentially small compared to the rainfall values [7]. With this simplification, the only atmospheric forcing driving the model was precipitation. In recent years, the National Oceanic and Atmospheric Administration have continuously updated the precipitation frequency estimates for the U.S. by using previous values found in technical memorandum NWS Hydro 35 (1997). NOAA has been updating these values for various regions of the U.S. since the early 2000s and publishing them as Atlas 14 data. NOAA is currently working on Texas, but results won't be published until the Fall of 2018. Therefore, a request for assistance was made to the National Water Center for this data. We were provided with data from Michael St. Laurent at UCAR CPAESS for precipitation frequency intervals of 2, 25, and 100 years at 6-hour (high intensity) and 24-hour durations (low intensity) (**Table 1**). The area of Brownsville fell under the Type III synthetic rain distribution [8]. The GSSHA model contains a built in Type III 24-hour distribution curve that was used for the 24-hour runs with storm frequency values converted into millimeters. The 24-hour curve was adjusted to fit a 6-hour rain event by taking every fourth time step from the 24-hour type III curve.

Table 1. *The Rainfall amounts associated with Storm Frequency*

Storm Frequency	24-h Rainfall Amount (mm)
2-year	107.19
25-year	223.50
100-year	308.36

4.3. Land Use and Soil Data

Land use data were obtained from the National Land Cover Dataset (NLCD). Soil type and texture data were obtained from Natural Resource Conservation Services (NRCS) Soil Survey Geographic Database (SSURGO). Tools in Arcmap 10.6 were used to process the data and prepare GSSHA input files describing the physical characteristics of the watershed. Green and Ampt parameters of infiltration were based on information from NRCS soil classification and land use/cover classification. Infiltration rates are affected by land use, especially impervious areas due to urbanization. Soil types associated with these land use types are referred to as developed soil texture classifications. In order to analyze flooded conditions, initial soil moistures were set up to be 95% saturated. **Figure 3** provides an overview of the complete model setup.

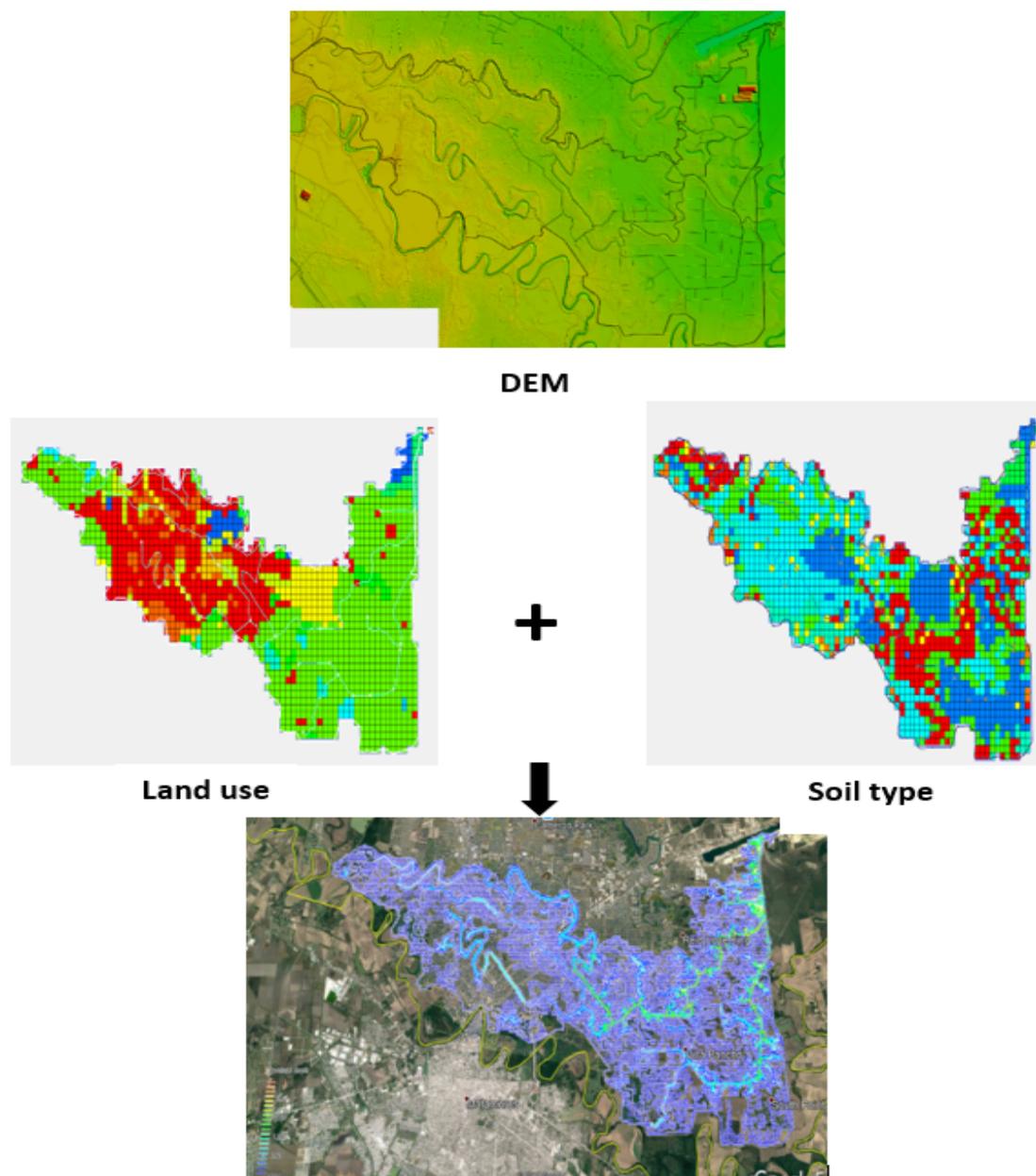


Figure 3: An overview of complete model setup. At the top, is the DEM (LIDAR data) imported into GSSHA used to determine elevation, watershed and stream channels. This was followed by importing Land Use and Soil type to determine soil properties. When these data sets were combined with atmospheric forcings, we had the necessary information to run GSSHA model (bottom).

4.5. Confusion matrix

Model comparison was done by creating a confusion matrix. 50 meter resolution was considered as a reference image for all three rain events. The matrix compared the 50 m inundation pattern with that of the 250 m resolution at different flood depth thresholds. **Table 3** represents the comparison done at a flood depth threshold of 1 meter.

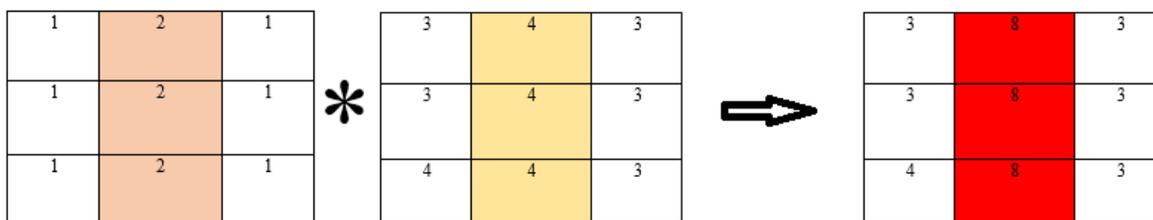


Figure 4: Concept illustration of model results through the creation of a confusion matrix comparing inundated pixels under two grid resolutions

This measure is equal to 100% when the two areas coincide. To perform these calculations, flood inundation rasters were produced from ASCII files obtained from model simulations. The boundary polygon derived from the WMS delineated watershed was used to extract the floodplain boundary area from each of the inundation rasters. All 250 m resolution rasters were resampled to 50 m grid. All rasters were classified using binary inundated/non-inundated classification for each pixel. With raster data classified at an identical resolution, true positive and false negative percentage were calculated. This comparison between two resolution raster grids were done for each rain intensity tested in the study.

4.6. Regression model to analyze the relation between test statistics at different flood threshold

Confusion matrix at flood depth thresholds of 0.1 m, 0.5 m, 1 m, and 1.5 m were created. A linear regression model was done to analyze whether the test statistics obtained change with flood thresholds. The two test statistics considered in this case study were Specificity and Sensitivity. Specificity is, by definition, how often the 250 m model correctly predicted non-inundated pixels when these pixels were actually non-inundated in the reference image (50 m resolution). Sensitivity, by definition, is how often the 250 m model correctly predicted inundated pixels when the same pixels were actually inundated in the reference image (50 m resolution).

5. Results

5.1. Comparison of inundation extent and overland flow depth by visual interpretation

In addition to the difference in inundated pixels, simulations at different resolutions also showed differences in water depth for identical rain intensity. For example, **Figure 5** shows the overland water depth and inundation extent near the Town Resaca at the same location and time step for a 2-year, 24-hour rainfall event; the top image is at a 250 m resolution and the one below shows a 50 m resolution. The two locations marked in red are topographic depressions in the area. It is apparent that the 250 m resolution run was unable to identify these locations as potentially inundated areas as compared to the 50 m resolution run. Conversely, the 250 m resolution overestimates inundation extent along the actual Town Resaca as evidenced by the outline of the individual grid cells visible in the top photo. Although the actual stream width is smaller than 250 m, it seems that the size of the individual grid cells caused inundation to extend further than necessary.

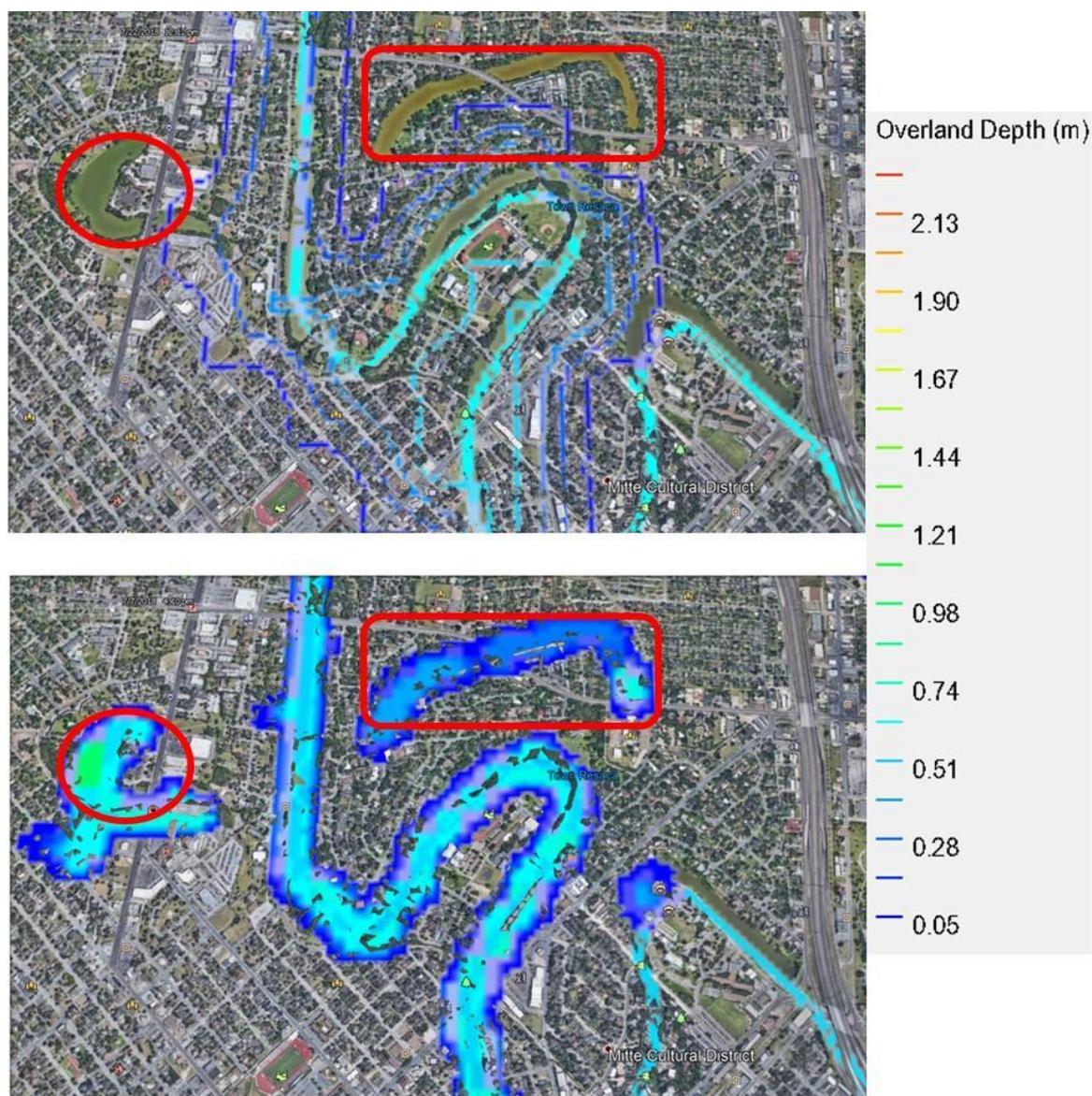


Figure 5. Comparison of inundation extent between 250 m Resolution Run over Town Resaca (Top) and 50 m Resolution Run over Town Resaca (Bottom) on a 2-yr, 24-hr rainfall event

Figure 6 shows the overland water depth near the outlet; the top image shows the 250 m resolution and the bottom image shows the 50 m resolution at the same location and time step for a 100-year, 24-hour precipitation event. The two locations marked in red show different flood inundation depths. The bottom image (50 m) shows severe overland depth, reaching values of approximately 3.5 m. However, the top image (250 m) displays an estimated maximum of 2.5 m for overland depth. Similar to the discussion above, the 250 m resolution simulates a larger flood extent, but lower overland depth (i.e. highest overland depth is 2.83m for 250 m resolution and highest overland depth is 3.91m for 50 m resolution).

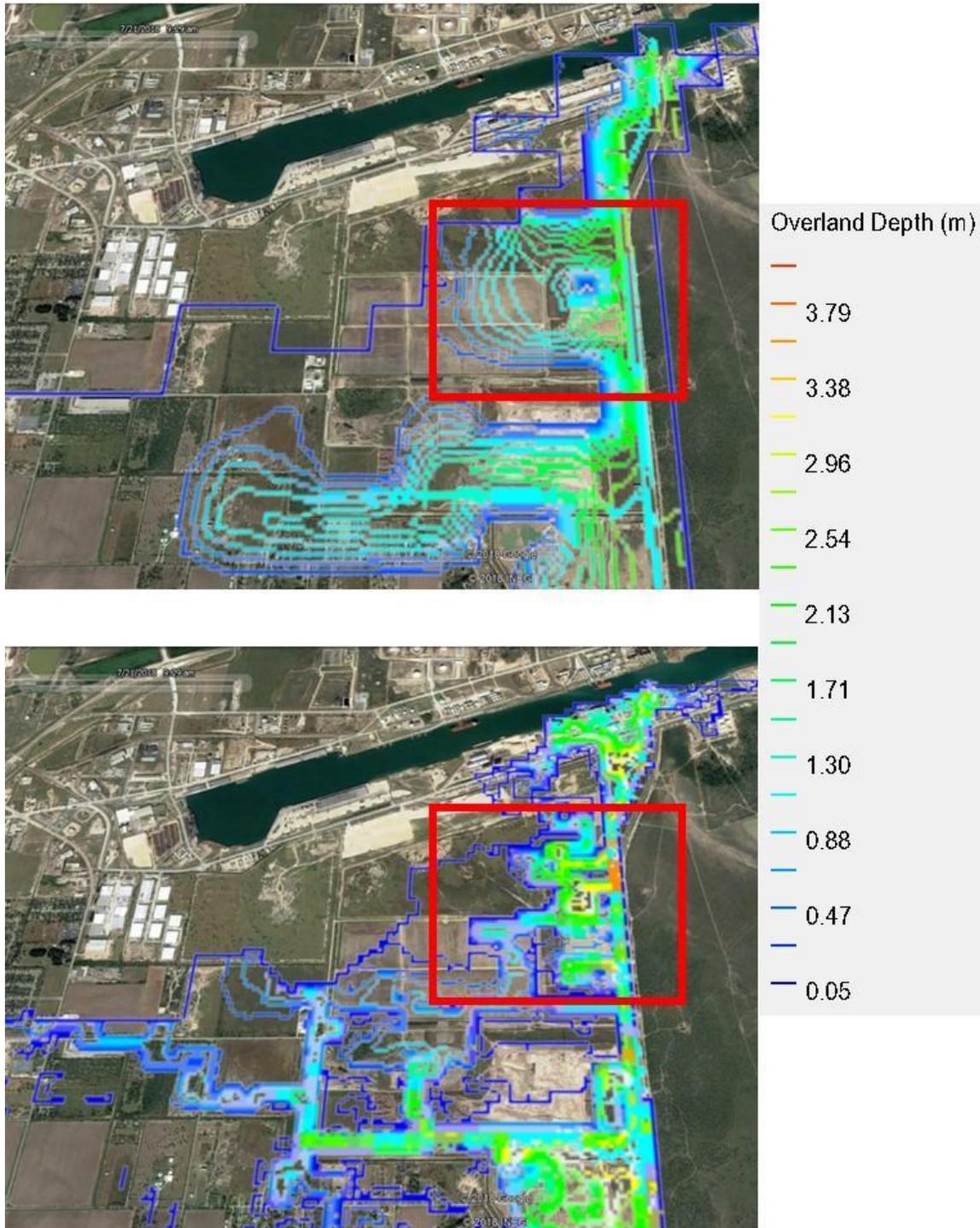


Figure 6. Comparison of overland depth between 250 m Resolution Run over Town Resaca (Top) to 50 m Resolution Run over Town Resaca (Bottom) on a 100-yr, 24-hr rainfall event

From the visual interpretation, it is obvious that for the 250 m resolution, inundation extent is more widespread than the 50 m meter resolution, perhaps even overestimating at times. As a precautionary measure, the overestimated extent may be preferable as it encompasses a larger area; however, it may also cause false alarms for nearby residents and affect decisions made by various stakeholders.

5.2. Comparison of Inundation extent by using Confusion Matrix

The results of the pixel-based inundation comparison are presented for three different rain events as a confusion matrix in **Table 2**. The confusion matrices present a binary comparison of flood inundation maps from the 250 m models against the 50 m models that list the number of True Positive, False Positive, False Negative and True Negative. It is important to note that the prevalence of inundated pixels between the two model runs are 4% for 2-year; 10.36% for 25-year rain and 12.85% for 100-year rain event. For 2 year, commonly inundated pixels were 1.33%; for 25-year, it was 4.11 and for 100 year, it was 5.63%. These results indicate that with larger rain events, the percentage of agreement (i.e. commonly inundated pixels) between the two models increases. This could be because with more rain, larger volumes of water would cause more individual pixels to be inundated in both of the resolutions.

Table 2. Confusion matrix demonstrating the comparison between 250 m and 50 m resolution.

			50-meter resolution model					
			Non-inundated		Inundated		Total	
			Pixel	Percentage	Pixel	Percentage	Pixel	Percentage
250-meter resolution model	2 year	Non-inundated	38130	92.32	1103	2.67	39233	95.52
		Inundated	1519	3.68	548	1.33	2067	5.03
		Total	39649	96.00	1651	4.00	41300	100
	25 year	Non-inundated	33622	81.41	2582	6.25	36204	88.15
		Inundated	3399	8.23	1697	4.11	5096	12.41
		Total	37021	89.64	4279	10.36	41300	100
	100 years	Non-inundated	30830	74.65	2983	7.22	33813	82.33
		Inundated	5162	12.50	2325	5.63	7487	18.23
		Total	35992	87.15	5308	12.85	41300	100

Furthermore, with an increase in rain intensity, the percentage of overprediction also increases. For 2 year, it was 3.68%; for 25 year it was 8.23%; and for 100 year, it was 12.5%. This is expected and supports the previous results showing that the 250 m resolution overpredicts the flooded pixels as seen in the 25-year flood where the 250 m model run contains 5096 inundated pixels compared to 4179 on the 50 m.

Linear regression modeling was performed to examine if the test statistics between the two models varied with the flood thresholds. Significant R-square values in Specificity indicates that with increasing thresholds the chances of 250 m to overpredict the inundated pixels increases. Therefore, it would be advisable to run a hyper-resolution modeling during severe water depths.

Insignificant R-square in regression model for sensitivity at different flood threshold indicates the uncertainty of sensitivity (or chances that 250m could predict it the inundated pixels correctly) with increasing flood thresholds, during each rain events. In other words, with increasing flood thresholds, there exist no relation that could project whether the sensitivity would increase or decrease. Had there been more sets of data, a possible statistically significant relationship could be established. Or it is also possible that there lies no relation between flood threshold and sensitivity at all. Albeit, with currently available datasets (three flood thresholds), the relationship seems insignificant.

Comparing both the test statistics, it can be inferred that chances that 250m could overpredict might rise high (due to strong coefficient of determination) while, there is no implied evidence (due to lower coefficient of determination) that it could predict correctly. Therefore, it can be concluded that hyper resolution modeling is effective in predicting the critical water depths during each rain events because, the 250 m resolution have higher chances of overprediction as well as false predictions.

Table 3. Linear regression model to analyze how test statistics produced vary with the flood thresholds

		R-square	p			R-square	p
Specificity	2-year	0.82	0.09	Sensitivity	2-year	0.36	0.3
	25-year	0.94	0.02		25-year	0.76	0.1
	100-year	0.91	0.04		100-year	0.006	0.9

5.3. Comparison of Hydrograph

When comparing the hydrographs at the outlet point, the difference in resolution of the models was apparent. As seen in **Figure 7**, the peak discharge for the 50 meter resolution was nearly double that of the 250 m discharge. Three likely causes exist for the major increase in discharge. First, the coarser resolution (250 m), water had to travel further to reach a channel as the grid cell size affected spatial extent. This would allow for a lower peak discharge that would then steadily decrease, as evidenced in the figure. Additionally, as water remained on the land for a longer period of time, it could better infiltrate the soil. The amount of infiltration water on the 25-year, 24-hour, 250 m grid was $6.27 \times 10^{-3} \text{ km}^3$ compared to $5.58 \times 10^{-3} \text{ km}^3$ on the 50 m grid at the same flood event and rain duration, thereby lowering the peak of the hydrograph. This pattern was evident on each pair of different resolution model runs. Second, the larger, individual grid cell size allowed more water to move back onto the land when flood inundation occurred, thus reducing the movement and amount of water draining towards the outlet. Finally, we hypothesize that the lower peak discharge may be a result of changes

in the discretization and averaging of land use/soil parameters at higher resolutions. Both the NLCD and SSURGO data have a resolution of 30 meters. When these parameters are combined onto a grid, unique ID values are generated for every combination of soil type and land use, thereby affecting individual grid cell hydraulic conductivity values. When these parameters were combined on a 250 m grid, 69 different values were generated; however, the 50 m grid generated 83 values. When a spatial averaging was performed on hydraulic conductivity between 50m and 250 m with values of 0.157 cm/hr and 0.169 cm/hr were found respectively. This result strengthens the conclusion that more water was allowed to penetrate the soil and lessen the discharge at the outlet. It should be noted that these hydrographs assume that the streams are empty at the beginning of the rainfall which would lessen the peak flow discharge.

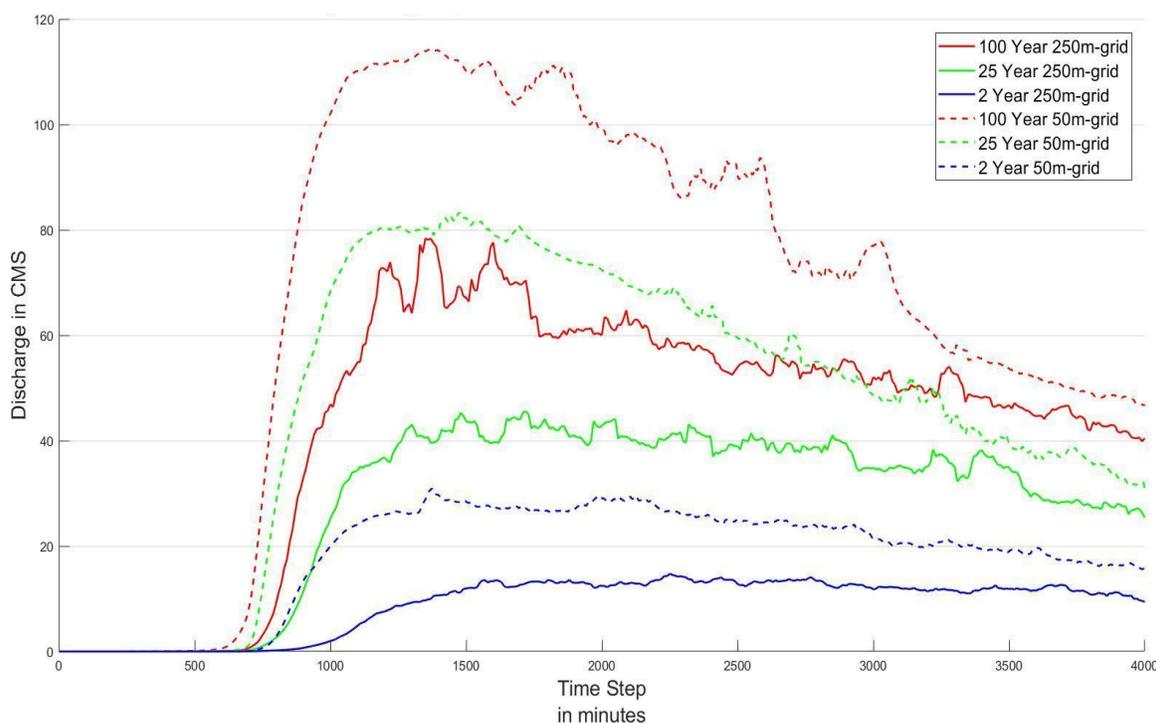


Figure 7. Comparing how different resolutions affect the hydrographs at the watershed outlet.

6. Conclusion

The percentage of agreement between two model increased with increasing rain events. The same was seen for the over predicted pixels by 250m. Unable to run multiple simulations (under same rain events) due to time constraints prevented us to identify, when these agreement and disagreement between two resolution makes statistically significant differences.

The results of the confusion matrix demonstrated that there is no statistical between two grid domains of the models. Since they do not show significant correlation at different three flood thresholds. However, it was noted that hyper-resolution (50m) was able to capture a larger range of water depths as well as limited the area of infiltration. Therefore, we conclude that there was no rain level at which initiating the higher resolution would be helpful predicted flooding before the rainfall. But once a known amount rainfall has occurred it would prudent to initiate a hyper-resolution model to more

accurately predict the depth and extent of infiltration.

Since the study area does not contain stream or rain gages, it was difficult to validate which model provided greater accuracy. However, the computational analogy of GSSHA gives the indication that a lower resolution of the model would cause more water to flow through the stream. We found that a 50 m resolution run over this ~ 100 km² watershed is not the most computationally efficient. Therefore, the future plan would be to incorporate a high performance, parallelizable model, such as AD-Hydro, to decrease computational time. In addition, comparing a known flood to the simulated flood would be useful in validating the research even further.

Acknowledgements: Dr. Jude Benavides, Dr. Fred Ogden, Dr. Sarah Praskievicz for leading us through this project. Dr. Sagy Cohen for use of his lab computer to run our high-resolution model runs. Robert Steinke and Nels Frazier for trying to set-up the AdHydro model for us. The National Water Center for supplying the atmospheric data and hosting us while we worked on the project.

References

1. Knowlton, K., Rotkin-Ellman, M., Geballe, L., Max, W., & Solomon, G. M. "Six climate change-related events in the United States accounted for about \$14 billion in lost lives and health costs." *Health Affairs*, 30(11), 2167-2176, 2011
2. O'Connor, J. E.; Costa, J. E. Large floods in the United States: where they happen and why; U.S. Dept. of the Interior, U.S. Geological Survey: Denver, CO, 2003.
3. Ambiotec Civil Engineering Group, Inc. "Flood Protection Plan"; submitted to the City of Brownsville, TX.
4. Whitko, Annemarie N. "Advanced floodplain mapping of a Rio Grande Valley resaca using LIDAR and a distributed hydrologic model." Masters, Rice University, Texas, 2005.
5. Rust; Lichliter; Jameson. Flood Protection Plan for the City of Brownsville, Texas; 1996.
6. Rafieeiniasab, A.; Norouzi, A.; Kim, S.; Habibi, H.; Nazari, B.; Seo, D.-J.; Lee, H.; Cosgrove, B.; Cui, Z. *Journal of Hydrology* **2015**, 531, 370–388
7. Interactive Mapper <https://www.sciencebase.gov/catalog/item/imap/55d3730fe4b0518e35468e1e> (accessed Jul 23, 2018).
8. United State Geological Survey. National engineering handbook; Water Resources Publications: Littleton, CO, 1985.
9. Bates, P.D. and De Roo, A. P. J., (2000). A simple raster-based model for flood inundation simulation. *Journal of hydrology*, 236(1), 54-77.

Chapter 4

Sensitivity of Urban Flooding to Presence of Subsurface Storm Drainage Systems in Hydrologic Models for Low-Gradient Watersheds

Isha Deo¹, Parth Modi², Mahkameh Zarekarizi³, and Jean Valle⁴

¹ The University of Texas at Austin; isha.deo@utexas.edu

² Virginia Tech; parth17@vt.edu

³ Portland State University; mahkam2@pdx.edu

⁴ University of Puerto Rico at Mayagüez; jean.valle@upr.edu

Academic Advisors: Paola Passalacqua, The University of Texas at Austin; Venkataramana Sridhar, Virginia Tech; Hamid Moradkhani, University of Alabama; Jonathan Muñoz Barreto, University of Puerto Rico at Mayagüez

Summer Institute Theme Advisors: Sarah Praskievicz, University of North Carolina, spraskie@uncg.edu; Fred Ogden, National Water Center, fred.ogden@noaa.gov; Jude A. Benavides, University of Texas Rio Grande Valley, jude.benavides@utrgv.edu

Abstract: In order to facilitate the National Water Center's mission to pursue hyper resolution hydrologic/hydraulic modeling, this work presents an urban flood inundation study using hyper resolution models to quantify the effects of subsurface storm drainage infrastructure. As urban populations increase globally, cities change natural flooding mechanisms through increased impervious land cover, detention and retention ponds, and storm sewer networks. Such expansion affects flood characteristics such as peak magnitude, time to peak, and the inundation area. This study addresses the sensitivity of these variables to the inclusion of subsurface storm drainage systems in hyper-resolution modeling. Using the Gridded Surface-Subsurface Hydrologic Analysis (GSSHA) Model in the Town Resaca basin in Brownsville, Texas, we show that including the storm sewer infrastructure is both feasible and beneficial for accurate flood inundation mapping in urban areas. However, at storm return periods above 5 years, the additional benefit in flood inundation accuracy is limited. At storm return periods less than the critical return period of 5 years, hyper resolution models of storm sewers can be a critical addition to urban flood inundation studies.

1. Motivation

An increasing global population strains natural resources and land availability, causing rapid urbanization. The associated increase in impermeable surfaces heightens flood-related damage caused by quantity and movement of the stormwater runoff. The former is governed by the type of the land use; the latter is governed by the density, size and characteristics of channels and provision of storm sewerage [1]. Increased urbanization results in higher flood peak, higher flood volume, and lower time to peak indicating a higher risk of flooding [2]. The subsurface storm drainage network has to move a higher volume of water effectively to prevent any flooding event. To address this problem, hydrologic models are used to study and forecast events that may pose a threat to the community. Given that the stormwater drainage system plays a vital role in the movement of runoff, this research evaluates the

difference between model runs with and without this network as part of the watershed. The authors aim to identify a storm recurrence interval at which hyper resolution flood inundation modeling with the subsurface storm sewer network ceases to improve quality of inundation predictions.

2. Objectives and Scope

The objective of this study is to quantify the sensitivity of hyper-resolution hydrologic models to the presence of subsurface storm drainage systems. By running the Gridded Surface-Subsurface Hydrologic Analysis (GSSHA) Model [3] on a hyper resolution scale, the impact of subsurface hydraulic structures on low-gradient watersheds can be quantified. Modeling the same scenario without the stormwater drainage system provides the benchmark for such comparison. By varying the design storms for each model scenario, the authors aim to identify a threshold return period after which the effect of storm sewers in the model is insignificant to flood extent. This return period is expected to be the design storm event used by cities and consulting engineers. Identifying this return period can help the National Water Center improve water prediction.

The study area is the Town Resaca basin in Brownsville, Texas, with an area of 5.88 sq. mi. A distinctive characteristic of the Lower Rio Grande Valley is its *Resaca* System. *Resacas* are historic channels connecting to the Rio Grande, that served to transport water away from the river during floods. Currently, the water flowing through these *Resacas* is mostly runoff generated by precipitation or the water pumped from the river. A subset of the system, Town Resaca, is shown in **Figure 1**.

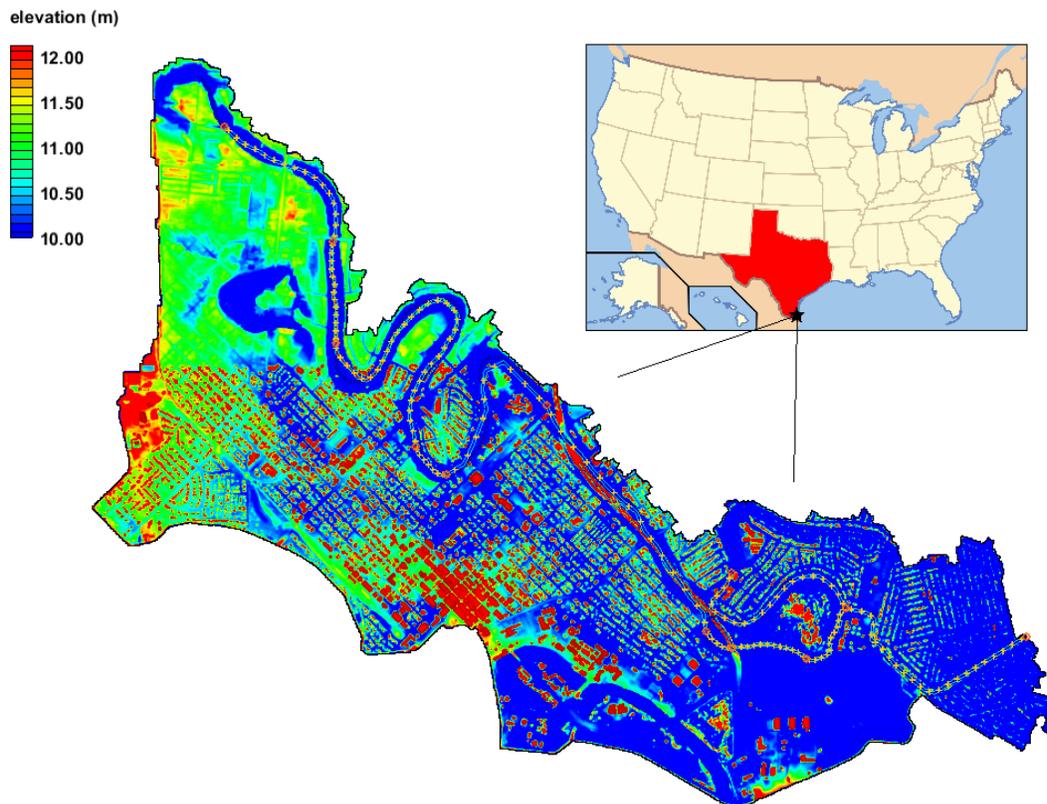


Figure 1. Town Resaca flowline is shown in blue. The central, straight part of the flowline is a culvert under a highway. The watershed outline is delineated in orange.

The City of Brownsville is a challenging test case for hyper-resolution modeling due to the following reasons. Firstly, the low elevation difference throughout the Texas Gulf Coast creates a scenario in which even a small storm causes flooding [4]. Secondly, the storm sewer system in Brownsville has proved inadequate historically, with street flooding occurring regularly during high intensity storms. Finally, Lower Rio Grande Valley faces political controversy, as the Rio Grande is not only a primary water source, but also an international boundary, creating complex water rights problems in the region.

3. Previous Studies

The study of urban flood inundation has long focused on changes in impervious cover [5], but has largely neglected the effects of man-made flood control measures such as storm sewer networks, due to high computational resources and time in modeling storm sewers at a fine scale. Recent studies of flood inundation with storm sewer networks have largely been localized [6], or use subcatchment level hydrologic models (such as EPA SWMM), to calculate overland water depth [7, 8]. However, recent improvements to GSSHA allow the introduction of storm sewer systems using a SUPERLINK approach [9, 10]. Using the SUPERLINK configuration on GSSHA, urban storm sewers are shown to increase the peak discharge in urban areas for moderate and high precipitation storm events, but show no difference in extreme events [10, 11]. Studies of storm sewer systems in hyper resolution modeling have focused primarily on the stream hydrograph, generally ignoring inundation depth and extent.

As the Lower Rio Grande Valley, TX remains an understudied area, region-specific studies are sparse. Brownsville's Flood Protection Plan examines the state of flood-related problems in Town Resaca as recently as 2011 and highlights the use of hydraulic and hydrologic models in flood studies. Certain areas in Town Resaca are known to flood for 25-year events, with some experiencing inundation at even 2-year storms [12]. Creating an accurate model of the flood inundation at various return periods is beneficial for Brownsville to facilitate further flood control measures.

4. Methodology

4.1. Initiation of the GSSHA Model

In order to run the GSSHA model with storm sewers, several components had to be initialized for the study area. The DEM and the channel cross sections were prepared specially for input into GSSHA. The storm sewer network was adapted from shapefiles into a SUPERLINK format used by GSSHA.

4.1.1. Adjustment of DEM

This study uses the hydrologic model GSSHA due to its ability to accurately simulate the sewer system present in the Brownsville watershed. For a successful model run multiple inputs are required, such as: Digital Elevation Model (DEM), land use and soil data, precipitation, channel cross-sections and sewer system information. The DEM used for the simulation was a 1 meter resolution LiDAR image of the Brownsville area. The resolution of this dataset was ideal for the requirements of a hyper-resolution model run. The pre-processing of the DEM was made using Arc-GIS and the Arc-Hydro toolbox. This step required detailed attention because the characteristics of the Resaca system complicated the delineation of the streams. The low elevation gradient in the area created a scenario where the Topographic Parameterization Program (TOPAZ) could not recognize some streams, leading to an incorrect delineation of the watershed in GSSHA. This problem was solved by burning the streamlines into the elevation dataset with the DEM Reconditioning option in Arc-GIS.

4.1.2. Addition of supplementary information

After successfully delineating the watershed in GSSHA, 27 previously surveyed channel cross sections were added to the model. The availability of this dataset provides the opportunity to create a model that represents the real conditions more accurately. In the middle part of the Resaca, the stream flows in a culvert for a distance of 1.27 miles approximately. This hydraulic structure was modeled as a rectangular channel throughout its length with culvert specifications at the beginning and end, with no overbank flow in this underground section. The soil and land use data were acquired from the Natural Resources Conservation Service and National Land Cover Database respectively. The grid size selected for the model run was 10 meter to produce results that meet the hyper-resolution criteria.

4.2. Compilation of subsurface storm drainage network

4.2.1 Preprocessing of storm sewer network

The subsurface storm drainage network in Brownsville had minimal existing information, requiring significant preprocessing before inputting into GSSHA. The available data was limited to a) surveyed storm sewer inlet locations, b) a preliminary shapefile of the storm sewer pipe network, without any network characteristics, c) georeferenced storm drainage system plans. The storm sewer inlet locations were validated using Google Earth before any further processing.

To minimize the number of inlet points in the sewer network, inlets were grouped at road intersections using a Python script and ArcGIS. All surveyed inlet locations within a 50 ft radius were combined into one point with multiple grate inlets. The pipe network was manually edited to split pipes at bends, junctions, and inlets. Topology relationships through ArcGIS forced the inlet locations to match up with pipe vertices where possible, forcing inlets to align with road intersections. The georeferenced city sewer plans were used to add pipe diameter information where available. Missing or contradictory dimensions were inferred based on outlet locations and assumed flow directions.

A Matlab script was developed to calculate the flow direction in each pipe based on pipe diameters, and nearest outlet location. After manually entering outlet nodes, each upstream pipe was identified using a recursive tree algorithm and assigned a flow direction based on pipe diameter and distance to outlet. This script greatly reduced manual entry and human error and provides a reliable method of modeling pipe flow directions from minimal information. The final flow directions and pipe dimensions can be seen in **Figure 2**.

4.2.2. Calculation of pipe invert elevations

As GSSHA requires pipe invert elevations at each node, a Matlab script was developed to calculate approximate elevations based on pipe flow directions. Using a similar recursive tree algorithm as the flow direction calculation, the invert elevation at the outlet was assumed relative to the ground elevation, and the invert elevation at each upstream node was calculated using a slope of 0.0125%.

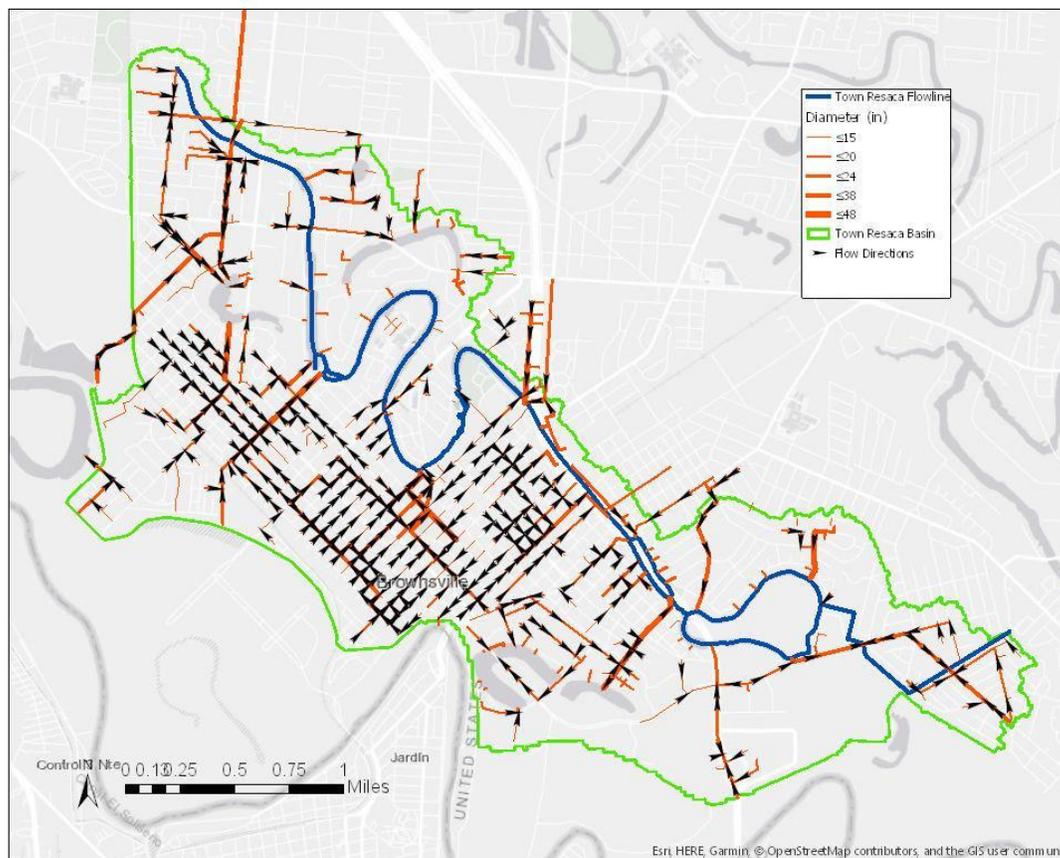


Figure 2. The pipe network and flow directions for the Town Resaca basin are shown.

4.2.3. Conversion of sewer network for GSSHA input

After preprocessing the sewer network into a directed graph with invert elevations, the pipe network had to be digitized into a format legible by GSSHA. In the Watershed Modeling System, or WMS, the sewer network is automatically generated based on cell to cell connectivity by individually adding pipes into the network and parametrizing them based on the geometry associated with the location of nodes. As WMS requires pipes, inlets, and outlets to be manually drawn into the GUI, we pursued adding the storm sewer network file externally, based on the SUPERLINK model used in GSSHA.

To simulate storm and tile drains, the SUPERLINK model^[7] is implemented into GSSHA. It solves the full-dynamic form of the Saint Venant equations in one dimension and employs the Preissmann slot to extend the open channel flow assumptions to closed conduits flowing full and surcharged [13]. In the SUPERLINK model, the water flow through the storm drains is a transformation of the overland flow through inlets. The SUPERLINK model includes four types of network components: superlinks, superjunctions, links, and computational nodes. Storm sewers are designed as superlinks connecting superjunctions. Types of superjunctions include inlets/manholes, outlets, and other critical locations. However, grate inlets along roadways are considered as computational nodes along the superlink. Node behavior is determined by the number of grates, as either orifice-type or weir-type.

The subsurface storm drainage system is connected to the stream network based on the stream link/node structure, where storm sewer outlets transfer water to the stream at stream nodes.

Therefore, each outlet contains information on the connecting stream link and node, the ground surface elevation, and the pipe invert elevation. Similarly, the inlets have the ground surface elevation, pipe invert elevation, and grid location. The grid location connects the inlets to the overland flow by transferring some fraction of water based on number of grate inlets.

Finally, in order to further simplify the pipe network for import into GSSHA, short pipes created to connect an inlet node to a road intersection node were removed. Pipes with no more than two connecting pipes were combined into one superlink, thus reducing the number of pipes significantly. This change can be seen in **Figure 3** below, and the changes to the pipe network can be seen in supplementary information

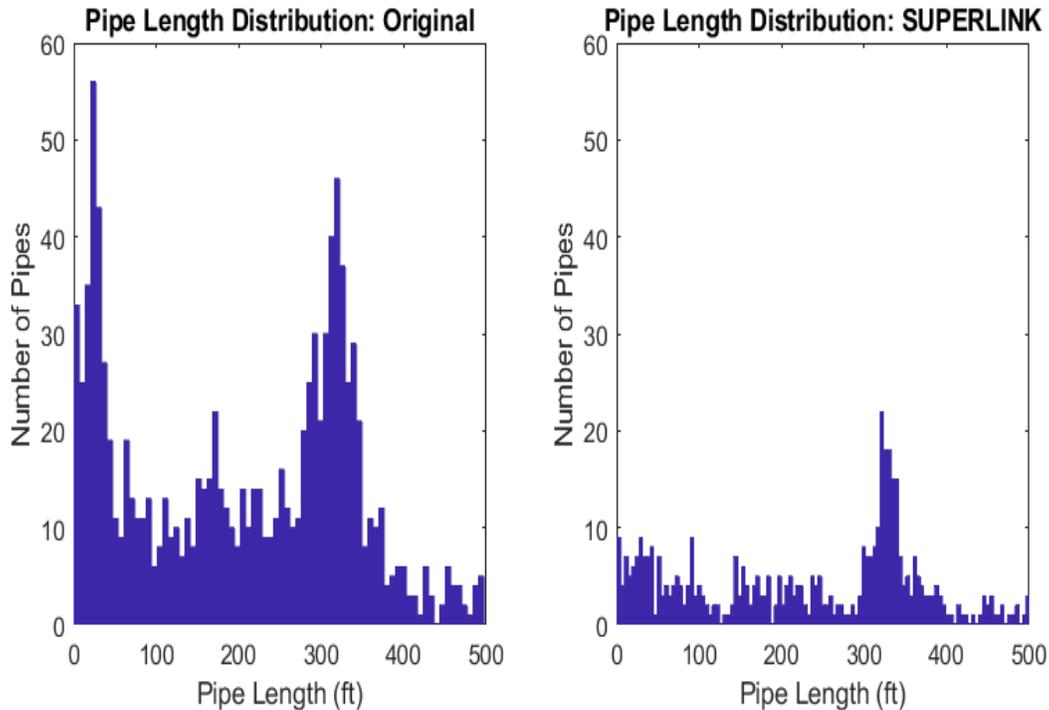


Figure 3. Histograms of pipe length distribution (a) before and (b) after simplifying superlinks

The GSSHA system can handle detailed hydrologic and hydraulic information, including the storm sewer network, surveyed channel cross sections, and hydraulic structures. Although including each of these components increases the accuracy of the model, it significantly increases computational time. Given a limited time frame, the authors intended to provide a reliable comparison of flood inundation with and without the subsurface storm sewer network, forcing us to run models without the hydraulic structures and weirs that exist on Town Resaca. However, each run is on a hyper resolution (10 m) with a complex sewer network and with surveyed channel cross sections.

4.3. Atmospheric Forcings

Synthetic rainfall events developed by U.S Department of Agriculture (USDA) Soil Conservation Service (SCS), also known as National Resources Conservation Service (NRCS), are commonly used for drainage design [14]. Four synthetic storm distributions (I, IA, II, and III) are available for the US, each assigned to a geographic boundary and Brownsville, the target study area, is located in the type III boundaries. Type III distributions [15] are based on Hydro-35 (1977) and TP-40 (1961).

Storms with various return periods (i.e. 2-, 5-, 10-, 25-, 50-, 100-, 250-, and 500-year) are incorporated

in this study. A wide distribution of return periods is considered in order to find a pattern of flood extents with and without storm sewers incorporated in the model. These specific return periods were chosen because they are common design return periods for various infrastructure. Storm depths over 24 hours are extracted from Hydro-35 report and are presented in the following **Table 1**. **Figure 4** shows the Type III rainfall distribution used in this study.

Table 1. Storm return periods and associated precipitation depths based on Hydro-35 Report

Return period (yrs)	2	5	10	25	50	100	250	500
Rainfall depth (mm)	91.44	133.35	165.1	203.2	241.3	279.4	330.2	381

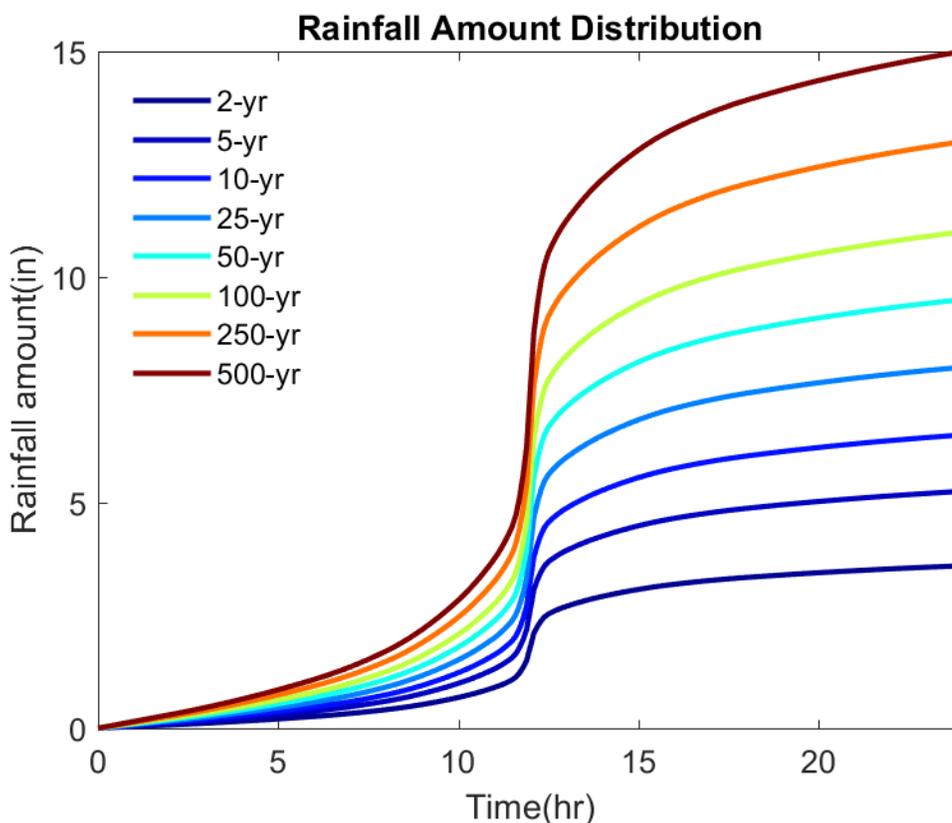


Figure 4. Rainfall distribution over time for Type III areas.

4.4. Running GSSHA

Using the finalized GSSHA model, several precipitation return periods were tested with and without storm sewers. After testing 2-, 5-, 10-, 25-, 50-, and 100-year storms, a threshold return period above which the storm sewer network ceases to improve flood inundation is expected. The authors expect this return period to be around the design period of the storm sewers (or approximately 5 years). However, for higher return periods, the storm sewer network may alleviate some road inundation.

5. Results

5.1. Hydrographs and overland water depth grids for 2-year storm with & without storm sewers

This section highlights the methodology to be used to evaluate the flood inundation rasters produced as an output of the GSSHA model, based on the 2-year, 24-hour storm. **Figure 5** compares the stream outlet hydrographs with and without the storm sewer network. **Figure 6** shows a comparison of the flood inundation grid for each simulation.

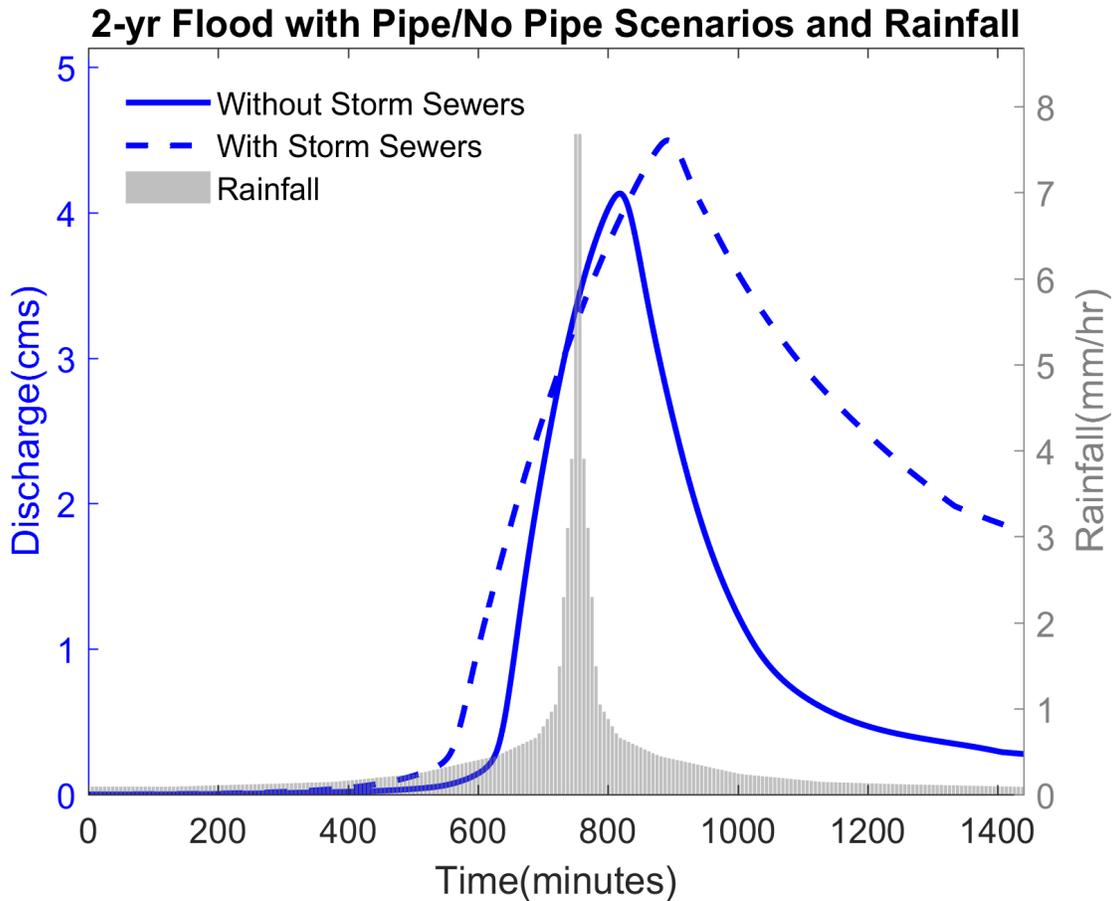


Figure 5. Stream outlet hydrographs for no storm sewers (solid blue line) and with storm sewer network (dotted blue line), and rainfall hyetograph for 2-year, 24-hr storm. Note that the hydrograph peak discharge is higher with storm sewers, with a higher total volume in the stream. Other hydrographs are provided in supplementary information.

Based on the streamflow hydrograph, the storm sewer network put in place is functioning as expected, by increasing the peak discharge and the total volume in the stream [10]. At a 10 m grid resolution, the impact of storm sewers on minimizing street flooding is apparent, especially for the low return period storm shown here. For higher return periods, this difference decreases. These can be seen in the supplementary information provided.

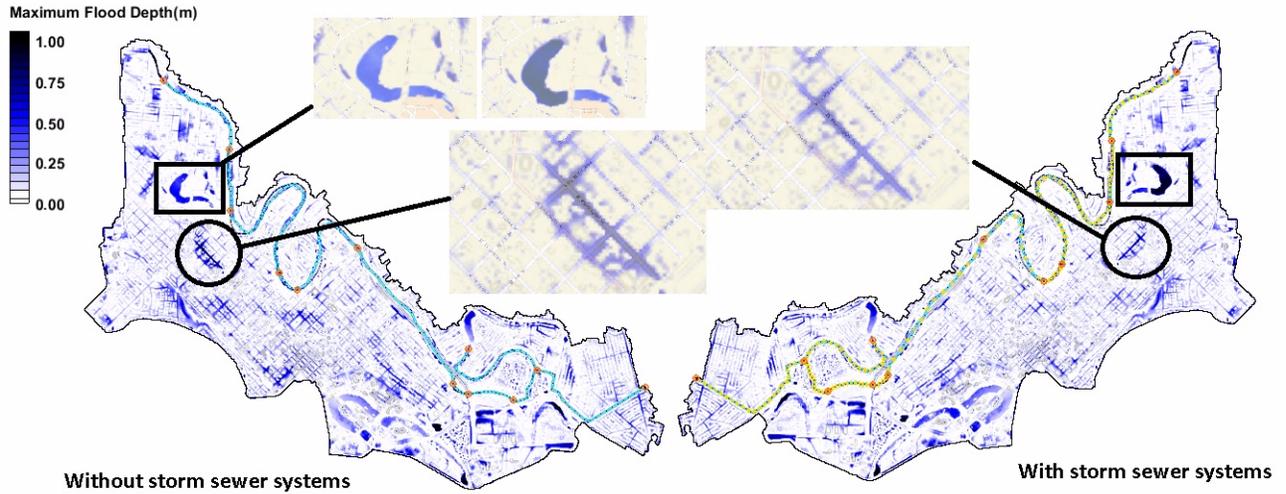


Figure 6. Flood inundation maps for 2-year, 24-hour storm shows with and without storm sewers. Specific neighborhoods with severe flooding are highlighted.

5.2. Inundation area comparison

Figure 7 shows the inundated area for each storm return period, for both storm sewer and no storm sewer simulations. It is clear that the difference in inundated area with storm sewers and without storm sewers decreases for higher return periods in the 24-hour storm, especially. The difference is less noticeable in the 3-hour storms.

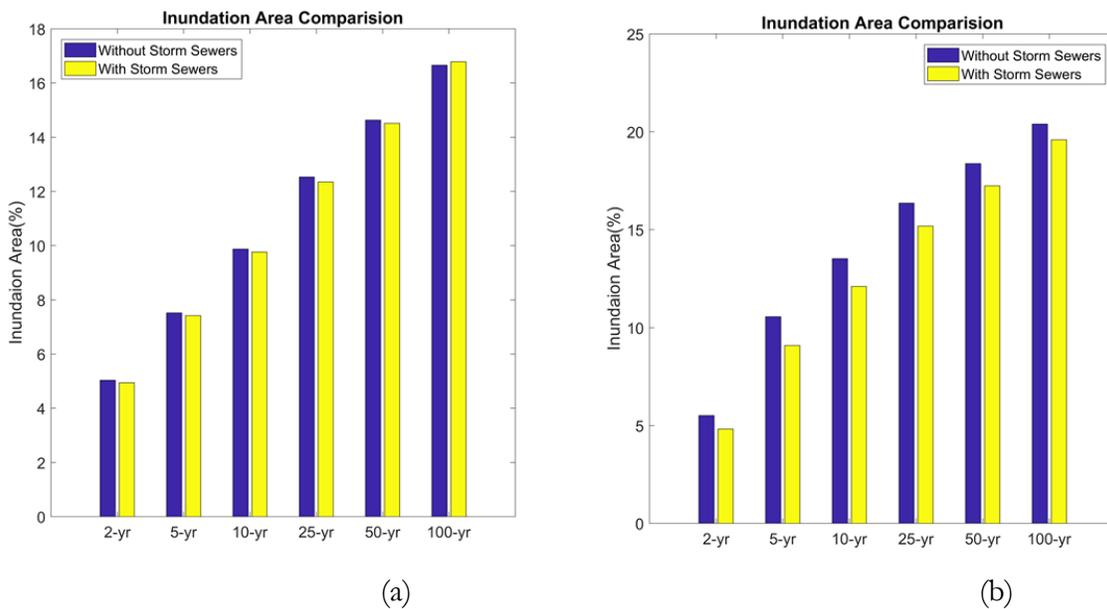


Figure 7. Inundated area comparison for (a) 3-hour storm and (b) 24-hour storm.

5.3. Confusion matrix analysis

At this point, we have created all of our model runs and are waiting for them to complete in order to conduct the quantitative analysis on the flood inundation rasters. Two approaches of flood inundation evaluation were used: 1) comparison of inundated raster cells, and 2) comparison of inundated address points between storm sewer and no storm sewer runs. Both approaches are evaluated using a confusion matrix, with the storm sewer inundation map taken as the benchmark. The elements of the matrix are summarized in **Table 2** below.

Table 2. Descriptions and consequences of possible combinations of flood extents from storm sewer & no storm sewer simulations. Each element is based on the storm sewer flood extent taken as the truth.

Element	Flooded in Storm Sewer Simulation	Flooded in No Storm Sewer Simulation	Description/Consequences
True Positive (TP)	Yes	Yes	Correctly identified flooded areas
True Negative (TN)	No	No	Correctly identified not flooded areas
False Positive (FP)	No	Yes	Superfluous flood threat causes distribution of resources away from areas in need.
False Negative (FN)	Yes	No	Flood threat is real but not modeled, leaving areas unprepared for response.

5.4. Confusion Matrices for 24-hour storms

The confusion matrices for the (a) grid cells and for the (b) address points are given in **Table 3**. Currently only the results for the 2-, 5-, and 100-year, 24-hour storm are available. To highlight specific neighborhoods/city blocks where storm sewer modeling is especially important, the address point confusion matrix is depicted on a map in **Figure 7**. Once all the return periods are run at a 10 m scale, we will be running a Signed Rank hypothesis test to determine if the inclusion of storm sewers makes a significant difference on the flood inundation rasters, as well as a Kruskal-Wallis test for distribution between the flooded cells for runs with and without storm sewers.

Figure 7 gives the false positivity calculated from the confusion matrix. A peak in false positivity indicates the critical return period for each storm duration. For a 24-hour duration, a peak is clear at 5 years. However, for a 3-hour duration, a peak occurs at 25 years but is not much different from the other return periods, so a critical return period is difficult to identify.

In **Figure 8**, the results of the address point inundation study are given. This highlights specific addresses where the storm sewer network is helping alleviate flooding. At higher return periods, we expect spatial trends to arise in neighborhoods that are either more or less susceptible to flooding than the current non-storm sewer flood inundation standard suggests.

Table 3. Confusion matrices for each 10 m run for 24 hour storms in (a) flood grid cells and (b) address points and for 3 hour storms in (c) flood grid cells and (d) address points using a 1 ft threshold for inundation. Storm sewer runs are included as ss while runs without storm sewers are nss. Flooding is indicated by f and no flooding is indicated by nf.

(a) Number of 10 m × 10 m raster cells - 24 hr

	2 year		5 year		10 year		25 year		50 year		100 year	
	ss_f	ss_nf	ss_f	ss_nf	ss_f	ss_nf	ss_f	ss_nf	ss_f	ss_nf	ss_f	ss_nf
nss_f	6770	1410	13348	2464	18075	2397	22471	2073	25977	1652	29162	1539
nss_nf	382	143764	256	136258	216	131638	290	127492	354	124343	299	121326

(b) Number of address points - 24 hr

	2 year		5 year		10 year		25 year		50 year		100 year	
	ss_f	ss_nf	ss_f	ss_nf	ss_f	ss_nf	ss_f	ss_nf	ss_f	ss_nf	ss_f	ss_nf
nss_f	34	12	85	34	138	47	204	40	256	14	294	33
nss_nf	0	9356	3	9280	0	9217	1	9157	5	9127	5	9070

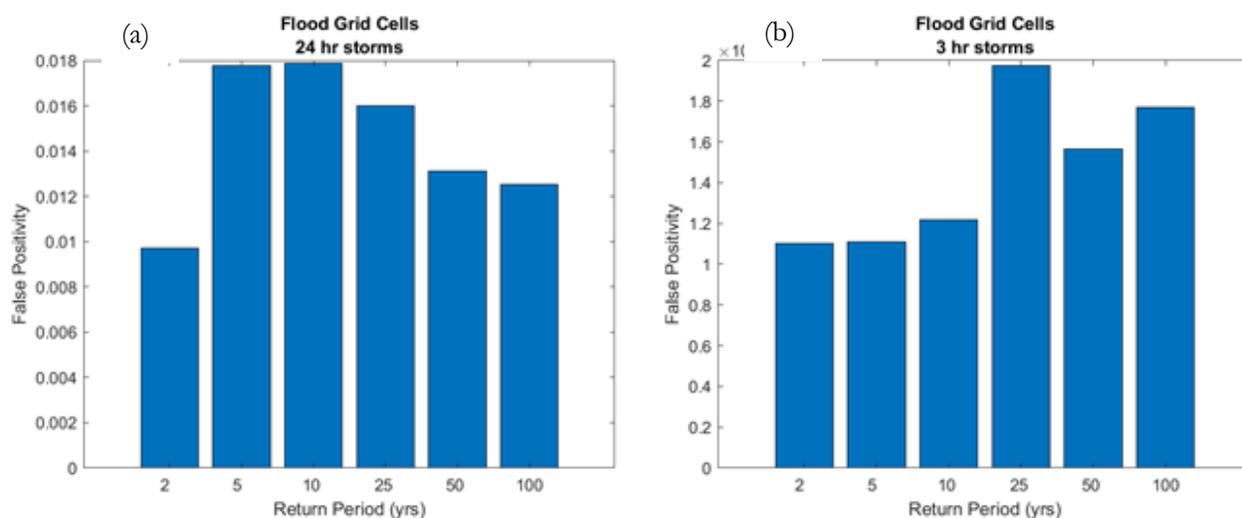
(c) Number of 10 m × 10 m raster cells - 3 hr

	2 year		5 year		10 year		25 year		50 year		100 year	
	ss_f	ss_nf	ss_f	ss_nf	ss_f	ss_nf	ss_f	ss_nf	ss_f	ss_nf	ss_f	ss_nf
nss_f	7308	160	10998	157	14525	168	18510	264	21848	204	25104	225
nss_nf	10	144848	8	141163	15	137618	18	133534	28	130246	23	126974

(d) Number of address points - 3 hr

	2 year		5 year		10 year		25 year		50 year		100 year	
	ss_f	ss_nf	ss_f	ss_nf	ss_f	ss_nf	ss_f	ss_nf	ss_f	ss_nf	ss_f	ss_nf
nss_f	38	4	69	13	111	6	150	4	199	10	241	7
nss_nf	0	9360	0	9320	0	9285	1	9247	0	9193	0	9154

Figure 7. Comparison of false positivity index across return periods for both 24-hour storms (a) and 3-hour storms (b).



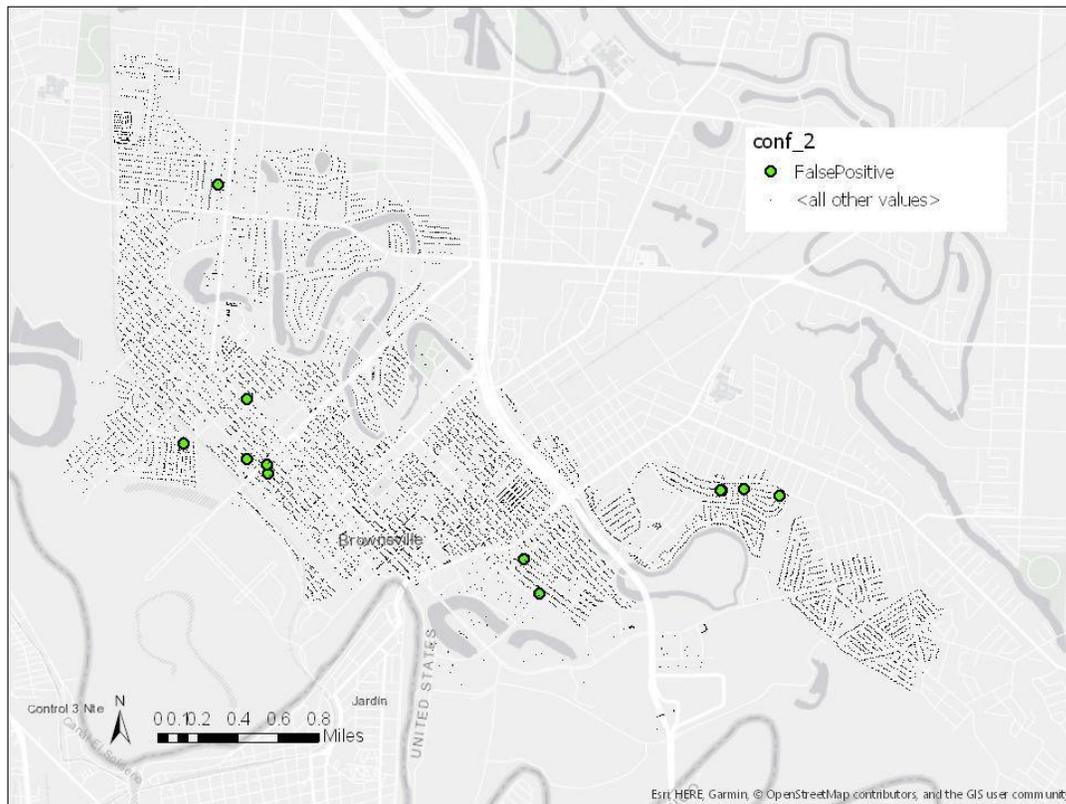


Figure 8. Map of address point confusion matrix for 2-year, 24-hour storm at 10 m resolution. Green points are false positives, showing where the storm sewer network is alleviating flooding. Red points are flooded with storm sewers but not flooded without storm sewers (none visible at this return period)

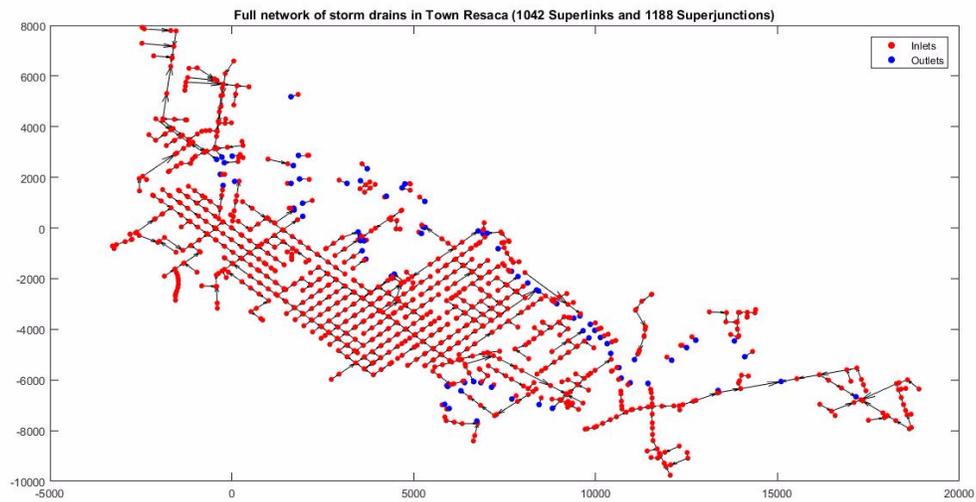
6. Conclusion

This work demonstrated that modeling subsurface storm drainage networks in hyper resolution models is not only feasible but also valuable for flood inundation studies. A basic sewer network was transformed into the SUPERLINK format used by GSSHA to model storm sewers. It is clear that at increasing return periods, the effectiveness of the storm sewers in preventing inundation decreases, as the number of raster cells where flooding is avoided decreases.

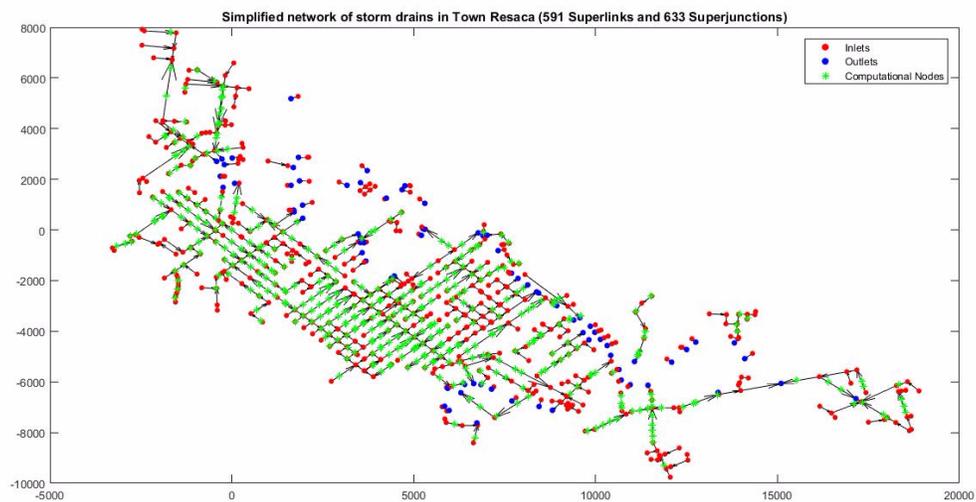
Modeling storm sewer networks as a part of hyper resolution flood inundation studies conducted by the National Water Center can be an effective way to address growing flooding concerns in the United States. However, the authors suggest an integration of the SUPERLINK framework in a triangulated mesh model such as AD-Hydro. Although the GSSHA framework is able to handle the storm sewer network in Town Resaca, it required significant computational time and resources.

Supplementary Materials:

S.1. Distribution of Superlinks before and after simplification

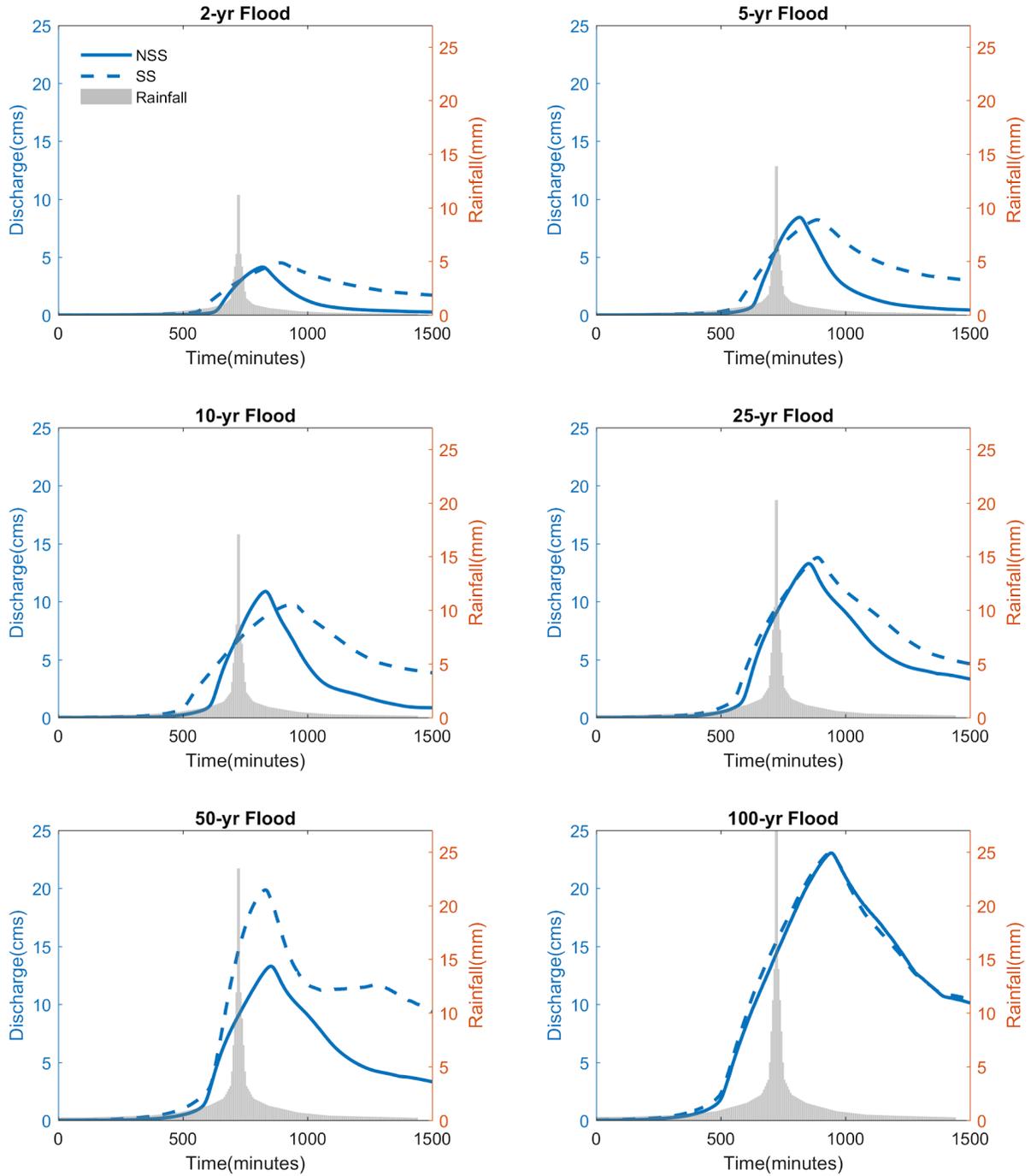


Full network of storm drains in Town Resaca. Inlets are shown by red dots whereas outlets are shown by blue dots. All of them are considered as Superlinks and Superjunctions (1042 Superlinks and 1188 Superjunctions).

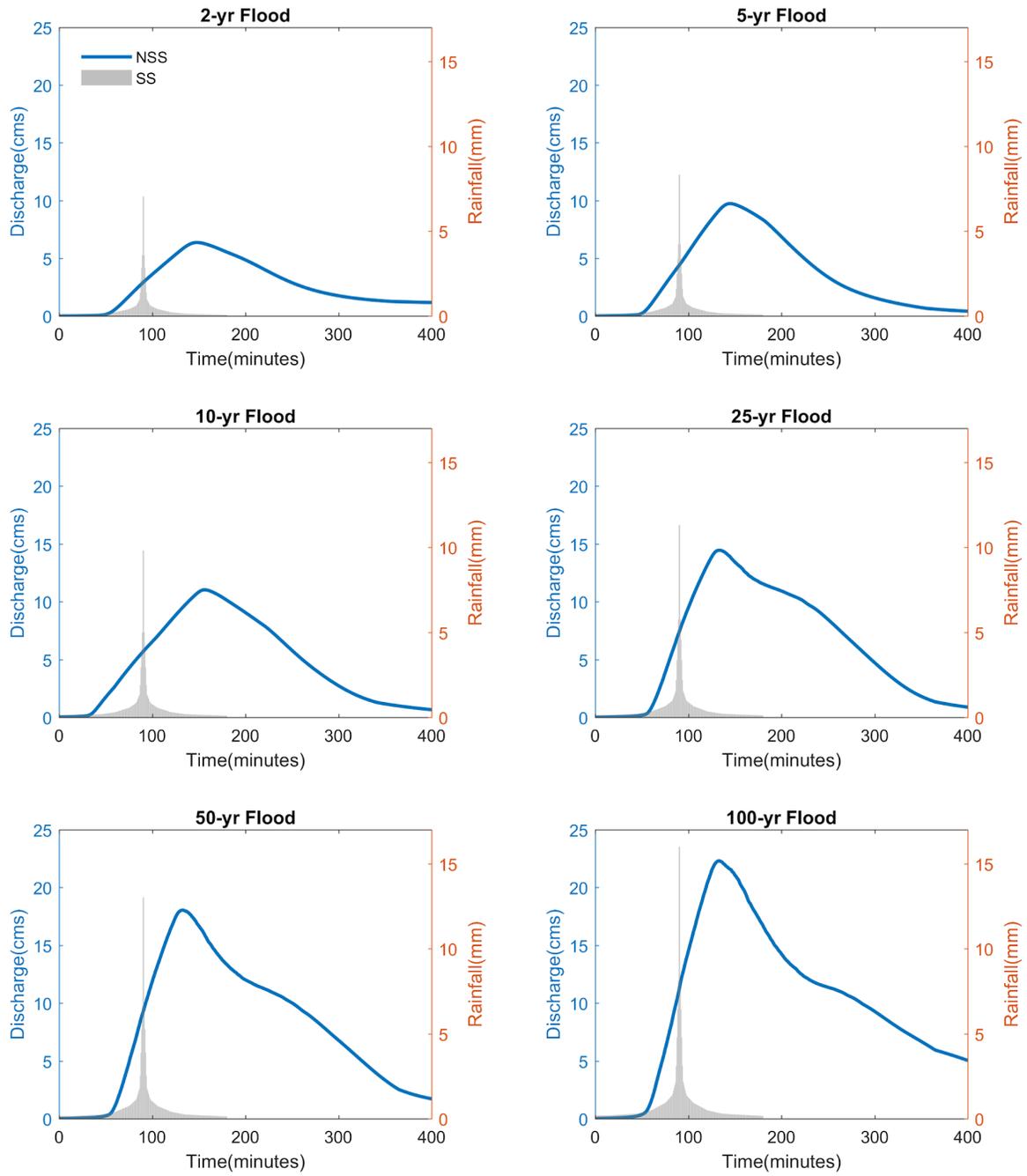


Simplified network of storm drains in Town Resaca. Inlets are shown by red dots, outlets are shown by blue dots and green dots indicate computational nodes. Some of them are considered as Superlinks and Superjunctions (591 Superlinks and 633 Superjunctions).

S.2. Additional Stream outlet hydrographs for no storm sewers (solid blue line) and with storm sewer network (dotted blue line)

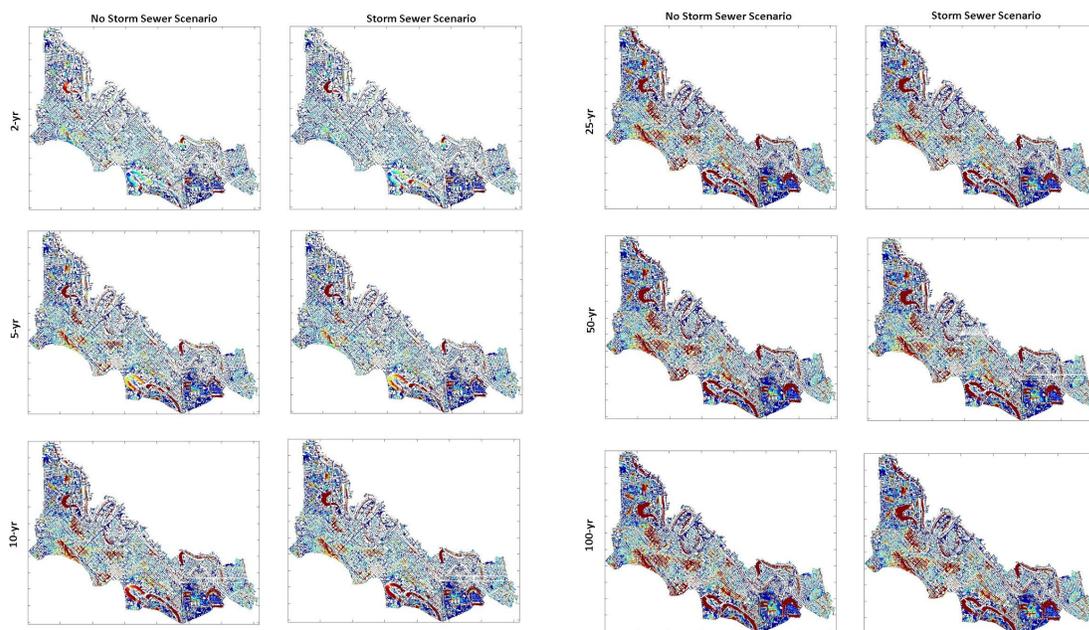


Hydrographs for 24-hour storm.

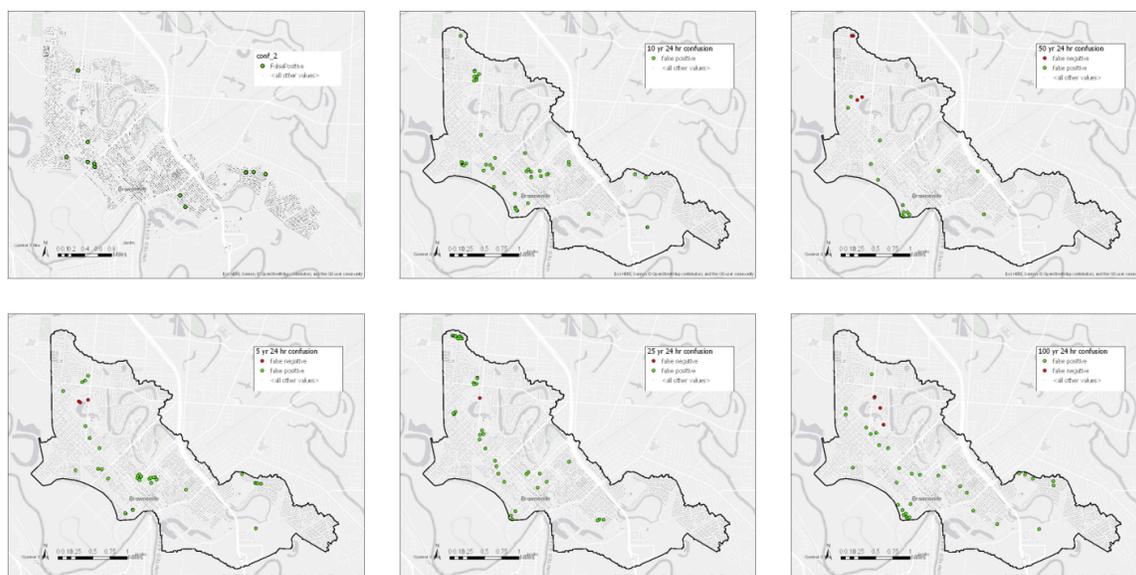


Hydrographs for 3-hour storm

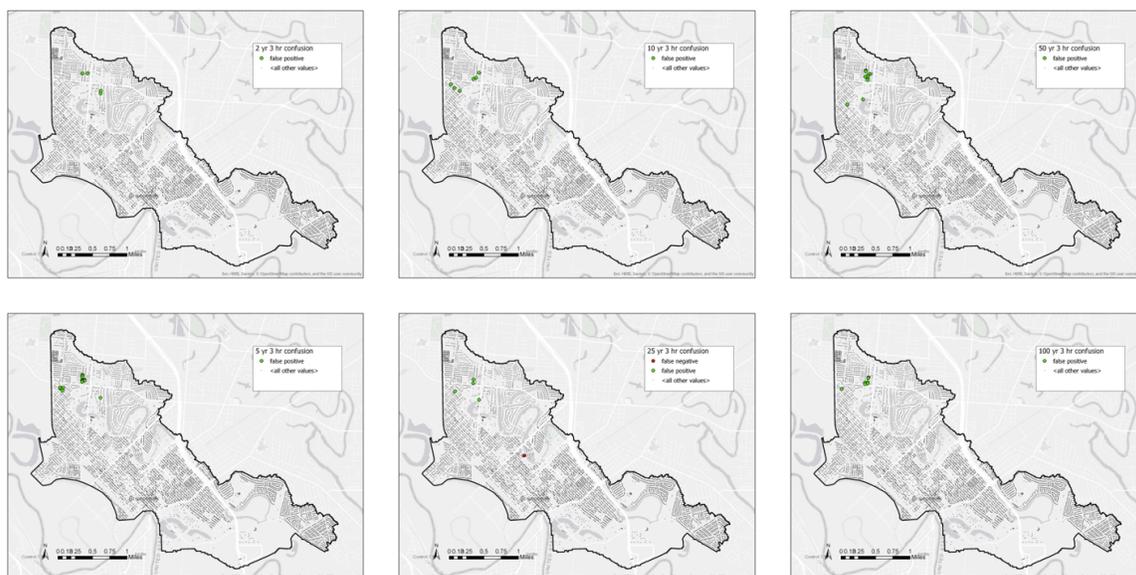
S.3. Inundation rasters and confusion matrices



Flood grid rasters for each return period for 24-hour storms, to compare inundated areas.



Address points confusion matrix results for 24-hour storm.



Address point confusion matrix results for 3-hour storm.

References

1. Leopold, L. Hydrology for Urban Land Planning - A Guidebook on the Hydrologic Effects of Urban Land Use. *Geol. Surv. Circ.* **1968**, 554, 1–21.
2. Suriya, S.; Mudgal, B. V. Impact of Urbanization on Flooding: The Thirusoolam Sub Watershed - A Case Study. *J. Hydrol.* **2012**, 412–413, 210–219.
3. Downer, C. W.; Ogden, F. L. GSSHA: Model To Simulate Diverse Stream Flow Producing Processes. *J. Hydrol. Eng.* **2004**, 9 (3), 161–174.
4. *Flood Protection Plan*; 1997.
5. Hollis, G. E. The Effect of Urbanization on Floods of Different Recurrence Interval. *Water Resour. Res.* **1975**, 11 (3), 431–435.
6. Fang, X.; Su, D. An Integrated One Dimensional and Two Dimensional Urban Stormwater Flood Simulation Model. *JAWRA J. Am. Water ...* **2006**, 77042, 713–724.
7. Jiang, L.; Chen, Y.; Wang, H. Urban Flood Simulation Based on the SWMM Model. *LAHS-AISH Proc. Reports* **2015**, 368 (August 2014), 186–191.
8. Hsu, M. H.; Chen, S. H.; Chang, T. J. Inundation Simulation for Urban Drainage Basin with Storm Sewer System. *J. Hydrol.* **2000**, 234 (1–2), 21–37.
9. Ji, Z. General Hydrodynamic Model for Sewer/Channel Network Systems. **2006**, 551 (5), 15–34.
10. Ogden, F. L.; Raj Pradhan, N.; Downer, C. W.; Zahner, J. A. Relative Importance of Impervious Area, Drainage Density, Width Function, and Subsurface Storm Drainage on Flood Runoff from an Urbanized Catchment. *Water Resour. Res.* **2011**, 47 (12), 1–12.
11. Ogden, F. L.; Ph, D.; Niedzialek, J. M.; Ph, D.; Byrd, A. R. Storm Drain Effects on Urban Flooding. **2012**, No. August, 1–14.
12. Consultants, A. E. *Flood Protection Plan - Phase I*.
13. Downer, C. W.; Pradhan, N. R.; Byrd, A. R. Modeling Subsurface Storm and Tile Drain Systems in GSSHA with SUPERLINK Coastal and Hydraulics Laboratory. **2014**, No. September.

14. Brown, S. A.; Schall, J. D.; Morris, J. L.; Doherty, C. L.; Stein, S. M.; Warner, J. C. Urban Drainage Design Manual. **2013**, *2009* (22), 478.
15. Woodward, D. E.; Cronshey, R. G. Flood Skew in Hydrologic Design On Ungaged Watersheds. *Jounral Irrig. Drain. Eng.* **1988**, *114* (2), 301–310.

Chapter 5

Using Dimensionless Scaling Parameters as Decision Metrics in a Heterogeneous Hydraulic Routing Scheme

Maryam Asgari Lamjiri¹, Kelly Flint² and Sean Matus³

¹ University of California San Diego; masgari@ucsd.edu

² San Diego State University; kflint@sdsu.edu

³ University of Illinois at Urbana-Champaign; matus2@illinois.edu

Academic Advisors: F. Martin Ralph, University of California San Diego; Alicia Kinoshita, San Diego State University; Praveen Kumar, University of Illinois at Urbana-Champaign

Summer Institute Theme Advisors: Ehab Meselhe, Water Institute of the Gulf, emeselbe@thewaterinstitute.org; Kyle Mandli, Columbia University, kyle.mandli@columbia.edu

Abstract: Homogeneous hydraulic routing schemes are subject to trade-offs between computational accuracy and resource consumption. In operational environments, such as with the National Water Model (NWM), accuracy is important, while computational efficiency and robustness are imperative. A heterogeneous routing scheme could theoretically give forecast centers the advantage of preserving accuracy while conserving resources by only implementing the Dynamic Wave model when it provides a significant increase in accuracy. This study proposes that there are instances when the Dynamic Wave or Diffusive Wave approximations are not necessary to produce sufficiently accurate results, and that Dimensionless Scaling Parameters (DSPs) can be used to initiate transitions between the Dynamic, Diffusive, and Kinematic Wave routing methods. A framework is constructed to investigate the relationships between the terms of the Saint-Venant momentum equation and DSPs. The framework automates simulation of the MESH model to produce a sample of points spanning a range of hydraulic scenarios. The workflow then analyzes that sample, both visually and statistically. It should be emphasized that the sample population considered in this work does not cover the full spectrum of all possible scenarios. The focus here is to establish the approach that could be expanded upon through a follow up effort. It is seen that the full Dynamic Wave provides unique accuracy over the Diffusive and Kinematic Wave models for only ~5% of sample cases, and the Diffusive Wave provides unique accuracy over the Kinematic Wave for ~75%. Histograms at different positions relative to transition thresholds show variable ranges for DSPs, primarily the Courant and Froude number, and may be interpreted as a justification for a routing scheme decision. Principal component analysis reveals redundant DSPs and demonstrates potential to statistically relate DSPs and terms of the momentum equation. Future work may utilize the framework established to capture a sample that better explores the entire parameter space and incorporates considerations for additional DSPs.

1. Motivation

Hydrologic and hydraulic model forecasts inform decisions regarding public health and safety, and management of watersheds and water resources. Decisions concerning public health and safety, as in

the cases of flood and evacuation warnings, are often time sensitive and have consequences which may immediately impact preservation of life, economy, and public perception of the efficacy of local leadership. Thus, forecast accuracy and computational efficiency of hydraulic models are subjects of particular interest.

The Saint-Venant equations, which govern conservation of flow volume and momentum throughout a channel, are used in one-dimensional flow-routing. The momentum equation is nonlinear, and while numerical methods reduce computational complexity, solving the full momentum equation, or Dynamic Wave equation, is computationally expensive relative to more simplistic models. The Diffusive Wave and Kinematic Wave equations are simplified forms of the momentum equation that are used in lieu of the Dynamic Wave equation to improve computational efficiency. Both the Diffusive Wave and Kinematic Wave neglect the effects of inertia, and the Kinematic Wave additionally neglects pressure gradient (P.G.) effects. While literature presents criteria for appropriating routing methods to different river wave types [1][2], the limitation of current routing schemes requiring selection of a single routing scheme to be implemented throughout the model domain and time frame forces a trade-off between computational cost and accuracy where flow conditions vary over time and space. One possible form of the Saint-Venant equations is provided below:

$$\frac{\partial Q}{\partial x} + \frac{\partial A}{\partial t} = 0 \tag{1}$$

$$\frac{\partial Q}{\partial t} + \frac{\partial}{\partial x} \left(\frac{Q^2}{A} \right) + gA \frac{\partial y}{\partial x} - gA(S_0 - S_f) = 0 \tag{2}$$



$Q =$ discharge [m^3s^{-1}]

$A =$ flow cross sectional area [m^2]

$S_0 =$ channel slope

$S_f =$ friction slope

$y =$ flow depth [m]

$x =$ distace along channel [m]

$t =$ time elapsed [s]

$g =$ gravitational constant [ms^{-2}]

$$Inertia Term = \frac{\partial Q}{\partial t} + \frac{\partial}{\partial x} \left(\frac{Q^2}{A} \right)$$

$$P.G. Term = gA \frac{\partial y}{\partial x}$$

$$Kinematic Term = -gA(S_0 - S_f)$$

A heterogeneous routing scheme capable of transitioning between routing methods would eliminate the need to choose an absolute routing method. Determining the transition point, or threshold of potential error accrued by use of Diffusive or Kinematic wave models, presently requires computing the solution to the full momentum equation, rendering routing method transition ineffectual. A computationally inexpensive surrogate, or set of surrogates, associated with relative importance of the

inertia, P.G., gravity and friction is a desirable solution to this problem.

2. Previous Studies

An approach using seven Dimensionless Scaling Parameters (DSPs) to differentiate between various river wave types was introduced by [3]. The DSPs, when written in a dimensionless form, indicate the relative importance of friction, inertia, and P.G. effects on a river wave. These DSPs are presented in **Table 1**, with their physical interpretation presented to show how each highlights part of the friction-inertia balance. While previous work depended on case study measurements to construct a framework of how each DSP corresponds to a specific wave type (Bulk, Dynamic, and Gravity waves), the study that follows utilizes a modeling framework. This framework allows for a systematic inclusion of physical river and wave characteristics to identify the ranges of DSP values that correspond to transitions between the Kinematic, Diffusive, and Dynamic Wave.

Table 1. Description of the seven dimensionless scaling parameters taken from literature [1]. The physical meaning of each term is presented in the context of differentiating between wave types.

DSP	Equation	Physical Interpretation
C_r	$C_r = v_0 \left(\frac{\Delta t}{\Delta x} \right)$	Ratio of mean flow velocity to measured wave celerity
F_0	$F_0 = \frac{v_0}{\sqrt{g y_0}}$	Ratio of surface wave to mean flow velocity
S	$S = \frac{S_0}{S_f}$	Ratio of channel bed slope to energy gradient
D_I	$D_I = \left(\frac{C_r}{F_0} \right)^2$	Ratio of surface wave to measured wave celerity
F_I	$F_I = \frac{2C_r}{(C_*)^2} \left(\frac{k \Delta x}{y_0} \right)$	Influence of friction effects on river flow
F_c	$F_c = F_I C_r$	Influence of channel bed slope on river flow
D	$D = \frac{D_I}{F_c}$	Ratio of wave diffusion to wave advection
	$v_0 =$ Average velocity (ms^{-1})	
	$y_0 =$ Average depth (m)	
	$\Delta x =$ spatial discretization (m)	
	$\Delta t =$ temporal discretization (s)	
	$k =$ water surface parameter [1: open water, 2: ice covered]	
	$C =$ Chezy coefficient ($\text{m}^{1/2}\text{t}^{-1}$)	
	$C_* = g^{-1}C^2 =$ dimensionless Chezy coefficient	

3. Objectives and Scope

The main goals of this research are as follows:

1. Ascertain that the Dynamic and Diffusive Wave models are not required to obtain an accurate approximation of the solution to the full momentum equation at all locations along a channel and all realizations within a designated time frame.
2. Define relationships between DSPs and relative magnitudes of inertia, P.G., and kinematic terms to determine potential thresholds that may serve as a transition signal in a heterogeneous routing scheme.

This study further explores the relationship between DSPs presented in [3] and the terms of the momentum equation by generating data from a series of modeled scenarios, or experiments, using a theoretical channel. Experiments cover a range of typical channel characteristics and flow conditions that fall under M1 profiles. Extreme scenarios, including adverse slopes, contractions and expansions, and chutes are not represented by the data generated for this study. Details about data generation and parameter ranges can be found in the Methodology section. Processed data was then used to visually and statistically relate DSP values to the relative significance of individual momentum terms.

4. Methodology

This study implements a workflow to investigate the relative significance of the terms of the momentum equation, and if a relationship exists between those terms and DSPs. The workflow was purposely created with the potential of being reconfigured into a heterogeneous routing scheme decision tool for use in an operational environment, such as with the NWM. For this reason, a modularized approach was taken to lay out the developmental and operational workflows seen in **Figure 1**. The modules that comprise the developmental workflow are the subject of the rest of this study.

To study if the Dynamic or Diffusive Waves are always required, the developmental workflow (**Figure 1**, left pane) is needed to automate the generation of momentum term and DSP data for a wide range of scenarios. A data generator module was created to produce a range of initial flow conditions and channel characteristics. The produced datasets served as inputs into the MESH model (a numerical model that solves the full Dynamic Wave equation, [4]), which was also automated through a simulation module. Automation of data generation and model simulation allowed for hundreds of scenarios to be considered. MESH output was then fed into a pair of process modules to calculate the momentum terms and DSPs. A final relation module was created to investigate if relationships existed between momentum terms and DSPs, both visually and statistically. Each one of these modules is explained in further detail.

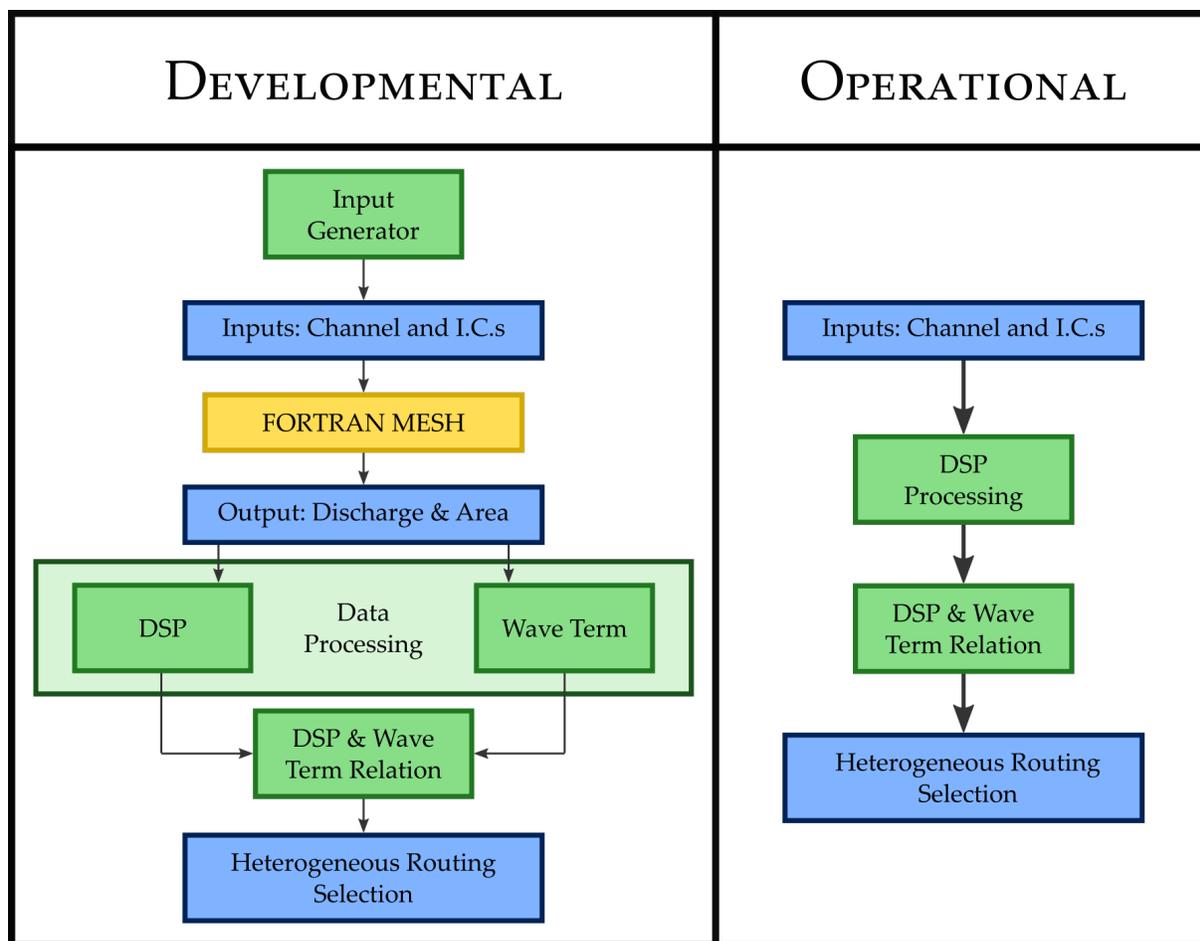


Figure 1. Developmental workflow (left) for utilizing the MESH model to identify relationships between wave types and DSPs, with the operational workflow (right) being able to use DSPs as a decision tool for determining sufficient wave type in a heterogeneous routing scheme.

4.1. Experimental Setup & Input Data Generation Module

The study's experimental setup added simulation automation features to the MESH model. The MESH model solves the full momentum equation for a single scenario with a specific set of channel and flow conditions by using the predictor-correction numerical scheme [4]. To guarantee that the version of the model used for this study simulated as expected, the MESH model was first run for scenarios in literature. Successful recreation of the scenarios seen in **Figure 2** provided sufficient confidence in the model that it could be used for experimental purposes.

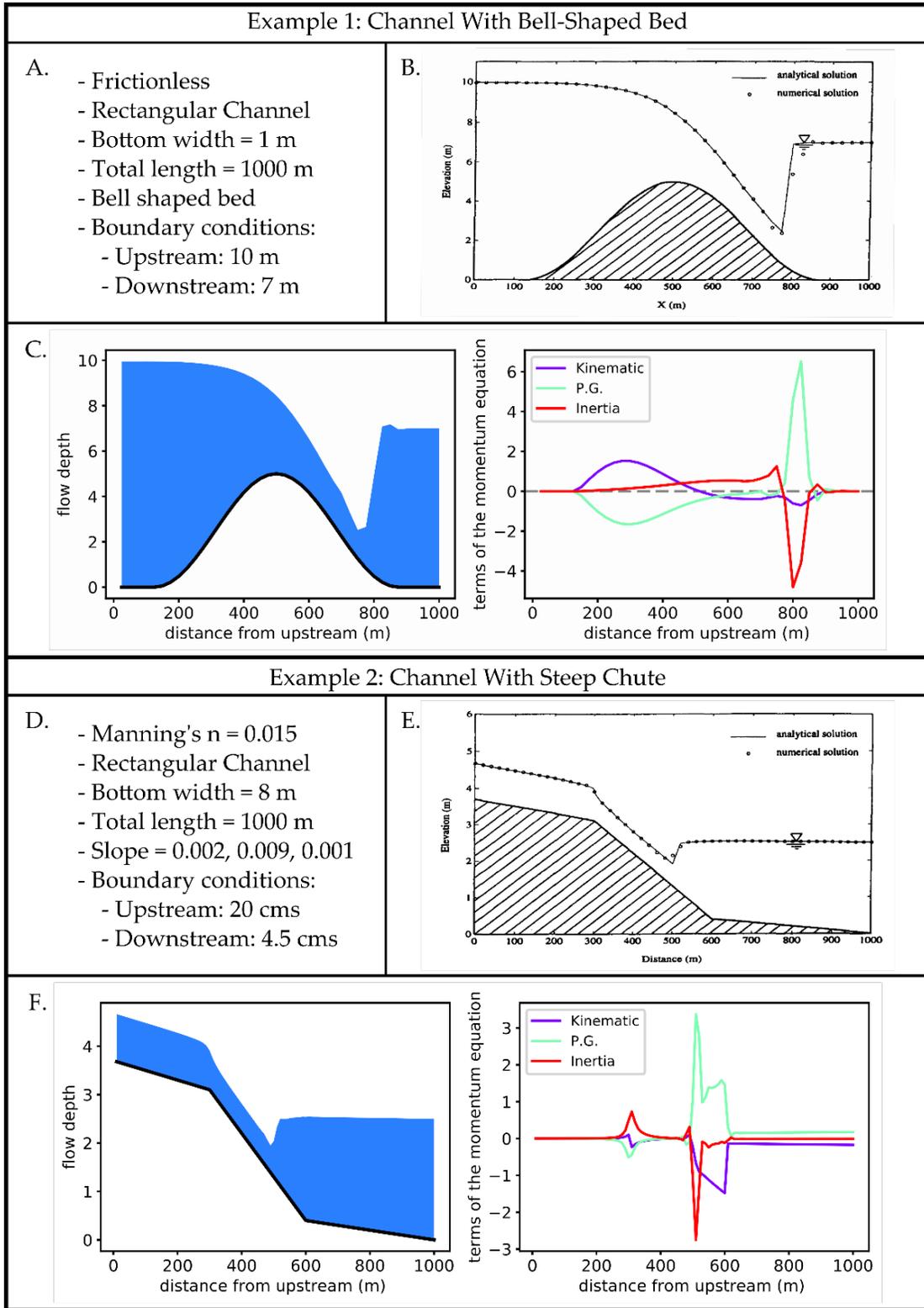


Figure 2. Examples (A, B and D, E) from literature of the MESH model being used [2]. The examples were recreated (C and F) as part of this study to confirm that the version of the MESH model used was simulating scenarios as expected. The physical setup of the two scenarios are explained in panes A and D.

Data for this study was generated by modeling flow in a theoretical, rectangular channel using the MESH model. Channel dimensions, length and bottom width, as well as spatial and temporal discretizations were constant throughout all experiments. Variable parameters, consisting of channel slope (slope), Manning's n (n), volumetric flow rate (Q), and downstream water depth (to control the amount of backwater), were systematically varied across a range of values for a set of experiments with the intent to capture a representative range of flow profiles. Only one variable parameter was altered per experiment, e.g. channel slope was varied for a set of experiments while n and Q remained constant. This approach provided a controlled means of altering channel and flow conditions with the ability to easily trace errors to the source. Values for variable parameters and constant parameters used in the experiments are shown in **Tables 2 and 3**.

Table 2. Values for variable parameters used to run experimental scenarios through the MESH model. Downstream stage was set as a multiple of normal depth (ND). Default values of slope and $1/n$ are 5×10^{-4} , and 33 respectively. Default steady flow is $100 \text{ m}^3/\text{s}$ and default unsteady flow is $500 \text{ m}^3/\text{s}$.

Flow State	Back Water	Downstream Stage (m)	Range of Changing Variable		
			Q (m^3/s)	Slope	$1/n$
Steady	Yes	2,3,4×ND	100-900	-	-
	Yes	2,3,4×ND	-	5×10^{-5} - 1×10^{-3}	-
	Yes	2,3,4×ND	-	-	10-90
	No	ND	100-900	-	-
	No	ND	-	5×10^{-5} - 1×10^{-3}	-
	No	ND	-	-	10-90
Unsteady	Yes	2,3,4×ND	100-700	-	-
	Yes	2,3,4×ND	-	5×10^{-5} - 1×10^{-3}	-
	Yes	2,3,4×ND	-	-	10-90
	No	ND	100-700	-	-
	No	ND	-	5×10^{-5} - 1×10^{-3}	-
	No	ND	-	-	10-90

Table 3. Values for constant parameters used to run experimental scenarios through the MESH model.

Channel Width (m)	Channel Length (m)	Spatial Discretization, dx (m)	Temporal Discretization, dt (s)
100	10000	20	60

Experiments were designed for both steady and unsteady flow. Unsteady flow hydrographs were generated using a skewed gaussian distribution, changing only the peak flow between experiments. Time-steps of steady flow were added to the beginning of unsteady hydrographs to ensure sufficient model spin-up time prior to hydrograph initiation. Individual time-steps within unsteady flow experiments were treated as separate instances of a flow profile. Flow profile designations were assigned during post-processing by calculating Froude number and critical slope using MESH outputs

for each time-step. Due to time and data-processing constraints, adverse profiles were excluded from this study.

4.2. Data Processing Module

The simulation output from the MESH model as well as channel characteristics for experimental scenarios are fed, in parallel, into two post-processing modules. These two modules use discharge and cross-sectional area of flow generated from the MESH model and calculate the individual terms of the full momentum equation based on equation 2 and the DSPs (**Table 1**).

In steady flow scenarios, only values at the last time step are used, whereas in unsteady flow scenarios, all time steps from the start to the end of the propagation of the unsteady flow along the channel are considered. The spin-up period at the start of the simulations and redundant steady state time steps after the passage of the discharge wave through the channel are removed from the analysis for unsteady flow cases.

For all realizations (points in time and space) used in the analysis the following adjustments are applied to momentum terms:

Steady Flow:

- Condition SF-1: Due to rounding approximations, the inertia terms may assume very small values for steady state solutions. Therefore, these terms were forced to be equal to zero for all steady flow scenarios as part of the post processing module.
- Condition SF-2: When the magnitude of the channel bed slope and friction slope (S_0 and S_f in equation. 2, respectively) are within 1% of each other, the flow is considered to be uniform and all terms of the momentum equation are forced to be equal to zero.

Unsteady Flow:

- Condition SF-2 is applied
- Condition UF-1: When P.G. and Kinematic terms have different signs and their magnitudes are within 1% of each other, the inertia term is considered to be negligible and set to 0.

The final product from the post-processing modules are saved as individual matrices for visual and statistical analysis.

4.3. Visual Relation Module

To identify if relationships between momentum terms and DSPs existed, the matrices outputted from the calculation modules were condensed into a single population and visualized. Visualizing the population was executed in two steps. The first depicted when the Kinematic Wave can be considered sufficient by highlighting when the P.G. and inertia terms both equal zero in a 3-dimensional space. The third dimension in this space is a DSP to visualize if transitions between the wave types corresponded to ranges of DSP values.

The second approach depicted the transition between the Diffusive and Dynamic Waves. The P.G. and inertia terms of the Saint-Venant equation were normalized by the kinematic term which produced an equation with the sum of the two ratios equal to negative one. When the ratio of P.G. to kinematic equaled negative one, it was known that the inertia term was zero. The increasing significance of the inertia term was then captured as the ratio of P.G. to kinematic term moved in either direction away from negative one to positive or negative infinity.

4.4. Statistical Relation Module

Several rounds of Principal Component Analysis (PCA) were performed on a 50% random sampling of the unsteady flow dataset. PCA re-projects data points onto axes that maximize variance within the data; correlating variables with axes that explain the greatest amount of variance can be used to interpret the relative significance of variables and inter-variable. PCA was performed on several combinations of DSPs and relative momentum terms (variables) to isolate relationships between DSPs and relative magnitudes of individual momentum terms, as well as to assess the relative influence of the inertia term. Relative momentum terms were calculated by normalizing the magnitude of each individual term by the sum of the absolute magnitudes of all three momentum equation terms. The first set of variables included all 7 DSPs and a relative inertia terms, the second set of variables included all 7 DSPs and a relative P.G., and the third set of variables included all 7 DSPs and all three relative momentum terms. Correlations between Principal Components (PCs) and variables were then used to identify DSPs that may offer similar information and relate variance of momentum terms to DSPs.

5. Results

5.1. Momentum Terms

While the magnitudes of Kinematic and P.G. terms have similar distributions, the inertia term magnitudes are significantly lower. This highlights the fact that in most cases, the Dynamic Wave equation is not required to accurately explain the system. The distribution of momentum terms for steady and unsteady flow are shown in **Figures 3 and 4**, respectively.

The influence of channel and flow characteristics on relative magnitude of the momentum terms is explored for each set of steady and unsteady flow simulations explained in **Tables 2 and 3**. **Figure 5** shows an example of the sensitivity of momentum terms to channel roughness for unsteady flow simulation without backwater effect at one time step. For all steady flow simulations, the values of the P.G. term are comparable with the kinematic term for the portion of the channel where backwater occurs. In these cases, the Kinematic Wave equation alone does not accurately explain the flow profile. The portion of the channel upstream of the backwater, however, has characteristics of uniform flow for which the Kinematic Wave equation is sufficient. The magnitude of kinematic and P.G. terms increase with increase in upstream discharge, larger backwater effects, higher values of Manning's n (more surface roughness) and steeper slopes.

For unsteady flow scenarios, the magnitude of the inertia term increases with increase in upstream discharge, steeper slopes, and decrease in Manning's n . It should be noted that the MESH model does not produce accurate results for the inertia term around sharp curvatures in water surface due to the activation of the artificial diffusion term [4]. Overestimation of the relative importance of the inertia term around regions where the water surface slope changes (e.g. transferring from uniform to backwater) is an artifact of this inaccuracy. Overall, the values of the inertia term are remarkably lower than those of kinematic and P.G. terms. The ratio of the inertia to kinematic (or P.G.) rarely exceeds 0.5 and is between 0-0.1 for majority of the realizations.

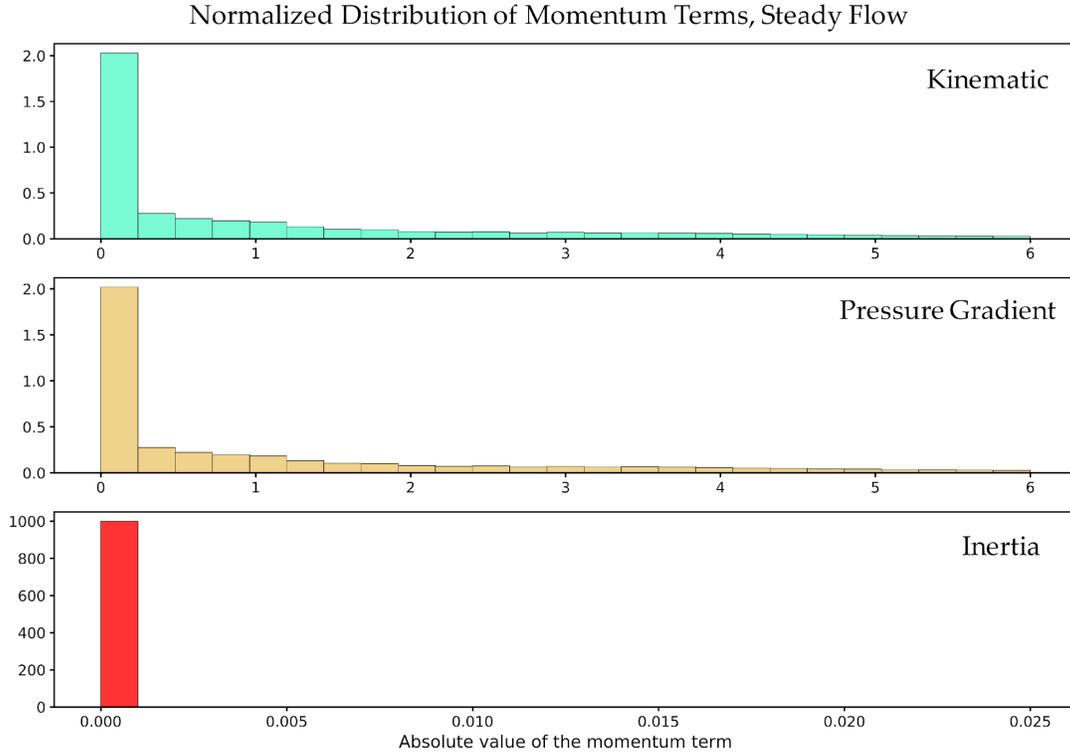


Figure 3. Distribution of momentum terms, steady flow.

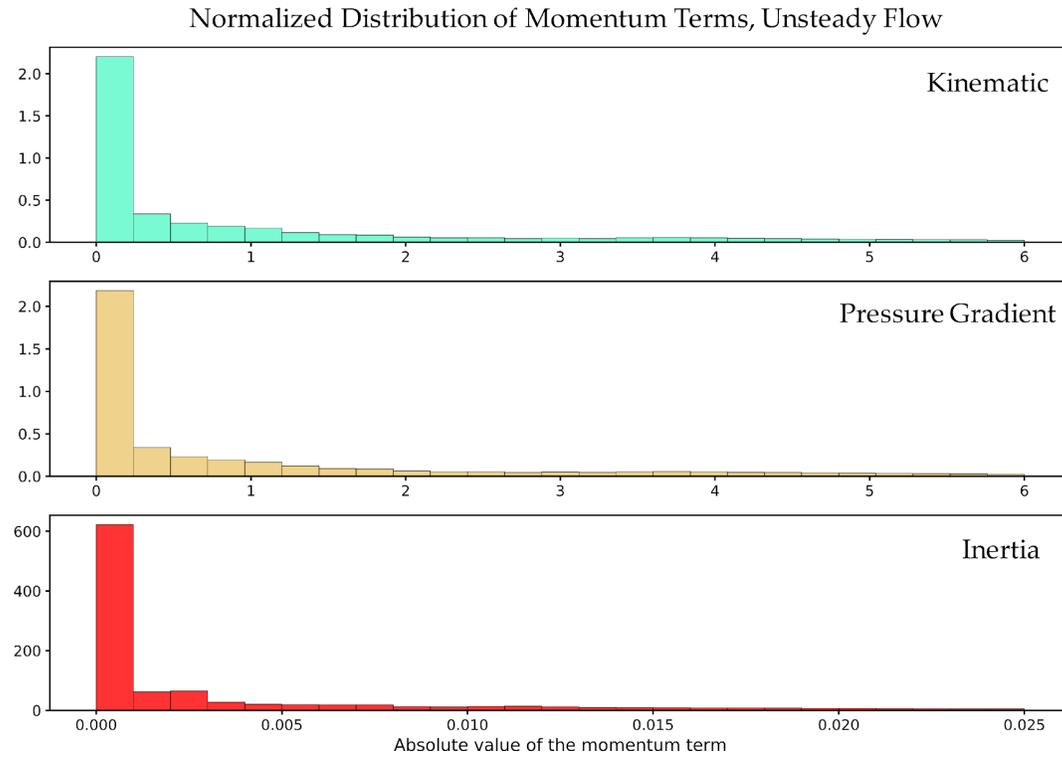


Figure 4. Distribution of momentum terms, unsteady flow.

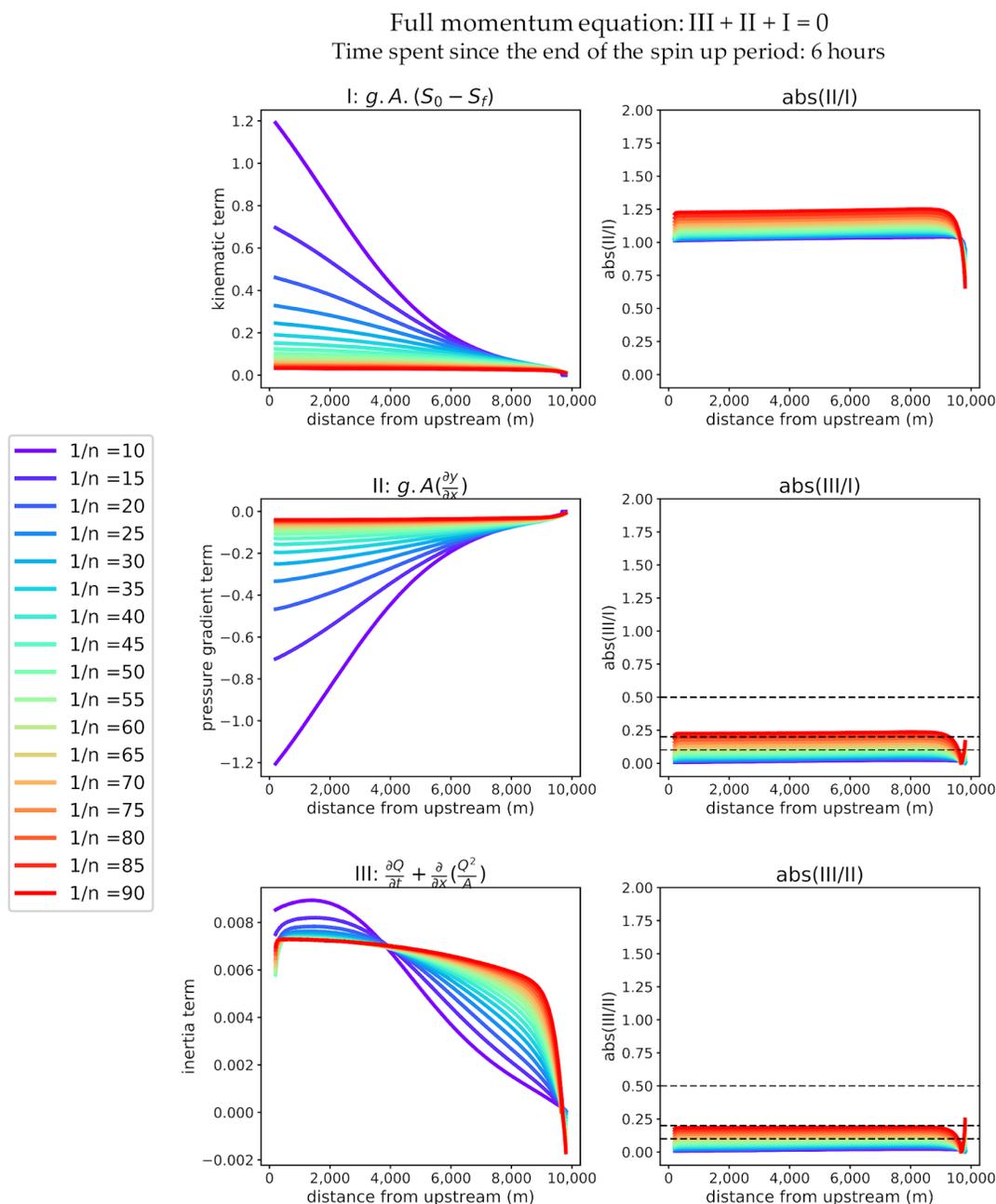


Figure 5. A snapshot of the sensitivity of momentum terms to Manning's n (surface roughness) for an unsteady flow scenario. Channel bed slope = 5×10^{-5} , channel width = 100 m, normal depth is forced at the downstream boundary.

5.2. Visual Analysis

Examples of visualizations from the momentum term and DSP processing module are shown in **Figures 6 and 7** with the Froude number. **Figure 6** shows the inertia and P.G. terms plotted against the Froude number for a 1% random sampling of all the unsteady simulations in this study (computational constraints require the plotting sub setting). The momentum terms are normalized as the ratio of the absolute value of the term itself to the summation of the absolute values of each term. The Froude number is normalized by itself to range from 0 to 1. The figure shows that the percentage of points where both the inertia and P.G. terms are zero is $\sim 22\%$. These points are representative of

a routing scenario where the Kinematic Wave would be considered sufficient. Additionally, for $\sim 74\%$ of the samples, the inertia term is less than 10% of the summation of the absolute values of the terms, while the P.G. term is nonzero. These can be representative of instances where the Diffusive Wave would be preferential over the Dynamic Wave, given a user designated acceptable error level. Both these support that the Dynamic and Diffusive Waves are not always necessary to achieve routing accuracy within a preset acceptable error range of the full momentum equation solution.

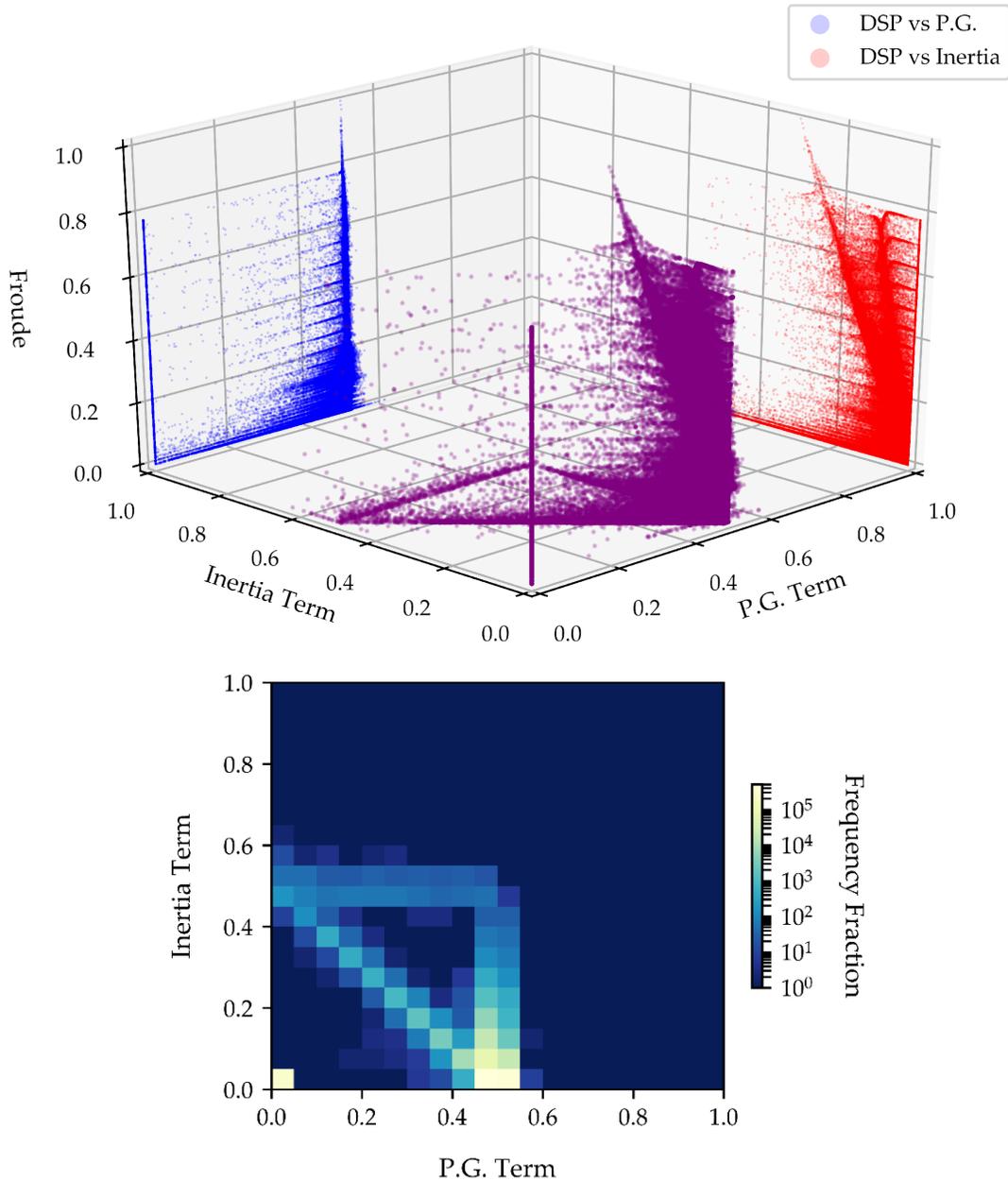


Figure 6. A visualization of the inertia term, P.G. term, and a DSP (Froude number) in 3-dimensional space (purple points, top plot). The 3D points are projected onto 2D planes with blue points being the Froude number versus the P.G. term and red points being the Froude number versus the inertia term. Bottom plot is a 2D histogram of the inertia and P.G.

Figure 7 shows a plot of the ratio of the P.G. to kinematic term (PGtoK) and histograms of the Froude number when this PGtoK momentum term ratio is greater than, near, and less than negative one. The behavior of the PGtoK ratio with respect to negative one is analyzed because a ratio value of negative one is representative of the inertia term equaling zero. **Figure 7** highlights this relation with the value of the PGtoK ratio to depict when the Dynamic Wave can be deemed unnecessary. It is also seen in **Figure 7**, that the DSP histograms when the PGtoK ratio is greater and less than -1 is different from the DSP histogram corresponding to when PGtoK equals -1. These histograms are proposed to represent threshold values of DSPs at the transition between Dynamic and Diffusive Wave. It is important to note that these results are sensitive to numerical errors in the MESH model for values near zero being exposed through division of terms and, thus adding numerical uncertainty to the histograms.

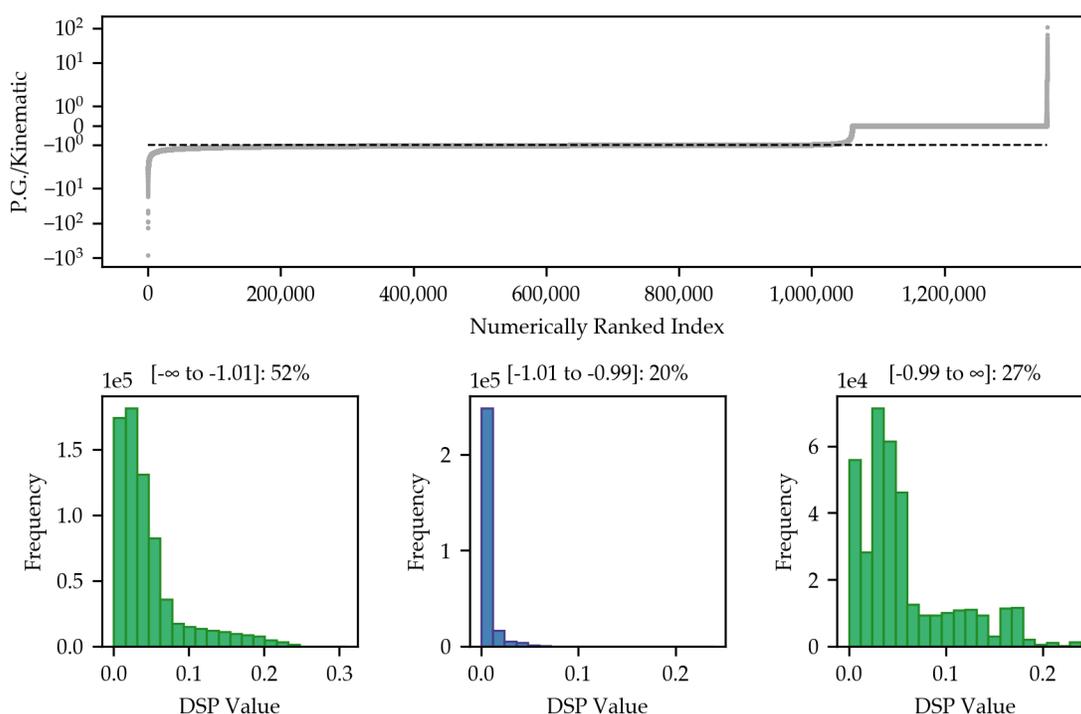


Figure 7. A visualization of the decision between using Dynamic versus Diffusive/Kinematic waves. The bottom panels show histograms of a DSP, the Froude number, for different ranges of the ratio of the P.G. term to the kinematic term shown in the top panel. The ratio was calculated for a random sampling of 1% of the unsteady simulations and then ranked numerically. The histogram of DSP values corresponding to when the P.G. to kinematic ratio is near -1 (bottom center, blue bars) is presented to be compared to the histograms of DSP values when the P.G. to kinematic ratio is both greater (bottom right) and less than (bottom left) -1 respectively.

5.3. Statistical Analysis

5.3.1. PCA: DSPs and Relative Inertia Term

DSPs 1, 2, and 4 exhibited strong correlations with PC1 while DSPs 3 and 6 (**Table 4**) showed strong correlations with PC2 ($r > 0.70$), indicating that these sets of DSPs likely vary together. This result is reasonable considering that DSP1 and DSP2 are functions of velocity, and DSP4 is a function of DSP1. The relative inertia term showed a moderate correlation ($r = \sim 0.50$) with PC1 but a much stronger correlation ($r = \sim 0.80$) with PC4. All of the DSPs however, showed weak correlations ($r <$

0.30) with PC4. This may indicate that, while the DSPs 1,2, and 4 may be able to serve as a signal for when the relative magnitude of the inertia term becomes significant or insignificant enough to be included or ignored, some variance is left unexplained by the DSPs and their signal may fail in some scenarios.

Table 4. Correlation values (r) between PC1-PC4 and All DSPs and the relative inertia term. PC5-PC8 all had correlations less than 0.2 and were removed. Green, bold format denotes $r \geq 0.50$.

	DSP1	DSP2	DSP3	DSP4	DSP5	DSP6	DSP7	Inertia
PC1	0.79	0.75	0.03	0.77	0.67	0.15	0.61	0.49
PC2	0.25	0.38	0.89	0.22	0.27	0.91	0.25	0.37
PC3	0.36	0.40	0.35	0.11	0.66	0.41	0.70	0.28
PC4	0.22	0.28	0.11	0.10	0.03	0.03	0.07	0.79

5.3.2. PCA: DSPs and Relative Pressure Term

DSP 4 exhibited a strong correlation with PC1 ($r = \sim 0.80$), while the remaining DSPs showed moderate ($0.49 < r < 0.69$) to weak ($r < 0.49$) correlations with PC1 (**Table 5**). The relative pressure term exhibited similar behavior to the relative inertia term, showing a moderate correlation ($r = \sim 0.52$) with PC1 and a strong correlation ($r = \sim 0.87$) with PC4. All DSPs showed weak correlations with PC4, again showing that the axis explaining the majority of the variance between DSPs may not explain the variance of the individual momentum terms.

Table 5. Correlation values (r) between PC1-PC4 and All DSPs and the relative pressure term. PC5-PC8 all had correlations less than 0.2 and were removed. Green, bold format denotes $r \geq 0.50$.

	DSP1	DSP2	DSP3	DSP4	DSP5	DSP6	DSP7	P.G.
PC1	0.65	0.62	0.32	0.80	0.69	0.47	0.62	0.52
PC2	0.43	0.54	0.89	0.02	0.05	0.87	0.05	0.02
PC3	0.33	0.46	0.14	0.03	0.69	0.21	0.73	0.50
PC4	0.36	0.18	0.15	0.19	0.11	0.16	0.14	0.88

5.3.3. PCA: DSPs and Relative Momentum Equation Terms

When all relative momentum terms were included in the PCA with the DSPs, relative kinematic and P.G. terms showed very strong correlations with PC1, while the relative inertia term exhibited a very weak correlation ($r = 0.01$, **Table 6**). Although no DSPs were strongly correlated with the PCs, DSPs 1,2, and 4 showed the strongest correlations with PC1, with r values between 0.56–0.65. The relative inertia term, DSPs 3 and 6 showed moderate correlations with PC2 ($r = 0.62$ – 0.65). The majority of the variance was explained by PC1 and PC2, with which both the momentum equation terms and the DSPs showed at least moderate correlations. This shows potential for DSPs to be statistically related to the relative importance of the momentum terms, however, this may require regression analysis or non-linear statistical analysis to define the relationship.

Table 6. Correlation values (r) between PC1-PC4 and All DSPs and all relative momentum equation terms. PC6-PC10 all had correlations less than 0.2 and were removed. Green, bold format denotes $r \geq 0.50$. (Kine = Kinematic).

	DSP1	DSP2	DSP3	DSP4	DSP5	DSP6	DSP7	Kine	P.G.	Inertia
PC1	0.56	0.65	0.25	0.65	0.42	0.40	0.35	0.96	0.95	0.01
PC2	0.56	0.56	0.65	0.23	0.25	0.62	0.24	0.28	0.36	0.65
PC3	0.35	0.32	0.42	0.15	0.54	0.46	0.58	0.17	0.22	0.45
PC4	0.00	0.22	0.53	0.37	0.63	0.53	0.62	0.48	0.49	0.14
PC5	0.26	0.15	0.08	0.25	0.05	0.01	0.10	0.16	0.17	0.64

6. Concluding Remarks

6.1. Conclusion

This work provides a framework for synthetic sample generation, visualization of individual terms of the full momentum equation and DSPs, and the statistical bridging of the two. Analysis of data confirmed that instances exist where the full Dynamic or Diffusive Wave routing methods do not significantly improve computational accuracy. While defining an acceptable level of error at which we may forgo Dynamic or Diffusive Wave for more simplified methods is beyond the scope of this study, using an arbitrary threshold of the Inertia term being less than or equal to 10% of the cumulative magnitude of the momentum terms results in only ~5% of the synthetic data required the full Dynamic Wave. With this arbitrary threshold, Diffusive and Kinematic Wave provide sufficient accuracy for the remaining scenarios. It should be noted that the sample generated in this study is not necessarily representative of the population, however, the framework developed can easily handle a larger, more representative sample in future work.

As can be seen in the DSP histograms, differences in DSP magnitudes exist near and away from the transitions of the Diffusive and Dynamic Wave, specifically for the Courant and Froude number. It is proposed that these differences in DSP values near and away from the Dynamic/Diffusive transition can serve as a decision tool to determine a method in a heterogeneous routing scheme.

PCA analysis revealed that redundancy may exist within the current set of DSPs. Continued research may benefit from elimination of redundant DSPs from the set. The moderate to strong correlations of several DSPs and the relative momentum terms to PC1 and PC2 show the potential for statistically relating this set of DSPs to the relative importance of individual momentum terms. However, inclusion of DSPs that represent different properties of flow may provide new information required to define a relationship to momentum terms that may serve as a transition signal in an operational setting.

6.2. Aspirations for Future Work

- The current set of simulations cover primarily M1 profiles. Future experiments should encompass all possible profiles and extreme scenarios.
- Preliminary results from statistical analysis suggest that the current set of DSPs do not fully explain the variability of the momentum terms. Additional DSPs as well as statistical analysis should be considered for analysis of the data and relationships of this nature.

Acknowledgements: We would like to thank CUAHSI and the National Water Center for hosting this Summer Institute and providing a great opportunity for us to learn in a highly collaborative environment. We'd also like to thank our theme leaders, Ehab Meselhe and Kyle Mandli, for their endless guidance and support throughout this research. Lastly, thank you to our course coordinators, Fernando Aristizabal, Lauren Grimley, and Danielle Tijerina for their technical and intellectual support as well as their efforts to create a friendly and fun work environment.

References

1. Ponce, V.M. Unsteady Gradually Varied Flow. *Fundamentals of Open Channel Hydraulics*, Online ed. 2014
2. Ponce, Victor Miguel, Daryl B. Simons, and Ruh-Ming Li. "Applicability of kinematic and diffusion models." *Journal of the Hydraulics Division* 104.3 (1978): 353-360.
3. Ferrick, Michael G. "Analysis of river wave types." *Water Resources Research* 21.2 (1985): 209-220.
4. Meselhe, Ehab Amin, F. Sotiropoulos, and F. M. Holly Jr. "Numerical simulation of transcritical flow in open channels." *Journal of hydraulic engineering* 123.9 (1997): 774-783.

Chapter 6

An Approach for Incorporating Realistic Channel Geometry into Continental-Scale Hydrological Modeling

John Brackins¹, Nishani Moragoda² and Azbina Rahman³

¹Tennessee Technological University, jbrackins42@students.tntech.edu

²University of Alabama, npmoragoda@crimson.ua.edu

³George Mason University, arabma19@gmu.edu

Academic Advisors: Alfred Kalyanapu, Tennessee Technological University; Dr. Sagy Cohen, The University of Alabama; Dr. Viviana Maggioni, George Mason University

Summer Institute Theme Advisors: Dr. Christopher Lowry, The State University of New York at Buffalo, cslowry@buffalo.edu; Dr. Sagy Cohen, University of Alabama, sagy.cohen@ua.edu

Abstract: A key element in hydraulic and hydrologic modeling is the specification of representative channel geometry. Without adequate geometry information, it is difficult to reliably simulate hydraulic properties such as bankfull discharge and stage at which flooding commences. The traditional solution to the geometry problem has been topographic and bathymetric surveying, and floodplains have become increasingly resolved with the advent of lidar. For continental-scale hydrology and hydraulics, however, the large amount of high-resolution data required, as well as the considerable computational effort needed to effectively incorporate such data, has led to simplifying assumptions such as rectangular or trapezoidal channels for long river reaches. The National Water Model (NWM) uses the simplified trapezoidal channel representation for 2.7 million river reaches, over which it forecasts water discharge for the entire continental United States. This has created uncertainties in when to initiate hydraulic predictions. The aim of this study is to: 1) evaluate the NWM predictions with the current trapezoidal channel representation and with real channel geometry (using HEC-RAS), and 2) suggest an improved representation of channel geometry while maintaining parsimony. As a preliminary analysis, the HEC-RAS model outputs of discharge and stage were compared with the USGS observed records for three cases: trapezoidal, real, and proposed generalized geometry representations. A brief analysis of NWM Muskingum-Cunge routing parameters for varied geometry cases was also undertaken. Statistical analyses show that more realistic channel geometry not only improves stage and flow predictions, but also improves simulated routing parameters, indicating the potential for geometric improvements to enhance the current NWM products.

1. Motivation

Accurate hydraulic and hydrologic modeling necessitates representative channel geometry. Without adequate geometry information, it is difficult to reliably simulate hydraulic properties such as bankfull discharge and stage at which flooding commences. The traditional small-scale solution to the geometry problem has been topographic and bathymetric surveying, and while floodplains have become increasingly resolved with the advent of lidar, channel bathymetry collection has lagged markedly

behind. Additionally, for continental-scale hydrology and hydraulics, the daunting amount of high-resolution data required, as well as the massive computational effort needed to effectively incorporate such data when it is even available has precluded its use in current large-scale studies [1,2]. The combination of these factors has led to simplifying assumptions such as rectangular or trapezoidal channels for long river reaches when the focus of a model is hydrology-driven rather than hydraulics-driven, e.g. [3,4].

The National Water Model (NWM) [5] uses the simplified trapezoidal channel representation for 2.7 million river reaches, over which it forecasts water discharge for the entire continental United States (CONUS). The bankfull width and cross-sectional area of the trapezoidal channel is developed using regional curves that relate contributing drainage area to the bankfull characteristics [6]. Then, average depths and bottom widths are calculated using mathematical relationships keeping the cross-sectional area constant. The channel side slopes are currently assumed to be 1H:20V nationwide [5]. As a result of this simplistic representation of the channel geometry, the NWM does not generate a reliable stage product and also suffers from the assumption of “infinite” channel depth, which makes it impractical to represent overbank flow appropriately. Since the NWM-simulated stage “product” is used for Muskingum method flow routing, errors in simulated stage affect time-varying properties such as flow area and hydraulic radius, which are used to estimate Muskingum routing parameters K and X on-the-fly. Clearly, flow routing could be adversely affected if the stage “product” is significantly affected by the simplistic channel representation. Therefore, it is of paramount importance to determine if the simplistic trapezoidal channel representation generates a reliable discharge and stage product in continental scale hydrological modeling and whether incorporating real channel geometry makes a significant improvement.

To adequately represent real channels, it is vital to obtain high-resolution (<20m) data within and near channels at a relatively high temporal resolution due to dynamic changes during high-flow events, but this is currently an onerous task. In order to obtain various kinds of high-resolution datasets, as a supplement to professional field surveying, many agencies are turning to citizen scientists for data collection [7]. Considering the expansion of the application of citizen science to different sectors of research, we consider the potential for collecting and updating high-resolution channel geometry data by the help of citizen scientists using a relatively simple method.

2. Objectives and Scope

The overarching goal of this study is to assess the effects of incorporating real channel geometry into continental-scale hydrological models and propose a generalized cross-section which adequately captures important aspects for modeling. The specific objectives of the study are:

- To simulate flow and stage using the Hydrologic Engineering Center’s River Analysis System (HEC-RAS) for both real and NWM channel geometry and compare the results with United States Geological Survey (USGS) observed gage data in different streams across the CONUS to evaluate the maximum expected potential for improvement.
- To evaluate the potential improvements in flow routing and stage prediction made possible by the proposed generalized cross-section.
- To determine potential methods for citizen scientists and others with little to no surveying experience to obtain a reasonably accurate cross-section representation using commonly available tools.

3. Previous Studies

Simplistic channel geometry assumptions are convenient when much of the data required to avoid

such assumptions is not readily available. An entire subfield of hydraulic geometry has arisen to aid in synthesizing geometry data when needed, which has largely focused on power-law relationships between streamflow and river mean depth, mean velocity, and width [8]. Some studies, chiefly [9], have lamented the focus on curve-fitting coefficients for these power-law relationships, arguing instead the importance of viewing hydraulic geometry as a dynamic equilibrium between erosive forces of moving water and resistive forces in the channel bed. Either approach requires large amounts of data which is currently expensive to produce, limiting their potential for use on a continental scale.

Recent advances in remote sensing, particularly the advent of the Surface Water and Ocean Topography (SWOT) [10] mission, hold potential for building datasets of channel geometry data, especially river flow width. Methods have been explored to use SWOT radar altimetry products to estimate river bathymetry and slope [11], as well as discharge [12]. While these methods show substantial promise, the SWOT mission was understandably designed with larger-scale features in mind, i.e. river widths more than 50 meters [10]. Many smaller headwater streams would not be reliably detected, so while SWOT can aid in building knowledge in larger rivers, it may not be the most appropriate tool in low-order streams. Implementing citizen science methods for collecting channel geometry may be uniquely suited for building understanding of lower-order streams.

Citizen science and community-based monitoring programs are increasing in number and breadth, generating volumes of scientific data. Specific to the field of hydrology, measurements such as stream stage are being increasingly collected by crowdsourcing. It is a viable tool for collecting distributed measurements of stream stage based on both the ease with which stream stage can be measured and the ubiquity of mobile phones. Citizen scientists can send hydrologic measurements via text message to a server that stores and displays the data on the web [7]. There are also different levels of engagement required for a citizen scientist to participate; generally, the more difficult or involved the task, the less likely it will attract enough interest to be a viable data source [7].

The authors are currently exploring methods to obtain channel geometry that would require minimal effort on the part of a citizen scientist, such as sending a photograph of a river with a staff gage and markers from which width could be estimated, to build datasets of simultaneous river stage and width. For the topic of this report, however, a more-involved approach requiring wading of small streams is suggested in order to test the potential for citizen science channel geometry data of this kind. Other methods for developing the required data are also explored briefly.

4. Methodology

In order to evaluate potential improvements to simulations of stage and flow based on more realistic channel geometry, it is important to first establish a baseline against which to compare new methods. To accomplish this end, the stage and flow simulation results based on NWM geometry assumptions were considered. It was also necessary to develop a best-case scenario for the potential effects of better channel geometry representations by using the best-available high-resolution bathymetry and topography data and considering the stage and flow results produced. These two cases bound the region of potential improvements possible through modifying channel geometry representations alone. If significant differences in stage and discharge errors are not observed, this would indicate that efforts to improve hydrologic and hydraulic models may be better spent on other parameters such as channel roughness or bed slope.

4.1. Model Selection

Several features would comprise the ideal model for performing this kind of evaluation: the model code should be capable of resolving small-scale channel geometry data efficiently, it should be publicly

available, and it should be capable of resolving hydraulic structures such as culverts and bridges (particularly because the comparison gage data is largely available only at bridges). Due to the ubiquity of HEC-RAS models nationwide because of FEMA flood insurance studies and the public availability through the FEMA Engineering Library [13], the requirement for accurate geometric input data in HEC-RAS (thus ensuring geometric data is available at any given site), and HEC-RAS’s abilities to solve the full Saint Venant equations and resolve hydraulic structures, HEC-RAS was selected as the main modeling tool for this study.

4.2. Site Selection

For each study site, two items of data were required in this study: a HEC-RAS model (as a source of geometry), and at least one USGS gage internal to the reach in HEC-RAS with a period of record for stage and discharge sometime within the 25-year retrospective run of the NWM (1993-2017). To identify potential study sites, the USGS Gages II shapefile [14] and the US Cartographic County Boundary dataset [15] were used to identify counties with varying numbers of USGS gage sites as an initial filter. For this study, the FEMA Hydraulic Studies for counties of interest were requested and compiled into a database. The associated geometry files and cross-section data were imported into a GIS to evaluate overlap between cross-section data and USGS gage coverage. The sites selected for the study are presented in **Table 1**.

Table 1. Selected HEC-RAS Study Sites

Stream	NWM Stream Order	USGS Gage ID	Contributing Drainage Area (USGS Gage, sq. mi.)
Middle Fabius River (Lewis County, MO)	4	05498150	141
Hackensack River (Rockland County, NY)	3	01376800	30.7
Cartoogechaye Creek (Macon County, NC)	4	03500240	57.1
Little Tennessee River (Macon County, NC)	4	03500000	140
North Second Creek (Rowan County, NC)	4	02120780	118
Big Canoe Creek (St. Clair County, AL)	5	02401390	141

4.3. Boundary Conditions and Model Parameters

As HEC-RAS models for FEMA flood insurance studies are typically completed using steady-state flow data, unsteady hydrographs had to be generated for each study site. For the vast majority of sites, only one USGS gage was available, meaning that if the USGS hydrograph data were to be used for independent comparison, it could not be used as forcing data also. Therefore, all hydrograph forcing

data came directly from the NWM 25-year retrospective run since it was the best-available data. The limitation of this approach is recognized, so all comparisons are prefaced with the NSE comparison between the NWM and the USGS gage site to denote areas where lack of improvements may have been the result of less-than-ideal forcing hydrographs.

Lateral inflows from tributary streams were specified throughout reaches at the nearest downstream cross-section. A potential weakness of RAS arises when a cross-section dries out during low-flow periods, which can introduce instabilities [16]. In some cases, minimum flows were specified to avoid this issue. The models were then tested so that the smallest minimum flow possible was used purely to stabilize the model runs. For unsteady runs, some additional assumptions about hydraulic structure behavior were made due to the scope of the study:

- Gate operations were not modeled due to lack of detailed information about operation rules. All gates were assumed as fully open during simulations.
- When geometry cases were changed, some bridge piers were ignored as needed to allow the model to run successfully. This should have minimal effect for the NWM geometry cases due to the relatively high expected stages quickly overtopping bridges completely. Bridge piers were included in the existing surveyed geometry in RAS.

4.4. Baseline and Realistic-Geometry Model Simulations

Several different cases of Manning's n values were run to avoid sensitivity to roughness values contaminating the results. The RAS model Manning's n values (with potential to vary across any given cross-section, especially in the overbank regions) were considered as the "real" values, and the NWM Manning's n cases (with a uniform Manning's n across any given cross-section) were also tested. For each of the selected study sites, four different cases were simulated with varying combinations of hydraulic geometry and Manning's n values:

- Surveyed real geometry and "real" Manning's n
- Surveyed real geometry and NWM Manning's n
- NWM channel geometry and "real" Manning's n
- NWM channel geometry and NWM Manning's n

In order to compare these representations with the observed USGS data and evaluate their performance, several statistical analyses, including Nash-Sutcliffe Efficiency (NSE), percent bias (PBIAS), and Root Mean Squared Error (RMSE), were carried out for both the complete 3-year simulation and selected peak events. Peak stage magnitude comparisons were also completed for the selected peak events. In order to evaluate the quality of the NWM hydrograph forcing, NSE was calculated between the USGS observed data and the NWM retrospective run for that particular stream link.

4.5. Proposed Cross-Section Representation

Currently, NWM channel hydraulics include only bottom width and side slope as channel parameters [5]. Our proposal is to improve the channel geometry by replacing the 2-parameter definition with a 7-parameter definition. This representation is shown in **Figure 1**, and the parameters are: the channel depths at b , c , and d (h_b , h_c , and h_d respectively) and the widths (W_{ab} , W_{bc} , W_{cd} , and W_{df}) between each sequential pair of points. The parameters define the location of 5 distinct points, of which two are assumed as the banks. The five-point representation assumes level bank elevations, so the 7 parameters are sufficient to represent an entire cross-section. Similar to the NWM geometry, the slope

extending left and right of the five-point representation is assumed to remain constant. However, unlike NWM geometry, this slope is not assumed as a constant 1H:20V, but it is instead based on the slope defined by the appropriate bank and the interior point closest to that bank (as an example, the slope to the left of the left bank at a is defined by the slope between points a and b). The 5-point generalization is adaptable to many scenarios, and it is capable of representing traditional assumptions such as rectangular and trapezoidal cross-sections with ease, ensuring backwards compatibility with existing NWM assumptions where better geometry data is not available. The main draw of this generalization is its adaptability to many representations, such as its ability to represent a natural cross-section in a curving river where a cut bank and point bar are likely to form (similar to cross-section B from [9], reproduced in **Figure 2**). The proposed generalization even has potential for representing a cross-section with two well-defined channels, helpful in reaches where braided channels are currently an issue.

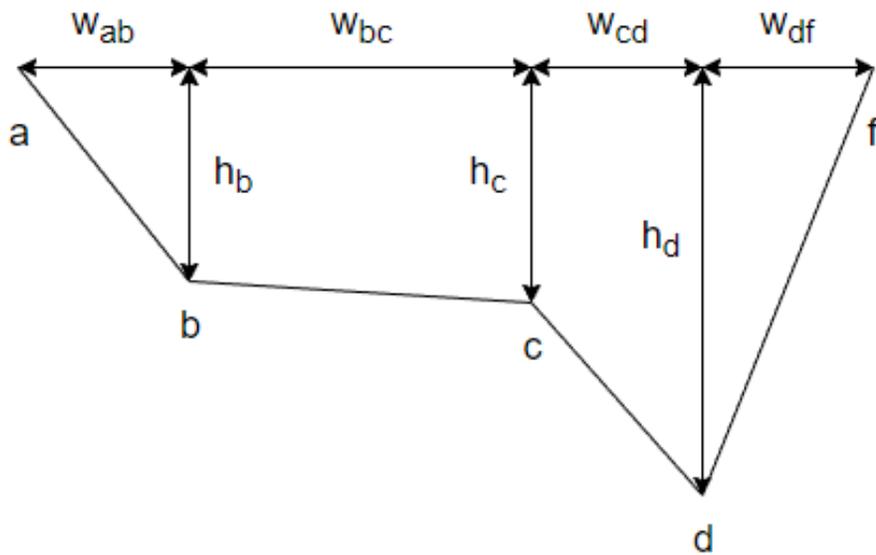


Figure 1. Five-point generalized cross-section representation.

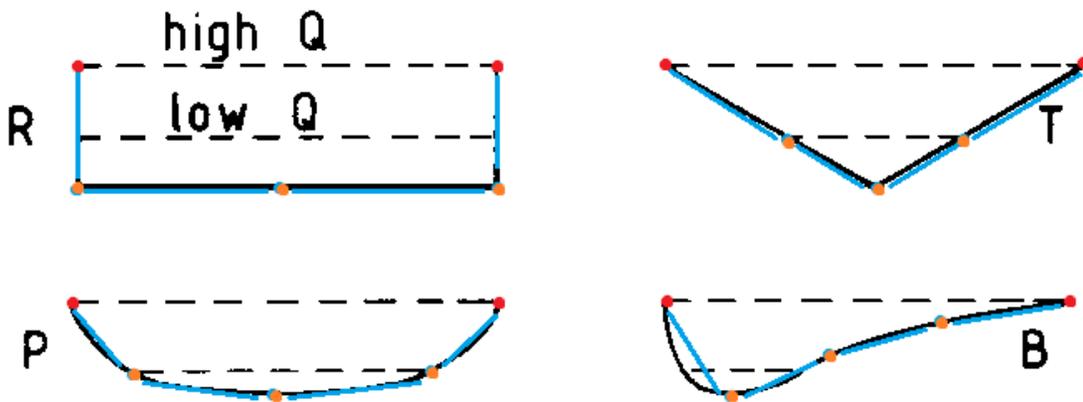


Figure 2. Characteristic cross-sections adapted from Ferguson 1986, annotated with five-point representations (red points indicate level banks, orange points indicate channel points).

To test the viability of the proposed five-point generalization, additional geometry cases were evaluated for the Big Canoe Creek site in St. Clair County, Alabama:

- 5-point generalized cross-section (including thalweg) and NWM Manning's n
- 5-point generalized cross-section (evenly-spaced points) and NWM Manning's n

Similar statistical comparisons were made for these additional two cases to evaluate potential improvements. The RAS Mapper tool in HEC-RAS was then used to visualize the differences in stage between: cases i) and iv) for both Hackensack River and Big Canoe Creek, cases iv) and v), and cases iv) and vi) for Big Canoe Creek. This was done by generating inundation extent maps using HEC-RAS simulated stage and Digital Elevation Models (DEMs) of the area downloaded from USGS.

4.6. Effects of Stage on Muskingum Routing Parameters

One of the key expected outcomes from improved channel geometry representations is improved flow routing. As the NWM currently updates Muskingum routing parameters K and X throughout a simulation based largely on the simulated stage, it follows that better prediction of stage during simulations will lead to improved estimates of these routing parameters and thus better flow routing. For the Big Canoe Creek site, an analysis of the effects on K was undertaken since link travel time is currently a known issue in the NWM. In order to evaluate K , some parameters were required: an estimate of the wave celerity $c_k = \frac{dQ}{dA}$, the simulated flow Q , and reach length dx . The reach length for the output cross-section was retrieved directly from the NHDPlus flowline network. The simulated flow was exported from HEC-RAS as a time-series. Within HEC-RAS, a plot of flow versus flow area was exported for a given cross-section and slope values were interpolated from this plot for given flow values to create a function for wave celerity. Due to small fluctuations in values, a cubic spline was fit to the data using the R package `stats` [17] prior to generating the wave celerity function. For the NWM geometry-NWM n and 5-point geometry-NWM n cases, the values of K were then calculated for each time step. While this method will not completely replicate the potential changes in K due to the use of the RAS-solver for the St. Venant equations instead of using actual Muskingum routing (and thus generating different flows at different times), it should provide some guidance as to general trends. Therefore, it is most appropriate to compare the two sets of routing parameter values at a given point for a given flow rather than inferring time-series trends.

5. Results

5.1. Summary Statistics and Improvements in Stage and Flow Simulation from Enhanced Channel Geometry

Figure 3 (a) and (b) shows the hydrographs for flow and stage in Hackensack River, NY for NWM geometry and real geometry with NWM Manning's n plotted against observed USGS data.

In general, the simulated stage in HEC-RAS is substantially higher than observed, due to the steep-sloped channel geometry representation. While flow also appears to be mainly over-predicted, this is due to the relatively poor hydrograph forcing from the NWM (NSE = -2.48, see **Table 2**).

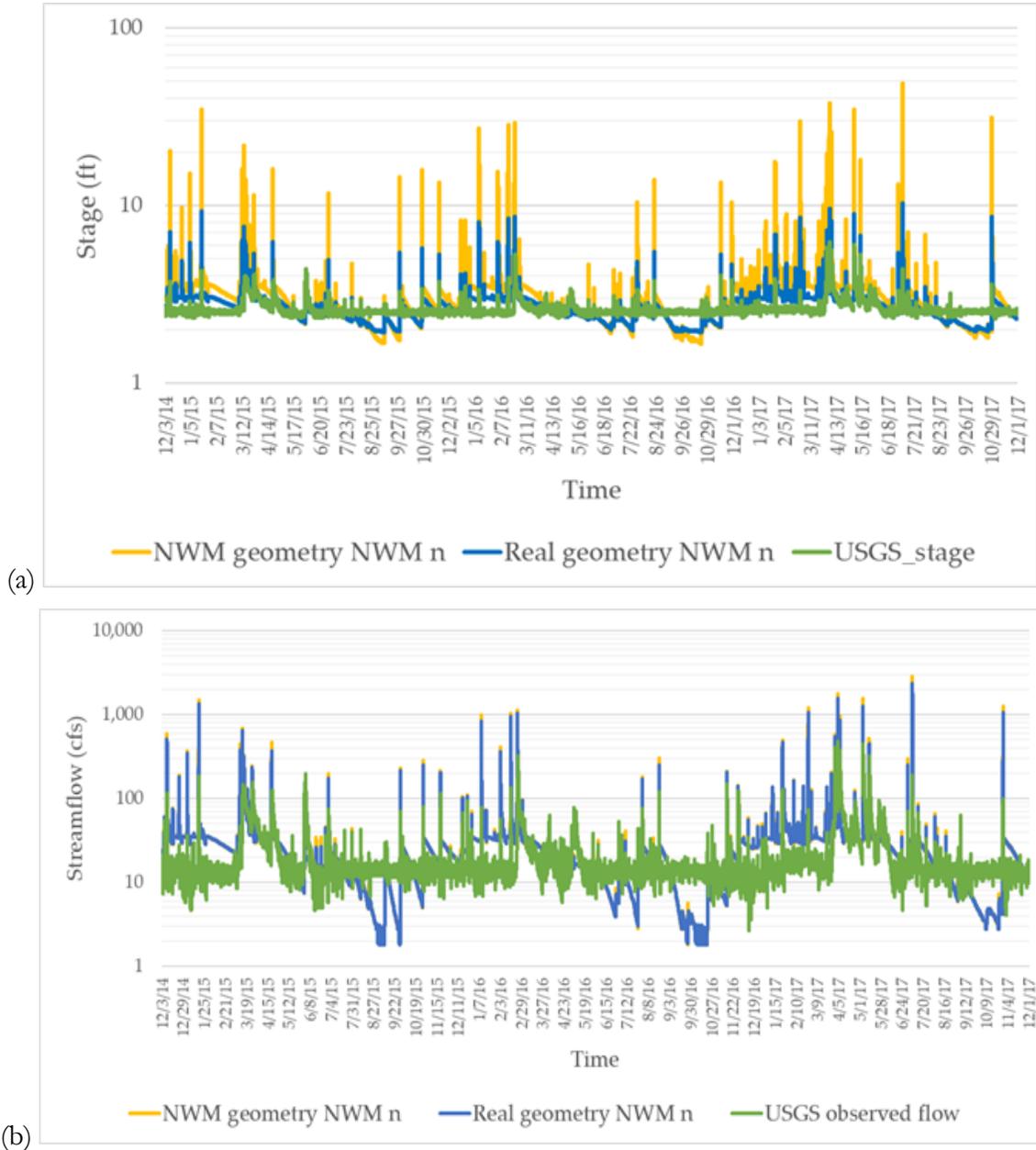


Figure 3. (a) Stage and (b) flow hydrographs for Hackensack River, Rockland County, NY

5.1.1. Stage Improvements

Table 2 shows the NSE, stage PBIAS and stage RMSE for the four combinations of hydraulic geometry and Manning’s n values that were simulated for each site, including additional geometry cases simulated for the Big Canoe Creek site in St. Clair County, Alabama. When the hydrograph forcing from the NWM is compared with the USGS observed data, the majority of the sites have a negative NSE for flow, which indicates that the NWM flow predictions for those river reaches are not very accurate. Hence, it can be inferred that since the quality of the input hydrograph forcing of the HEC-RAS were poor, it could have affected the HEC-RAS output hydrographs.

Table 2. Summary of Time-Series Results

			NWM vs. USGS	HEC-RAS (NWM forced) vs. USGS Stage Comparisons (sampled at USGS times)											
Site	Type (Full Period, Highest Flow)	Time Period Simulated	Hydrograph forcing quality	NWM geometry, NWM <i>n</i>		NWM geometry, "real" <i>n</i>		Real geometry, NWM <i>n</i>		Real geometry, "real" <i>n</i>		5-point generalization (with thalweg), NWM <i>n</i>		5-point generalization (evenly spaced), NWM <i>n</i>	
				PBI AS (%)	RM SE (ft)	PBI AS (%)	RM SE (ft)	PBI AS (%)	RM SE (ft)	PBI AS (%)	RM SE (ft)	PBI AS (%)	RM SE (ft)	PBI AS (%)	RM SE (ft)
Big Canoe Creek	Full	12/03/2014-12/31/2017	0.30	6.1	3.4	-0.06	2.8	-9.4	1.4	-11.2	1.4	0.5	1.5	27.7	1.6
	High	12/21/2015-01/09/2016	-0.48	73.7	15.0	53.7	11.2	3.4	1.1	2.4	1.2	10.4	3.2	10.2	3.2
Middle Fabius River	Full	12/03/2014-12/31/2017	-0.03	92.7	6.5	79.3	5.6	41.0	2.4	37.9	2.2	-	-	-	-
	High	04/09/2015-08/18/2015	0.46	105.1	14.8	84.9	12.4	20.6	3.2	17.4	3.1	-	-	-	-
Hackensack River	Full	12/03/2014-12/23/2017	-2.48	20.0	1.7	100.0	7.0	3.6	0.5	-3.5	0.4	-	-	-	-
	High	02/23/2016-03/05/2016	-3.84	2.6	4.4	14.0	14.8	0.3	1.1	-0.1	1.0	-	-	-	-
Cartoogechaye Creek	Full	12/03/2014-12/23/2017	-0.79	58.5	1.5	38.0	1.2	31.0	0.6	9.0	0.5	-	-	-	-
	High	12/24/2015-12/28/2015	-5.11	0.32	7.4	0.23	2.8	-0.03	2.6	-0.1	2.7	-	-	-	-
Little Tennessee River	Full	12/03/2014-12/23/2017	0.38	109.0	2.9	103.8	2.9	98.5	2.1	95.3	2.0	-	-	-	-
	High	12/23/2015-12/28/2015	0.17	97.8	10.0	0.24	10.3	-0.05	2.1	-0.04	1.8	-	-	-	-
North Second Creek	Full	12/3/2014-12/31/2017	-0.08	2.7	2.0	31.1	2.7	-50.2	1.5	-57.7	1.6	-	-	-	-
	High	09/20/2015-03/06/2016	0.46	9.4	3.3	40.8	4.6	-51.1	2.0	-54.7	2.1	-	-	-	-

The RMSE for stage shows improvements in all the sites for both the full time period and highest flow events with real channel geometry. Therefore, it is evident that incorporating real channel geometry has the potential to generate better predictions of stage in the NWM as well as other hydrologic models.

Table 3 shows the comparison of peak magnitudes of stage between the two different geometry configurations (both with NWM n) and the comparison of each configuration with USGS observed stage. The NWM geometry largely overpredicts the peak stage in all cases, whereas the real geometry fairly accurately predicts peak stage in three of the sites and slightly overpredicts in three sites. This further depicts that stage can be significantly improved by incorporating real channel geometry into the model.

Table 3. *Stage Statistics (Absolute Peak Difference and Percent Peak Difference)*

Site	Difference between peak magnitudes of RAS stage and USGS observed stage			
	NWM geometry, NWM n		Real geometry, NWM n	
	Difference (ft)	% Difference	Difference (ft)	% Difference
Big Canoe Creek	112.3	678	3.6	22
Middle Fabius River	122.2	463	-0.6*	-2
Hackensack River	23.8	451	3.3	63
Cartoogechaye Creek	26.3	218	4.5	37
Little Tennessee River	33.8	277	0.0	0
North Second Creek	54.1	369	-0.1*	0

*Negative differences mean RAS simulation for that particular geometry case has underpredicted the stage

5.1.2. Flow Routing Improvements with Better Estimates of On-The-Fly Muskingum Routing Parameters

The Muskingum K parameters were calculated for each output time from the HEC-RAS model for Big Canoe Creek, and the results from the peak event are shown in **Figure 5**. From **Figure 5**, it is apparent that during the peak event, the NWM currently underestimates the reach travel time by a factor of approximately 1.75. For the full 3-year time series, the current NWM assumptions consistently underestimate the travel time through this reach by a factor of approximately 1.82. These are in line with unpublished speculation about the NWM currently rushing flow through reaches about twice as fast as expected, and this revelation implies that channel geometry is indeed worth pursuing due to the potential dramatic effects on flow routing if this trend holds across other study sites. The

brief spike in K values during the peak event appears to be an artifact of the smoothing process but will be investigated further in future work.

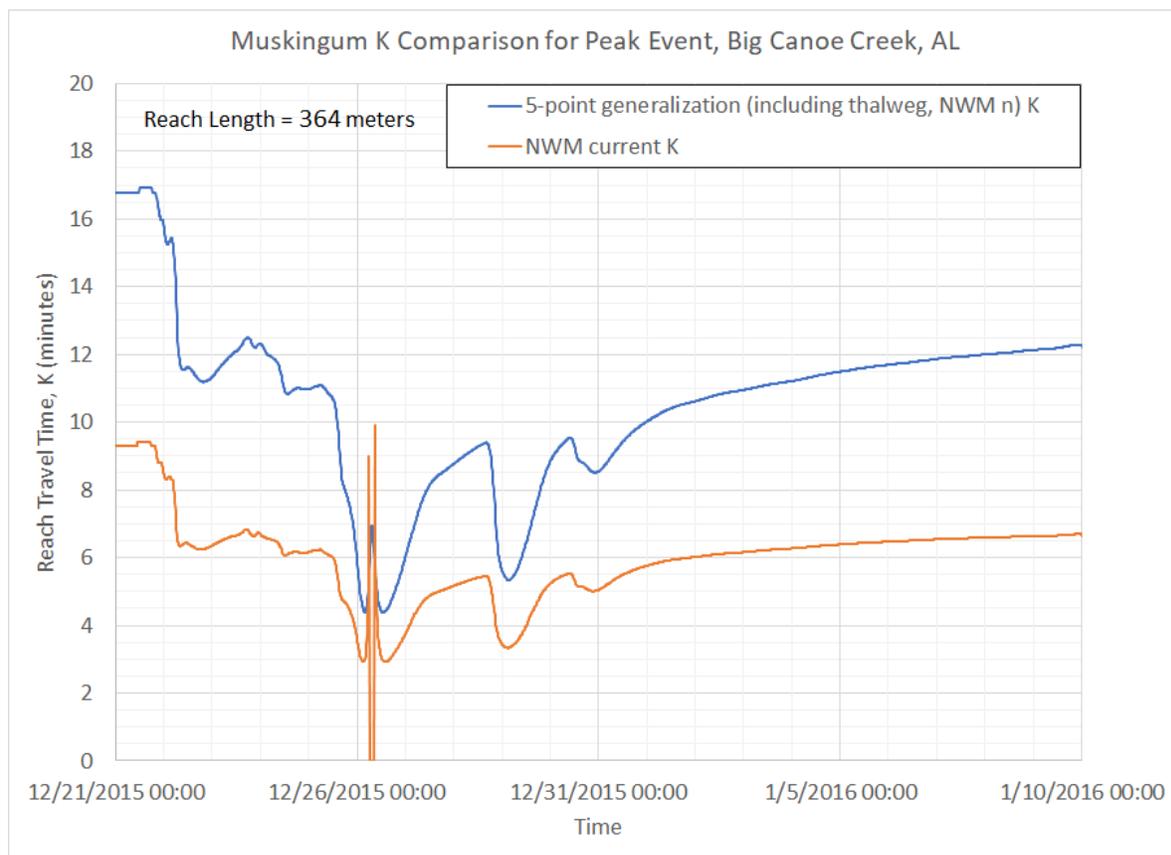


Figure 5. Comparison of Muskingum routing parameters estimated on-the-fly

5.2. Potential Improvements for Flood Inundation Mapping

Figure 6 shows the inundation map for Hackensack River for both real and NWM geometry cases with NWM Manning’s n , further showing that NWM geometry largely overpredicts the stage when compared with real geometry. If the NWM predicted stage was used in inundation extent predictions, it can be seen that Palisades Center Mall in Rockland County, NY, would expect flooding in this event if NWM geometry is used, but with the use of HEC-RAS geometry, the Mall is not falsely warned.

Another example of difference in inundation map using real geometry and NWM geometry is shown in **Figure 7** for Big Canoe Creek, which is a fifth order stream. It shows that using NWM geometry HEC-RAS predicts larger area to be inundated than using real geometry. This opens up an important possibility: using more realistic channel geometry in the NWM may have the potential for generating a reliable stage product that can directly be used to generate flood inundation predictions rather than relying on stage generated using synthetic rating curves which introduces many uncertainties. However, further investigation is required.



Figure 6. Inundation Map of Hackensack River

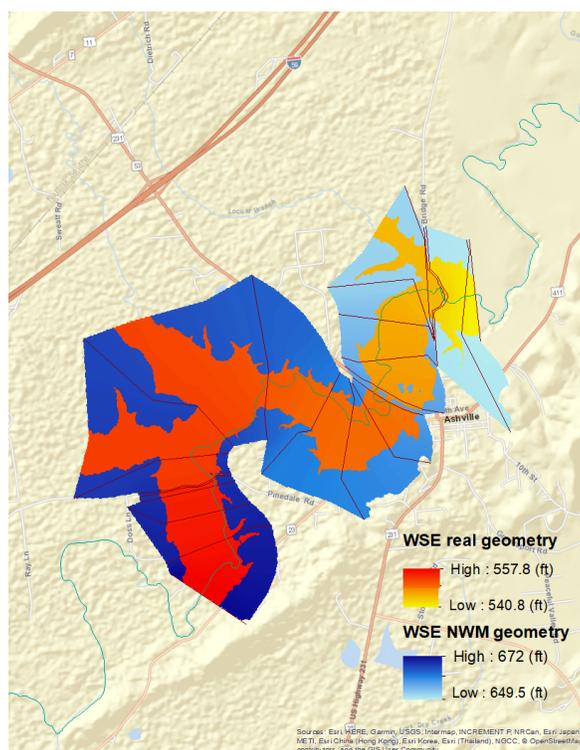


Figure 7. Inundation Map of Big Canoe Creek

5.3. Action Plan for Gathering New Data and Leveraging Existing Data

If citizen scientists in the locality are involved in collecting real channel cross section data, they can take measurements of channel width and depth of the river using the relatively simple method suggested above. This will be a straightforward and easy task, that does not require much equipment or intensive training, as only a measuring tape to measure the channel widths and a ruler to measure depths will be sufficient. To supplement citizen science efforts, initiating county-level programs of professional field surveying in stream channels can provide a means for local government agencies to obtain more accurate flow and inundation forecasts for their own county. Another companion dataset of geometry data is existing channel cross section data in HEC-RAS models for locations where FEMA flood insurance studies have been performed, provided that this data can be parsed into an appropriate form. One of the products of this research is a Python script which is capable of parsing native HEC-RAS geometry files, and this code could possibly be adapted to this end.

6. Conclusion

This project was carried out to assess the importance of representing real channel geometry in continental scale hydrological modeling with the idea of improving the NWM predictions and proposing a simple method to engage citizen science to fulfill the data needs. This work demonstrates that representing actual channel geometry in large scale hydrological modeling can significantly improve model predictions over trapezoidal geometry assumptions, both in terms of stage prediction and Muskingum-Cunge routing parameter estimation. Future work can be directed toward studying the exact details of collecting citizen science channel geometry data and finding out the advantages and disadvantages of the methods suggested herein.

The Summer Institute was a great opportunity for working in groups, networking, implementing existing skills and learning new ones. In our study, our work with HEC-RAS extended our exposure in hydraulic modeling. We have also gained good experience in working with geospatial data. Learning how to access and analyze real datasets from USGS has helped us to understand how to collect and work with field data. Overall, the experience was rewarding and very helpful for our futures in this field.

Supplementary Materials: A Hydroshare repository of models and codes developed for this project will be created shortly.

Disclaimer: This product uses the Federal Emergency Management Agency's API, but it is not endorsed by FEMA.

References

1. Wing, O. E. J., Bates, P. D., Sampson, C. C., Smith, A. M., Johnson, K. A., and Erickson, T. A. (2017), Validation of a 30 m resolution flood hazard model of the conterminous United States, *Water Resour. Res.*, 53, 7968–7986, doi:10.1002/ 2017WR020917.
2. Savage, J. T. S., Bates, P., Freer, J., Neal, J., & Aronica, G. (2016). When does spatial resolution become spurious in probabilistic flood inundation predictions? *Hydrological Processes*, 30(13), 2014–2032. <https://doi.org/10.1002/hyp.10749>
3. Neal, J.C., Odoni, N.A., Trigg, M.A., Freer, J.E., Garcia-Pintado, J., Mason, D.C., Wood, M. and Bates, P.D., 2015. Efficient incorporation of channel cross-section geometry

- uncertainty into regional and global scale flood inundation models. *Journal of Hydrology*, 529, pp.169-183. <https://doi.org/10.1016/j.jhydrol.2015.07.026>
4. Cohen, S., Kettner, A.J., and Syvitski, J.P.M. (2014). Global suspended sediment and water discharge dynamics between 1960 and 2010: Continental trends and intra-basin sensitivity. *Global and Planetary Change*, 115, 44-58. <http://dx.doi.org/10.1016/j.gloplacha.2014.01.011>
 5. Gochis, D., Dugger, A., & McCreight, J. (2016). Technical Description of the National Water Model Implementation *WRF-Hydro*, 53.1
 6. Blackburn-Lynch, W., Agouridis, C.T. and Barton, C.D., 2017. Development of Regional Curves for Hydrologic Landscape Regions (HLR) in the Contiguous United States. *JAWRA Journal of the American Water Resources Association*, 53(4), pp.903-928.
 7. Lowry, C. S., & Fienen, M. N. (2013). CrowdHydrology: Crowdsourcing Hydrologic Data and Engaging Citizen Scientists. *Ground Water*, 51(1), pp. 151–156. <https://doi.org/10.1111/j.1745-6584.2012.00956.x>
 8. Leopold, L.B., and Maddock, T. (1953). The hydraulic geometry of stream channels and some physiographic implications. Professional Paper 252, Washington, D.C. 57p.
 9. Ferguson, R.I. (1986). Hydraulics and hydraulic geometry. *Progress in Physical Geography: Earth and Environment*, 10(1), pp. 1-31.
 10. National Aeronautics and Space Administration (NASA) (2016). Surface Water and Ocean Topography Mission (SWOT§) Project Science Requirements Document.
 11. Durand, M., Andreadis, K. M., Alsdorf, D. E., Lettenmaier, D. P., Moller, D., and Wilson, M. (2008). Estimation of bathymetric depth and slope from data assimilation of swath altimetry into a hydrodynamic model. *Geophys. Res. Lett.*, 35, L20401, doi: 10.1029/2008GL034150.
 12. Durand, M., Rodríguez, E., Alsdorf, D.E., and Trigg, M. (2010). Estimating River Depth From Remote Sensing Swath Interferometry Measurements of River Height, Slope, and Width. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 3(1), pp. 20-31. doi: 10.1109/JSTARS.2009.2033453
 13. Federal Emergency Management Agency (FEMA) (2018). Hydraulic Studies (multiple), accessed July 12, 2018. Available from: <https://hazards.fema.gov/wps/portal/frisel>
 14. Falcone, J. (2011). GAGES-II: Geospatial Attributes of Gages for Evaluating Streamflow. United States Geological Survey, Reston, VA. https://water.usgs.gov/lookup/getspatial?gagesII_Sept2011
 15. United States Census Bureau. (2017). Cartographic County Boundary Dataset. Available from: https://www.census.gov/geo/maps-data/data/cbf/cbf_counties.html
 16. HEC-RAS River Analysis System User's Manual, version 5.0. (2016). Hydr. Engr. Ctr., U.S. Army Corps of Engrs., Davis, California.
 17. R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>

Chapter 7

Applications of Citizen Science Data for National Water Model Streamflow Forecasts

Di Wu¹ and Elizabeth A. Del Rosario²

¹Southern Illinois University-Carbondale; di.wu@siu.edu

²Harte Research Institute, Texas A&M University-Corpus Christi; elizabeth.delrosario@tamucc.edu

Academic Advisors: Ruopu Li, Southern Illinois University-Carbondale; Richard McLaughlin, Texas A&M University-Corpus Christi; Paul Montagna, Texas A&M University-Corpus Christi

Summer Institute Theme Advisor: Christopher Lowry, The State University of New York at Buffalo, cslowry@buffalo.edu

Abstract: The National Water Model (NWM) is a hydrologic model that simulates observed and forecast streamflow for approximately 2.7 million streams, based on an observational network of nearly 8,000 United States Geological Survey (USGS) stream gauges. This observational data network could be increased by integrating crowdsourced distributed hydrologic measurements on ungauged streams. Citizen Science (scientific work undertaken by members of the general public) engages the larger water community (at national, regional, and local scales), but comes with uncertainty. In order to investigate this uncertainty, a decision tree method was applied to evaluate existing citizen science data of stream stage base on the CrowdHydrology network. Quality control (QC) flags were developed for data measurements to pass from Level 1 (raw dataset), to Level 2 (flagged dataset), to Level 3 (processed dataset). QC flags were tested with synthetically generated crowdsourced stream stage measurements and unaltered USGS gage height. This methodology was then applied to CrowdHydrology sites and compared to co-located pressure transducer measurements. Error estimates were calculated to determine uncertainty in the in the citizen science data at these sites. Using this methodology, the NWM can incorporate crowdsourced data as independent verification and validation points to increase accuracy in forecast predictions. In addition, this research advances the Office of Water Prediction's goal of supporting a water-resilient nation by involving the public in the collaborative research process; allowing for better informed water management decisions, promoting water resource awareness / education, and increasing public trust.

1. Motivation

Water resource is one of the basic ecosystem services for human activities, which need to be managed appropriately. Hydrological science forms the fundamental of water resources management decisions by assessing water-related risks and challenges like pollution, floods and drought [1]. The National Water Model (NWM), as a hydrologic model, is designed for water cycle simulations and streamflow forecasts over the entire continental United States (CONUS) [2]. Taking hydrologic physical processes into account, NWM provides high-resolution forecasts of soil moisture, surface runoff, snow water equivalent, and other parameters by leveraging the network of USGS stream gauges [3].

However, the scarcity of observational data in NWM always exists, both in spatial and temporal scales. For some streams especially those with small stream orders, the USGS gages are sparse (**Figure 1**), which can cause problems when considering the heterogeneous and complicated requirements for water management and governing processes [1]. The current lack of observations in lower order streams provides motivation for exploring alternative sources of data to support the NWM.

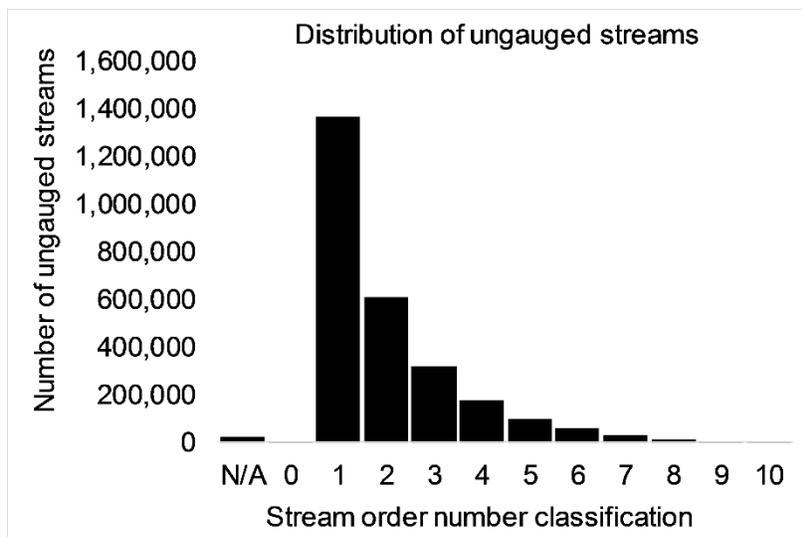


Figure 1. Number of ungauged streams by stream order classification in the United States.

Water runs in a network of streams spanning a length of at least 89 million kilometers worldwide [4]. The USGS has around 8,000 gauges in the United States that make up hydrologic monitoring networks. Even with this monitoring network, it is unrealistic to monitor all streams with in-stream sensors. Citizen Science “scientific work undertaken by members of the general public”, and crowdsourcing, “a form of citizen science, where data is provided by the crowd”, are practical methods to increase data collection. Increasing data observations increases the accuracy of stream forecast and expands our understanding of when, where, and how streams flow. Crowdsourcing can fill in information gap on intermittence streams and can vastly increase the number of monitored tributaries in a watershed. Crowdsourcing hydrologic data is also an easy means to promote public engagement and education about streams and watershed processes. This increases in public involvement in meaningful scientific work leads to collaborations among the public, scientist, and governmental which increases public trust [5].

2. Objectives and Scope

The goal of this study is to develop a methodology to apply a quality control process to citizen science data in order to expand the observational network for the NWM into ungauged watersheds. Many crowdsourcing platforms are available, however, crowdsourcing as a means of collecting scientific data has yet to become widely accepted [6]. Uncertainty and error in citizen science measurements are a primary concern for the scientific community [6]. Fienen and Lowry [7] found that high quality observations could be obtained without requiring trained observers, and stated that with a simple filter, errors such as transcriptions errors could be removed from the dataset. This research explores a decision tree method that allows citizen science data to move through a quality control process consisting of L1 (raw dataset) to L2 (flagged dataset) to L3 (processed dataset).

Our study looked at CrowdHydrology (www.crowdhydrology.com) sites. The CrowdHydrology project consists of signs displayed next to water level gauge staff (a large ruler) that encourage people

passing by to send a text message recording the station number and stream level from the gauge staff [8]. Our study looked at five sites on the Boyne River, a 3rd order stream in Northern Michigan, with corresponding pressure transducer data. These sites were compared with the nearest USGS gage for stage data and nearest CoCoRaHS (www.cocorahs.org) site for precipitation data (**Figure 2**).

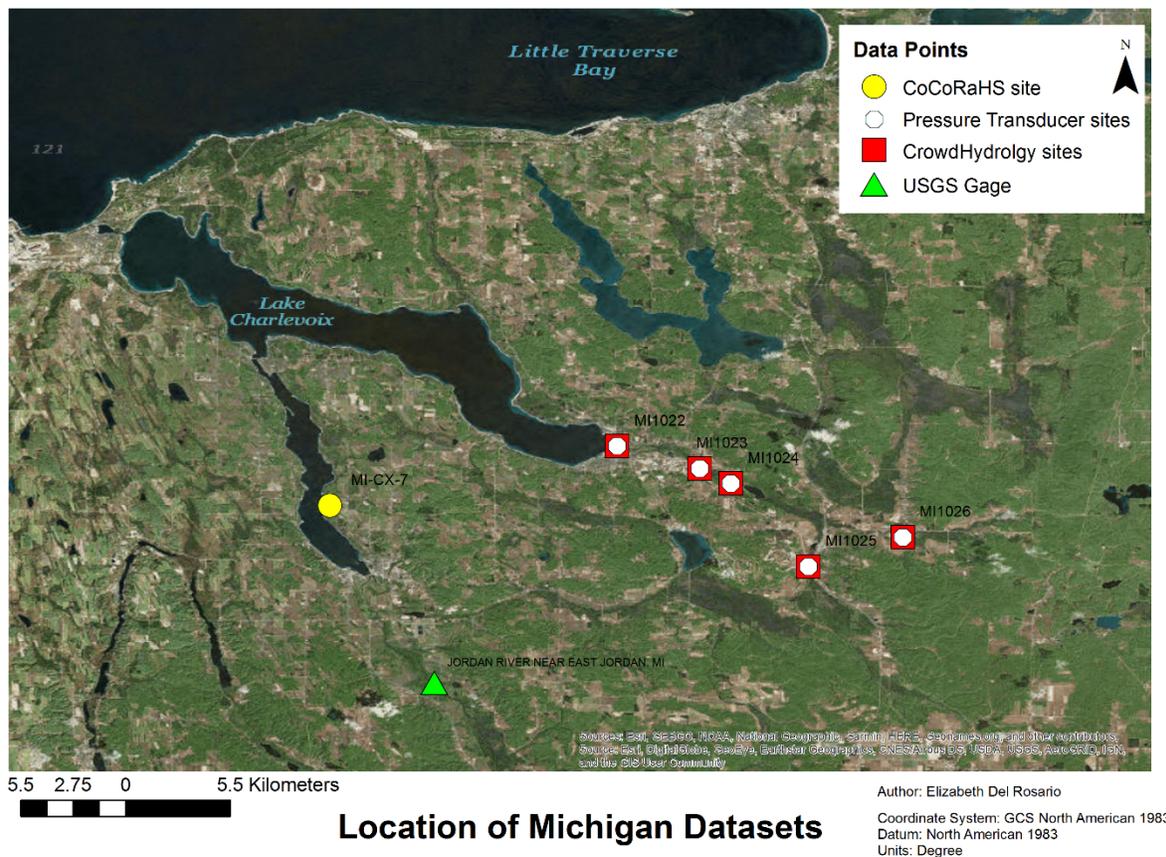


Figure 2. Location of Michigan sites on Boyne River.

The current NWM uses a network of approximately 8,000 USGS gages to forecast streamflow [3]. The USGS gaged sites are integrated into the model as verification / validation points in which the model is adjusted. The USGS gages monitor surface water in the United States. The distribution of gauges is biased towards the middle order streams (**Figure 3**), and for some states is disproportional to the surface water usage (**Figure 4**).

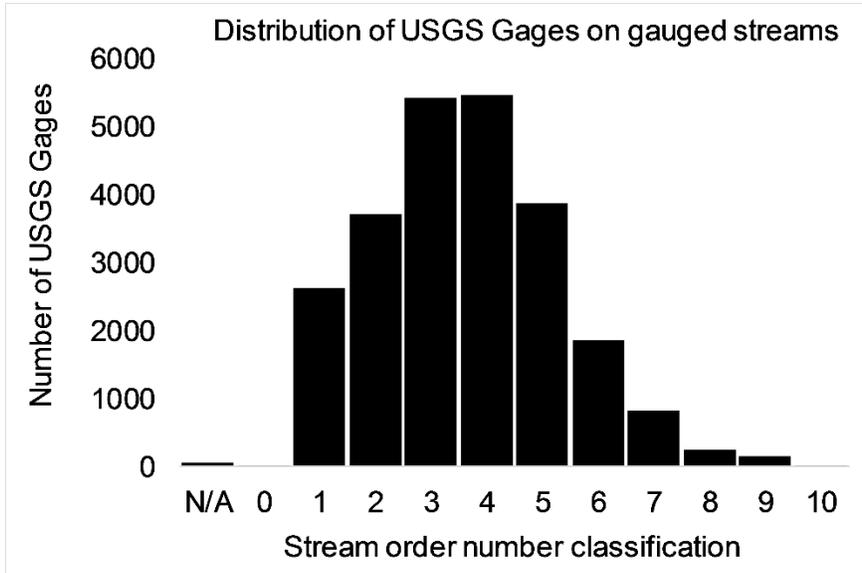


Figure 3. Distribution of USGS gages by stream order classification.

Crowdsourced hydrologic data are being collected (Appendix 1) and could be used to fill in these data gaps on the smaller order streams. These additional data could allow for a more even distribution among the stream orders and help bridge the gap between water monitoring and water withdrawals. In addition, the NWM could use the crowdsourced data points as independent verification / validation points like with the USGS gage data to increase the accuracy of the streamflow forecast. The scope of this project consisted of using various methods to explore citizen science data, develop a method for quality control in the citizen science data, and determine how uncertainty can be quantified in the quality control datasets.

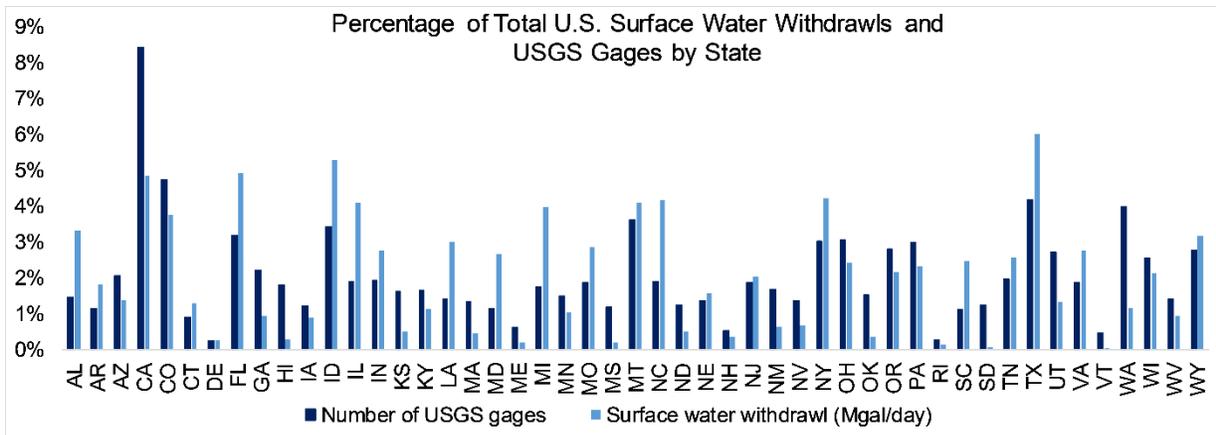


Figure 4. Percentage of total surface water withdrawals per day and percentage of total number of USGS gages by state (Alaska excluded). Data sources [9-11].

3. Previous Studies

As an enhancement to traditional research, citizen science projects based on crowdsourcing have the potential to meet some of the challenges of limited data availability. Defined by National Oceanic and

Atmospheric Administration (NOAA), U.S. department of commerce, citizen science is “a form of open collaboration where members of the public participate in the scientific process to address real-world problems in ways that include identifying research questions, collecting and analyzing data, interpreting results, making new discoveries, developing technologies and applications, and solving complex problems” [12,13]. Citizen science projects engage the community in scientific researches, which are designed for producing reliable data and information, creating new techniques, supporting natural resources management and raising public awareness of environmental concerns [14-16]. Citizen science is the base on which these large-scale or long-lasting environmental projects are conducted. Without work from these amateur volunteers, field data collection over a wide geographical region or a long period of time would be unachievable.

By participating in these citizen science projects, citizen scientists have made key contributions in various fields including climate change, invasive species, biological diversity, ecological restoration, and water resource management [17,18]. The dataset of National Weather Service’s Cooperative Observer Program (NWS-COOP), generated by volunteer weather observers, collects basic weather data across whole US to document the climate changes since 1890 [19]. North American Breeding Bird Survey (BBS) is another project which reflects the contributions of citizen scientists. It was a roadside-based survey developed by bird counts from volunteer observers starting in 1966. As the primary source to describe population change and distribution patterns for North American songbirds, BBS reflects bird group dynamics and conservation situations among different species [19,20]. The revolution of mobile phones and Internet in recent years provides citizen volunteers with easier and higher efficient approaches to collect, store and communicate a large amount of data [21]. These developments enlarge the capabilities of citizen science, driving its usage growth in new fields with innovative methods [22-24].

These technological advances have also encouraged the worldwide increasing engagements of citizen science data in hydrological researches with a wide range including stream flow estimation, floods prediction, hydrological database generation, and water quality monitoring [8,25,26] (Appendix 1). The confidence of data collected by citizen scientists has been evaluated in different scales [27-29]. Le Boursicaud et al. [30] gauged the flow velocities and hydraulic processes of extreme floods based on the YouTube home movie, which can be used for post-event determination of stream discharges. Using the same data sources, Michelson et al. [31] monitored water level changes in a Saudi Arabia cave. Turner et al. [32] tracked the annual reaches variation of perennial surface flow by wet/dry mapping. In this process, volunteers were trained to collect the coordinates of flow end points. There are also various citizen science projects related to hydrology launched by governments such as Volunteer Water Monitoring Programs (USA), the Risk-Scape Project (New Zealand), Water-Watch Victoria monitoring network (Australia), which not only vastly increase the number and type of available hydrologic data with low-cost, but also promote public understanding and participation about hydrological processes [33].

4. Methodology

To evaluate uncertainty in citizen science-based stream stage data in time series and flag erroneous observations, a decision tree was developed and applied in an EXCEL platform. Multisource datasets were utilized in these processes to determine the ruleset of decision tree. The classification ability of the decision tree was evaluated based on a set of CrowdHydrology stage data of Boyne River, Michigan.

4.1. Data Sources

To flag the citizen science data, which has high risk of error, the reference datasets are required. As

the data quality can be influenced by lots of factors like sampling scale, frequency, location, etc. multiple sources of reference data were utilized to reduce the consequences caused by different data collection methods in this study. The data sources and locations used are listed in **Table 1**.

Table 1. Data source and location for each station.

Data source	Station ID	Latitude	Longitude
CrowdHydrology	MI1022	45.214508	-85.011725
CrowdHydrology	MI1023	45.203904	-84.972731
CrowdHydrology	MI1024	45.196873	-84.958077
CrowdHydrology	MI1025	45.157571	-84.921393
CrowdHydrology	MI1026	45.1714449	-84.876804
CrowdHydrology	OR1000	44.576106	-123.32638
CocoRaHS	MI-CX-7	45.18639	-85.1475
USGS TX	08211503	27.8969652	-97.625551
USGS OR	14171000	44.525	-123.334
USGS MI	04127800	45.10250676	-85.098112

4.1.1. Crowd-Hydrology Dataset

The citizen science CrowdHydrology datasets were obtained for stations MI 1022-1026 and OR 1000 from www.crowdhydrology.com/data. These datasets span from May 2014 to June 2018 and were used for testing and obtaining the decision tree metrics. The number of measurements used for each station is listed in **Table 2**. Measurement duration is sporadic with less observations during the winter months (**Figure 5**).

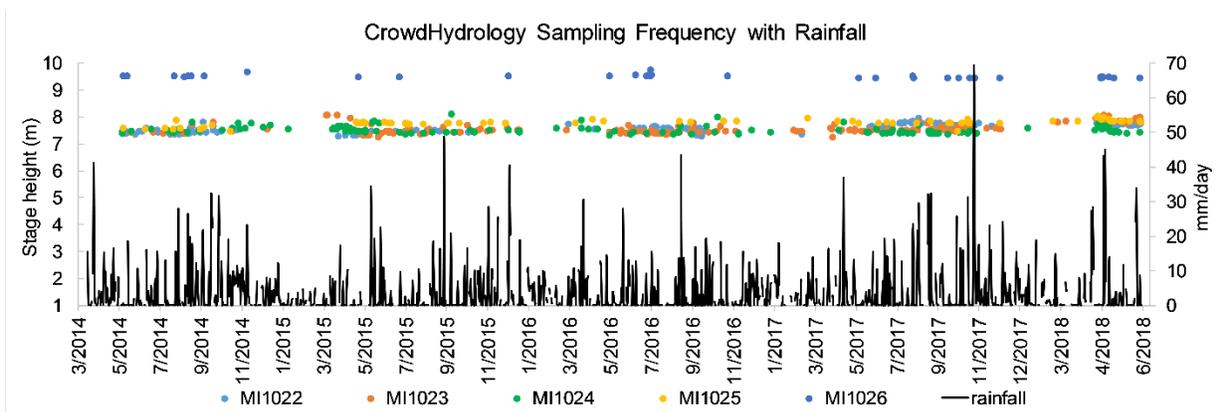


Figure 5. Michigan Boyne River CrowdHydrology measurements with CoCoRaHS rainfall.

Table 2. Number of measurements for each CrowdHydrology site used.

Station ID	Number of total measurements	Number of measurements corresponding to pressure transducers
MI1022	169	15
MI1023	211	20
MI1024	150	7
MI1025	93	23
MI1026	40	15
OR1000	1262	N/A

4.1.2. USGS Dataset

The USGS datasets were obtained from <https://waterdata.usgs.gov/nwis>. USGS is one of the most widely used reference data source, provides long-term stage monitoring data across whole country. TX USGS Gage No. 08211503 was used with added random noise, as simulation of citizen science data. It was compared to the original dataset, as corresponding truth values, and used for control selection. MI USGS Gage No. 04127800 was used in decision tree with the MI citizen science datasets for flagging the unusual data. The number of measurements used and date range for each gauge is listed in **Table 3**. Measurement duration is every 15 minutes.

Table 3. Number of measurements and date range for each USGS gage station used.

USGS Gage No.	Number of measurements	Discharge date range	Gage height date range
08211503	139,882	1/01/2014 to 12/31/2017	1/01/2014 to 12/31/2017
14171000	148,386	4/01/2014 to 7/02/2018	3/04/2018 to 7/02/2018
04127800	138,421	4/11/2014 to 7/02/2018	3/04/2018 to 7/02/2018

Stream discharge measurements made over the range in stage of the stream are plotted against the corresponding stages to define the stage-discharge relation that is used in conjunction with the recorded stage record to determine the discharges throughout the year [34]. The stage was back-calculated and values predicted for the MI and OR USGS gages by creating a rating curve from the 2018 gauge height and discharge data [34,35]. The rating curve was made by inserting a linear regression line with the equation in the form $Y = a + bX$, where X is the explanatory variable (discharge) and Y is the dependent variable (gauge height). Root mean square error (RMSE) was used to indicate the absolute fit of the model to the data—how close the observed data points are to the

model's predicted values [36]. RMSE values were calculated, the accuracy of the prediction considered acceptable, and the equations in **Table 4** used to predict stage height.

Table 4. Back-calculations for stage height for USGS gages.

Gage #	Flow Rating Curve Equations	R ²	RMSE
04127800	$y = 0.0059x + 1.9865$	0.9622	0.0578
14171000	$y = 0.0078x + 26006$	0.9977	0.4731

4.1.3. Pressure Transducer Dataset

Pressure transducers (PT) were installed and corresponded with the CrowdHydrology (CH) MI data points on the Boyne River. These data obtained from the pressure transducers was used for qualifying uncertainty in the citizen science data and evaluating the classification capabilities of decision tree by combined with CH data. The number of measurements used and date range for each PT are listed in **Table 5**.

Table 5. Number of measurements and date range for pressure transducer on the Boyne River.

Station ID	Number of measurements	Date range	Measurement duration
PT1022	40,000	6/07/2017 to 7/05/2018	1 min
PT1023	4,622	5/17/2018 to 7/04/2018	15 mins
PT1024	4,630	5/17/2018 to 7/04/2018	15 mins
PT1025	20,794	7/27/2017 to 7/12/2018	15 mins / 5 mins
PT1026	25,546	9/13/2017 to 7/11/2018	15 mins

4.1.4. Additional Datasets

Except the biases from data collection, recording and processing, the singular data can also be described by some external factors. Additional datasets were used for setting controls to determine if they can be used as potential reference to flag wired data.

1) Precipitation data

Precipitation is one of possible factors which may impact the stream stage. To make up the gaps in dataset, two sources of precipitation data were selected. The first is from National Climatic Data Center (NCDC), NOAA. The second is CoCoRaHS dataset, a national citizen science-based precipitation observation network. These datasets report cumulative daily precipitation in inches from April 1, 2014 to June 29, 2018. For NCDC dataset, the record was chosen from the gauge which is

nearest to Citizen Science data gauge to solve the disagreement between two data collection system. For CoCoRaHS data, the precipitation is recorded as county scale (**Figure 2**).

2) Temperature data

Temperature may have potential influence not only on stage itself but also volunteers' activities, which are close related with sampling frequency of citizen science data. The daily average temperature data was collected from NCDC-NOAA. As the precipitation data, the observation station with closest distance from citizen science gauge was select to obtain temperature records. The unit of temperature dataset is Fahrenheit.

4.2. Decision Tree

Decision tree methods, also called recursive partitioning, were developed to segment the target dataset into subdivisions based on the predefined controls for each branch [37,38]. As one of the most commonly used prediction models, decision tree has incomparable advantages for binary classification by facilitating user's comprehension and simplifying the classification processes [39]. The typical framework of a simple decision tree includes one input dataset, several test rules and a set of categories, which correspond to the root, branches and leaf nodes. To ensure the on-way data flow and avoid loops in the decision tree, a node is only allowed to have one parent node. Based on this structure, the input dataset can be subdivided sequentially according to the controls and fall into a certain class in the end [38].

The controls in a decision tree define the classification rules. By breaking a complex decision into a set of sequentially independent controls, decision trees implement data categorization in a multistage approach [40]. Every obtained case should satisfy the ruleset which is composed of controls along the path from the root to the corresponding leaf. For the same task, different control combinations may lead to different conclusion and accuracy [41].

4.2.1. Control Design

For the stream stage, two types of data, which have high possibility to be sources of uncertainty in citizen science dataset, were flagged. One is the incorrect data, which should be kicked out before data analysis; the other type are data with unusual values, to which close attention should be paid. To set the classification rules for decision tree, a simulated citizen science dataset of TX was applied to assess the capability quantitatively for each control. Some records were selected randomly and deleted to simulate irregular sampling frequency. Among the remained 104,030 stage records, there were 528 points whose values were modified as random noise. Except these data with noise, original dataset contains 44 zero records, which also should be flagged. Seven controls are designed to be test in this study, including:

1) Positive

The stage value cannot be negative. The zero stage values, as an extreme situation, also should be given great emphasis. Thus, if a stage record is not positive, it was flagged.

2) Local stability

The sharp changes on local scale in time series data always imply the potential issues. To pick these data points, the data bias from average stage was measured by standard deviation. Moving windows with four sizes (3-day, 5-day, 7-day and 14-day), which reflect the variation in different time scales and decrease the rate of mis-classification, were used to calculate local average $Aver_i$ and standard deviation S_i ($i = 3,5,7,14$). If the distance between a data and $Aver_i$ is larger than triple S_i , a flag

should be assigned to it.

3) Sampling frequency

Sampling interval may produce the gap in time series data, which may reduce the reliability of dataset. In this study, we assume that the threshold is 3-day.

4) Comparison with reference dataset

Reference data source can be diverse and include reports on governmental website or field measurements. We selected USGS gage data as reference to test if the trend of citizen science data is consistent. Several steps were used to assess agreement:

1. Calculate absolute difference dif between citizen science data and reference USGS data;
2. Draw out the paired data from citizen science dataset whose slope k_{CS} has different tendency with corresponding k_{USGS} , and tab them ‘temporary flag’;
3. Calculate average \underline{dif} and standard deviation S for temporary flagged group, if $\left|dif - \underline{dif}\right| > S_{flag}$, flag it, the unflagged data was put back to the citizen science dataset;
4. Calculate average \underline{dif} and standard deviation S for the rest of citizen science dataset, if $\left|dif - \underline{dif}\right| > S$, flag it.

5) Sampling time

As a manually acquired data source, quality of citizen science data may be related with sight-affected factors like darkness. We assumed that during the early morning (0:00-3:59) and late night (21:00-23:59) the probability of recoding wrong value is higher than other time.

6) Precipitation

As a major source of fresh water, precipitation may complement the stream. We assumed that the changes of stage should correspond with precipitation. The rate of precipitation (**Table 6**) has been classified by American Meteorological Society (AMS) to reflect the rainfall intensity, which were chosen to determine the rank of precipitation and stage slopes [42]. We assume the stage data should be flagged if the rank of its slope does not agree with corresponding precipitation slope.

Table 6. Precipitation categories

Precipitation categories		Rate(mm/h)	Rank
Drizzle	A trace	<0.10	1
	Light	0.10-0.25	2
	Moderate	0.25-0.50	3
	Heavy	0.50-1.00	4
Rain	Light	1.00-2.50	5
	Moderate	2.50-7.60	6
	Heavy	>7.60	7

7) Temperature

We assume that the extreme temperature may be in relation to sampling frequency due to its impact on the human activities. Here the range of 15 to 25 Celsius was selected as a comfortable temperature for citizen scientists [43]. Data with corresponding temperature out of this range were flagged.

The flag assignment rules are shown in **Table 7**. For each control, the results involve two class ‘Flag’ and ‘Unflag’, which correspond to Flag 1 and Flag 0.

Table 7. The flag assignment Rules for each control

Control		Flag Assignment Rules			
1	Positive	$Stage > 0$	0		
		$Stage \leq 0$	1		
2	Local Stability	$ Citizen\ Science\ data - Aver_i \leq 3 * S_i \quad i = 3,5,7,14$	0		
		At least one of $ Citizen\ Science\ data - Aver_i > 3 * S_i \quad i = 3,5,7,14$	1		
3	Sampling Frequency	$\leq 3\ days$	0		
		$> 3\ days$	1		
4	Reference Comparison	$k_{CS} \times k_{USGS}$	$ dif - \underline{dif} $		
		< 0			
		Only one slope equals to 0	Temporary Flag	$> S_{flag}$	1
		$= 0$	Both slopes are equal to 0	$\leq S_{flag}$	Temporary Unflag
> 0	$\leq S$	0			
5	Sampling Time	21:00 – 3:59(+1 day)	0		
		4:00 – 20:59	1		
6	Precipitation	$Rank_{precipitation} = Rank_{Citizen\ science\ date}$	0		
		$Rank_{precipitation} \neq Rank_{Citizen\ science\ date}$	1		
7	Temperature	15°C – 25°C	0		
		$< 15^\circ C\ or\ > 25^\circ C$	1		

4.2.2 Ruleset Generation

Ruleset is composed of ranked controls with good capability to categorize data points into appropriate

classes. Every control was tested individually. To reflect the classification's results, we utilized a confusion matrix. Four performance indicators were considered to quantitatively estimate the classification accuracy of each control. Assume a binary confusion matrix as **Figure 6**, we computed Precision P , Recall r , F1-score F_1 , and *Accuracy* [44,45].

		Predicted class	
		P	N
Actual Class	P	True Positives (TP)	False Negatives (FN)
	N	False Positives (FP)	True Negatives (TN)

Figure 6. Binary confusion matrix

For this study, the precision describes the proportion of the data points flagged by controls that should be flagged according to the ruleset. The recall expresses the ability of controls to find flag-data in the dataset. The equations of precision and recall are:

$$P = \frac{TP}{TP + FP} \quad (1)$$

$$r = \frac{TP}{TP + FN} \quad (2)$$

Taking the harmonic average of precision and recall leads to the F1-score, which gives equal weight to both measures and avoid the punishment of extreme values. The F1-score can be calculated using:

$$F_1 = 2 \times \frac{P \times r}{P + r} \quad (3)$$

The *Accuracy* reflects the overall accuracy of classification and is computed by:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (4)$$

In this study, FN represents these data points which are misclassified into unflag class. To reduce the FN rate, controls selected to form the rules combination tend to have complementary abilities. Dividing FN into non-positive data and positive data with bias would be helpful to figure out the

different ability of controls and determine their operation orders. The performance of selected controls combination was also assessed by confusion matrix and performance indicators as individual controls.

4.2.3. Prediction Accuracy Estimation for Decision Tree

For accuracy evaluation, Crowd-Hydrology data from five stations on Boyne River, MI, were applied to obtain decision tree as input dataset. To test the flexibility of the decision tree, USGS data from the nearest gauge and pressure transducer data from the same five stations were both used as references. The accuracy of prediction was assessed qualitatively and quantitatively. Using this methodology, the question if the decision tree can improve citizen science data quality to meet the requirement of NWM can be answered.

Error and uncertainty were calculated among the raw (L1), flagged (L2), and processed (L3) datasets. The root-mean-square error (RMSE) is a measure of accuracy between the values from an estimator (citizen science data) and the values observed (pressure transducer data) or considered true.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}} \quad (5)$$

Percent error is used when comparing an experimental result E with a theoretical value T that is accepted as the "correct" value [46].

$$Percent\ error = \frac{|T - E|}{T} * 100\% \quad (6)$$

Fractional or relative uncertainty is used to quantitatively express the precision of a measurement [46].

$$Percent\ uncertainty = \frac{error}{E} * 100\% \quad (7)$$

Percent difference is used when comparing two experimental results E1 and E2 that were obtained using two different methods [46].

$$Percent\ difference = \frac{|E_1 - E_2|}{\frac{E_1 + E_2}{2}} * 100\% \quad (8)$$

5. Results

The results section presents the obtained decision tree and its reliability. Based on this decision tree, the improvements of citizen science dataset are evaluated quantitatively.

5.1. Control Selection

Confusion matrix allows visualization of the performance for each control. Based on the USGS TX data with added noise, the classification confusion matrix is shown in **Table 8**. Classification performance was estimated by four indicators and the results are shown in **Table 9**, which provides the quantitative comparison between different controls.

Table 8. Classification confusion matrix for simulated citizen science data

	Control	Actual Class	Prediction Class	
			Flag	Unflag
1	Positive	Flag	403	169
		Unflag	0	103458
2	Local Stability	Flag	96	476
		Unflag	725	103209
3	Sampling Frequency	Flag	0	572
		Unflag	1	103457
4	Reference Comparison	Flag	94	478
		Unflag	0	103458
5	Sampling Time	Flag	160	412
		Unflag	30192	73266
6	Precipitation	Flag	209	363
		Unflag	50543	52915
7	Temperature	Flag	334	238
		Unflag	60603	42855

Table 9. Classification performance estimation

	Control	Precision	Recall	F1-score	Accuracy
1	Positive	1	0.7045	0.8267	0.9984
2	Local stability	0.1169	0.1678	0.1378	0.9885
3	Sampling frequency	0	0	NA	0.9945
4	Reference Comparison	1	0.1643	0.2823	0.9954
5	Sampling time	0.0053	0.2797	0.0103	0.7058
6	Precipitation	0.0041	0.3654	0.0081	0.5107
7	Temperature	0.0055	0.5839	0.0109	0.4152

Control orders relate with the efficiency and result of decision tree. In this study, FN represents the data points which are misclassified into unflag class. To reduce the FN rate and improve the flagging accuracy, controls selected to form the rules combination tend to have complementary abilities. Dividing FN data points into non-positive data and positive data with bias. **Table 10** is used to reflect the mis-sort types in FN for each control.

Table 10. Flag-class data misclassified into unflag class

	Control	Misclassified flag-class data				
		Total	Type			Bias
			Negative or 0			
1	Positive	169	0	0%	169	100%
2	Local stability	476	353	74.16%	123	25.84%
3	Sampling frequency	572	403	70.45%	169	29.55%
4	Reference Comparison	478	373	78.03%	105	21.97%
5	Sampling time	412	308	74.76%	104	25.24%
6	precipitation	363	228	62.81%	135	37.19%
7	Temperature	238	144	60.50%	94	39.50%

To establish optimal ruleset, controls with good and complementary classification abilities were chosen. According to **Table 8** and **Table 9**, as the worst performance on identifying flag type data, Control 3 (Sampling frequency) should be removed. Taking F1-score and Accuracy into account simultaneously, Control 1 (Positive), Control 4 (Reference Comparison) and Control 2 (Local stability) have high potential for flag abnormal value. Based on the result of **Table 10**, among these three controls, control 1 has preeminent capability to locate non-positive values but doesn't work with data containing biases. On the contrary, control 4 and 2 have low misclassification rate for biased data and high error rate on non-positive type data. Thus, the combination including control 1, 2, 4 could be potential to provide substantial information without large duplication and lead to satisfactory results of categorizing data into flag and unflag classes. The structure decision tree is show in **Figure 7**. Applying this decision tree to the TX data with added noise, the results was recorded as **Table 11**, which presents a great performance.

Table 11. Classification result of Decision tree

Confusion Matrix			
Actual class	Prediction class		
	Flag	Unflag	
Flag	572	0	
Unflag	726	102732	
Classification Performance Estimation			
Precision	Recall	F1-score	Accuracy
0.4407	1	0.6118	0.9930

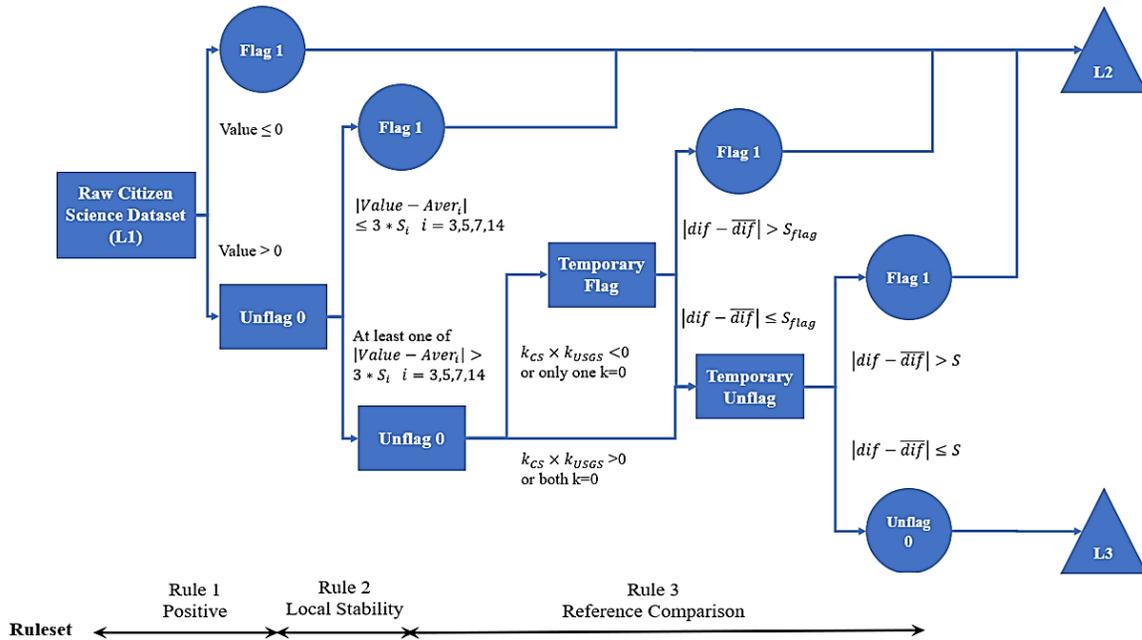


Figure 7. Decision tree structure

Although the general performance of the decision tree is satisfactory, there are misclassifications mainly concentrating on extreme values due to natural fluctuation of stream stage (Figure 8). Potential methods to reduce this type of error is to change thresholds (multiple of standard deviation in rule 2 and rule 3) which are used to measure the value bias. As the threshold is determined empirically, it may not be the appropriate one for the specific dataset. Exploring more thresholds can be helpful to find a superior metric.

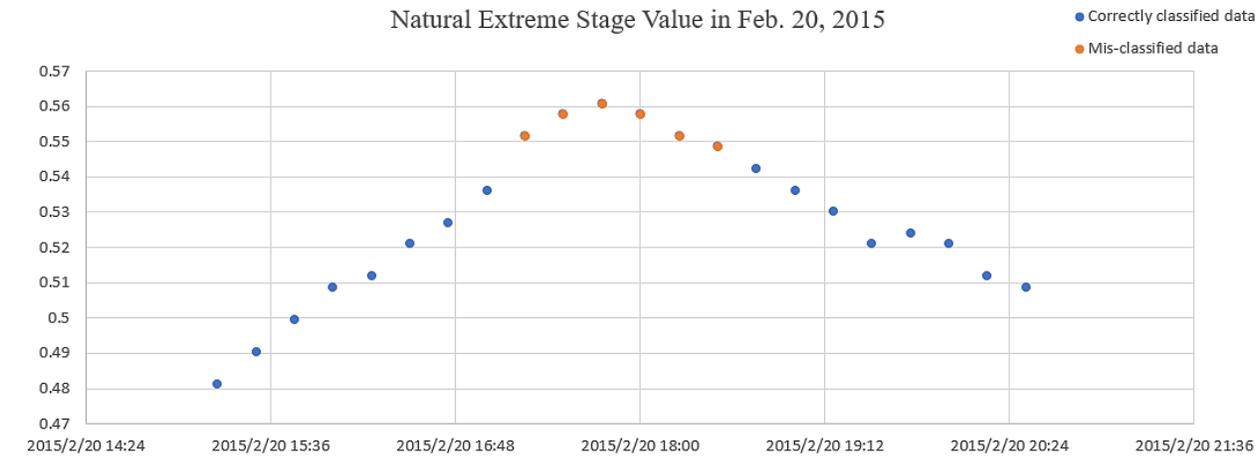


Figure 8. Example of natural extreme value in time series stage data

5.2. Application of Decision Tree Methodology

The simulated citizen science dataset consisted of 104,029 data points in the L1 (raw) dataset; 88,099

points flagged by the decision tree methodology in the L2 dataset; and 15,930 data points in the L3 (processed) dataset. The original USGS TX data consisted of 104,030 stage records. **Figure 9** shows the simulated dataset with added noise. The decision tree methodology was successfully at flagging potentially erroneous data points that were removed with processing in L3 dataset (**Figure 9**)

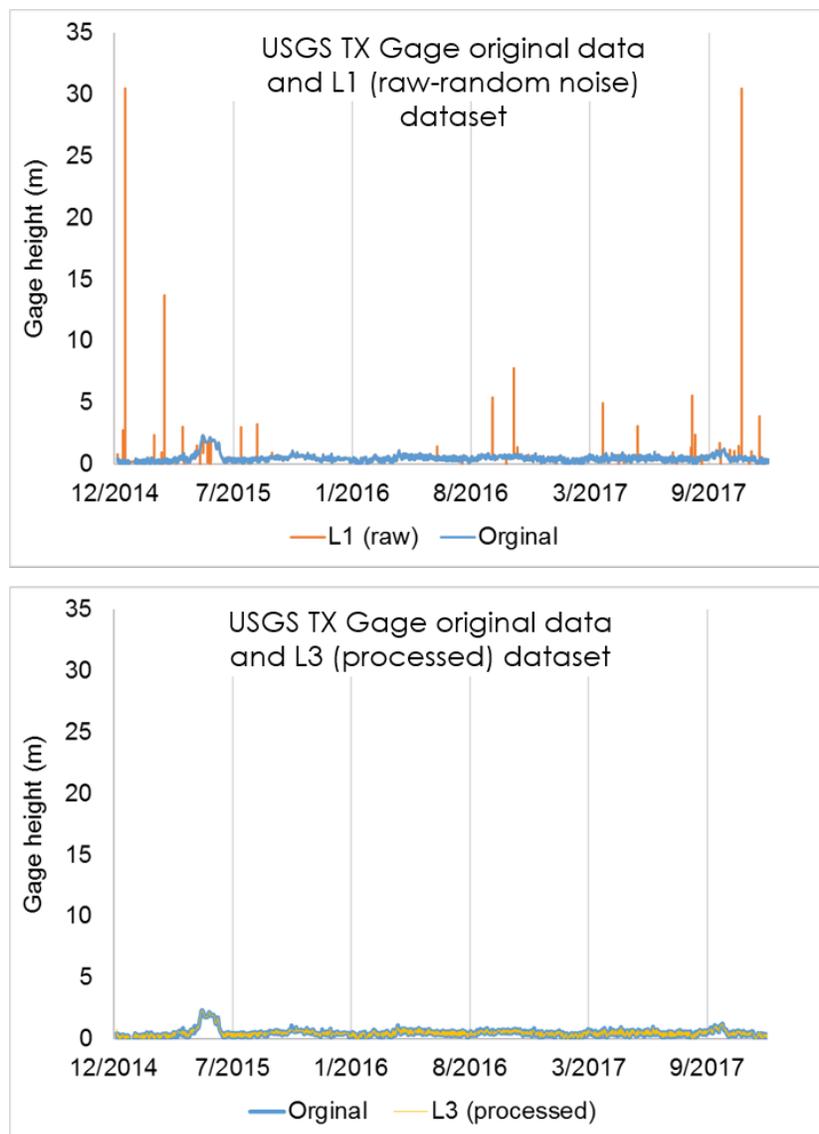


Figure 9. L1 (raw) simulated citizen science dataset with original USGS TX gage dataset (top). L3 (processed) simulated citizen science dataset with original gage dataset (bottom).

The number of data points in each dataset (L1-L3) are listed in **Table 12** and were used to create the hydrographs plots in **Figure 10**. The L1 (raw) dataset with the points flagged (from L2) by the decision tree are on the left side, and the L3 (processed) dataset with the flagged data points removed for each MI station are on the right side (Fig. 10). The number of data points flagged by the decision tree (a, c, e, g, i) created more noise in the hydrographs. Removing these data points (b, d, f, h, j) created gaps in the plots. To fix these gaps, other data sources are needed to correct flagged data points to adjust them to the pressure transducer's data. The decision tree was able to flag potential problem data

points, however, more research is needed examine the L2 dataset to correct points to move to the L3 dataset.

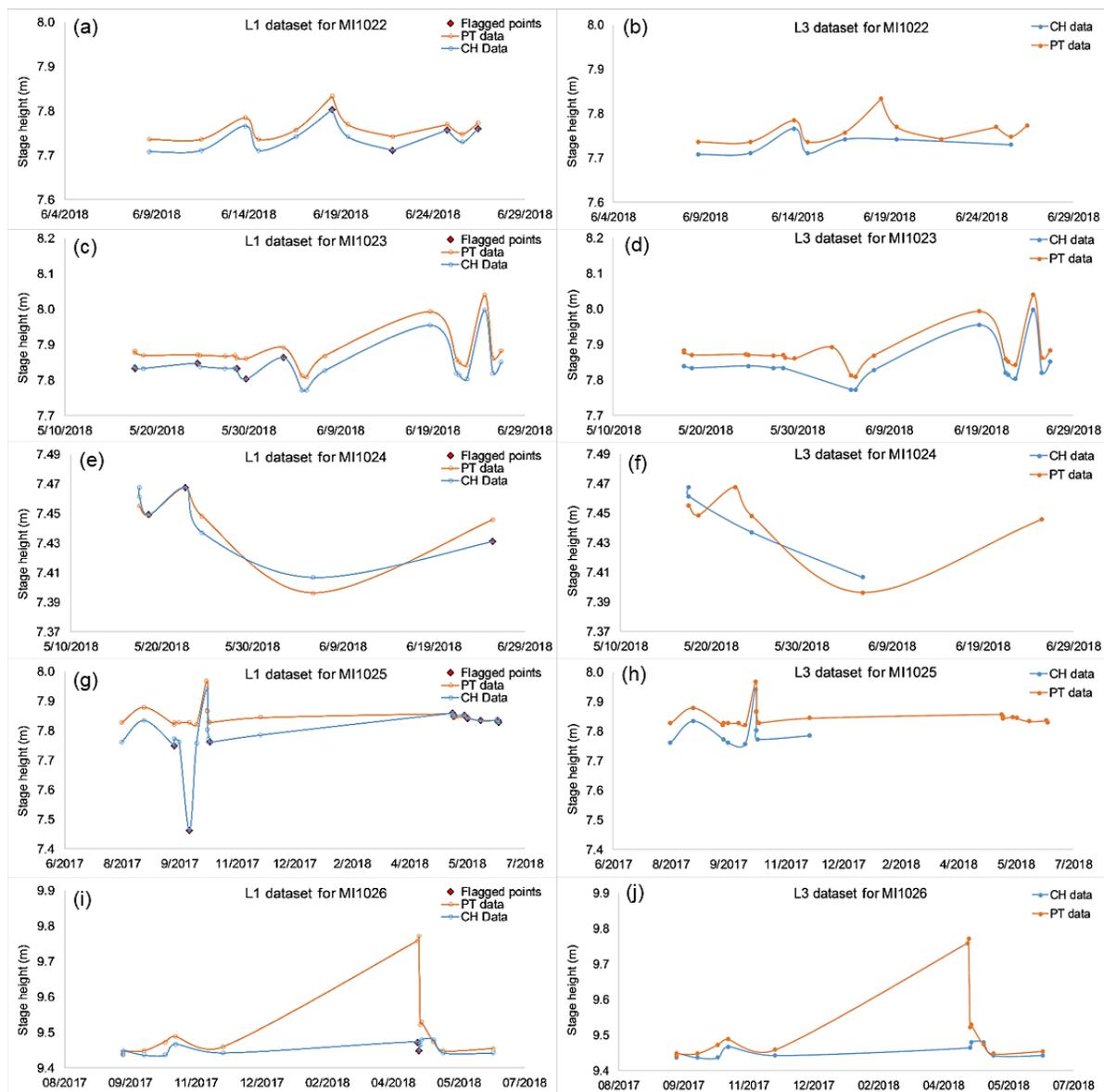


Figure 10. L1 raw dataset with flagged points (right) and L3 dataset with the L2 points removed (left). Orange line is pressure transducer dataset, blue line is CrowdHydrology dataset.

Table 42. Number of data points in each dataset.

Station ID	L1	L2	L3
MI1022	15	4	11
MI1023	20	5	15
MI1024	7	3	4
MI1025	23	14	11
MI1026	13	2	11

5.3. Prediction Accuracy of Decision Tree

Error estimates between the CrowdHydrology (CH) datasets and the pressure transducer (PT) datasets are shown in **Table 13(a,b)** were calculated from the number of data points in the datasets in **Table 12**. The smallest error estimates for the L1 and L3 datasets are in blue. RMSE was decreased by removing data points for all sites except MI1024. Removing data points increased all error estimates for MI1024. The increase in error estimates for MI1024 is most likely because there were only 7 data points that corresponded with the pressure transducer (**Table 12**). Percent error decreased except for MI1022 and MI1024. Percent uncertainty increased at all sites. Percent difference was decreased at sites MI1023 and MI1026.

The random error in experimental results is due to lack of observer precision, resulting in a spread of results (**Figure 11**). Due to the random nature of these errors, there is an equal chance that they will be above or below the ‘true’ value or measured mean value [47]. As in **Figure 11**, error is estimated in a range for the datasets, \pm instead of %, we divide the table % values by 100 and result in \pm meters (**Table 13b**). This makes the error, uncertainty, and difference values less than the RMSE value and puts them in level with the USGS staff gauge [48,8].

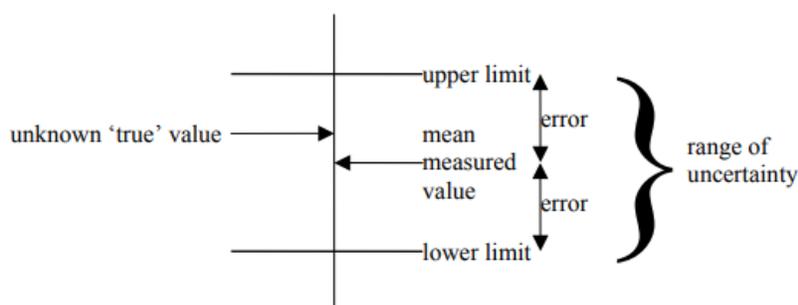


Figure 11. Error and uncertainty estimates [47].

5.4. Signal Analysis

The CrowdHydrology (CH) and pressure transducer (PT) had different time frames, therefore, only a subset of the CH datasets could be compared one-to-one with the PT datasets. (The CH date range was from 2014 to 2018, and the date range for the PT data was from 2017 to 2018). The decision tree methodology used the nearest USGS gage to the CH dataset as a flagging criteria. Since the CH and PT datasets are on an ungauged stream a USGS station from a neighboring watershed was used. The USGS gage showed similar trend in the signals for stream stage (**Fig. 12**), however, was on a different scale from that of the CH datasets and the PT datasets. A signal analysis was used to evaluate patterns between the USGS gage, the CH, and PT data (**Fig. 13**). SAS Wavelet Analysis produced the scalograms from the daily mean stage height for the CH, PT, and USGS datasets. Wavelet scalograms communicate the time frequency localization property of the discrete wavelet transform performed by SAS 9.4 [49]. The images produced are shown in **Figure 13**. The location and size of the rectangle are related to the time interval and the frequency range for this coefficient; low levels are plotted as wide and short indicating a wide time interval but a narrow range of frequencies in these data, high levels are plotted thin and tall indicating small time ranges but large

Table 13a. Error calculations for Michigan CrowdHydrology and pressure transducers datasets.

Site	Dataset	RMSE (m)	% Error	% Uncertainty	% Difference
MI1022	L1	0.0230	0.2820	0.0111	0.0706
	L2	0.0232	0.2741	0.0353	0.0686
	L3	0.0227	0.2865	0.0371	0.0717
MI1023	L1	0.0383	0.4776	0.0186	0.1197
	L2	0.0388	0.4678	0.0597	0.1173
	L3	0.0381	0.4698	0.0601	0.1177
MI1024	L1	0.0095	0.1063	0.0044	0.0266
	L2	0.0085	0.0696	0.0094	0.0174
	L3	0.0102	0.1339	0.0180	0.0335
MI1025	L1	0.0892	1.6125	0.0648	0.1642
	L2	0.1071	0.9728	0.1286	0.2473
	L3	0.0582	0.7095	0.0911	0.1781
MI1026	L1	0.1219	0.6565	0.0207	0.1661
	L2	0.3042	3.1098	0.3287	0.7898
	L3	0.0273	0.2105	0.0223	0.0527

Table 13b. Error estimates based on range \pm meters for the CH and PT datasets.

Site	Dataset	Error (m)	Uncertainty (m)	Difference (m)
MI1022	L1	0.00282	0.00011	0.00071
	L3	0.00287	0.00037	0.00072
MI1023	L1	0.00478	0.00019	0.00120
	L3	0.00470	0.00060	0.00118
MI1024	L1	0.00106	0.00004	0.00027
	L3	0.00134	0.00018	0.00033
MI1025	L1	0.01613	0.00065	0.00164
	L3	0.00710	0.00091	0.00178
MI1026	L1	0.00657	0.00021	0.00166
	L3	0.00210	0.00022	0.00053

frequency ranges in the data [49]. The energy is defined as the sum of the squares of the detail coefficients for each level [49]. The overall pattern in the analysis for the pressure transducers and the MI CH sites corresponds with the pattern for the MI USGS gage (l, o). The patterns of the CH sites compare with the patterns of the PT sites (e, h, k, n), except the MI1022 PT site (b) lacks a pattern. This is most likely do to the daily mean producing a small amount of data, since the sample duration was every minute. The patterns of the MI CH sites all appear similar to each other (a, d, g, j, m), and different from the OR CH site (c) where cycles appear as blue (decrease in energy). The pattern for the OR CH site (c) and the OR USGS gage (f) corresponds with each other in that both produced cycles for red/pink (increases in energy) and blue (decreases in energy). The OR USGS gage (f, i) appear different from that of the MI USGS gage (l, o) with no apparent cycles. The 2017-2018 USGS gages, OR (i), MI (l) match that of the pressure transducers (e, h, k, n). The MI USGS gage (l) compare well with the PT sites (k, n), while the OR USGS gage (i) lacks energies in the upper right and left corners. Using this type of signal analysis allows for pattern comparison in the stage heights as increases and decreases in energies.

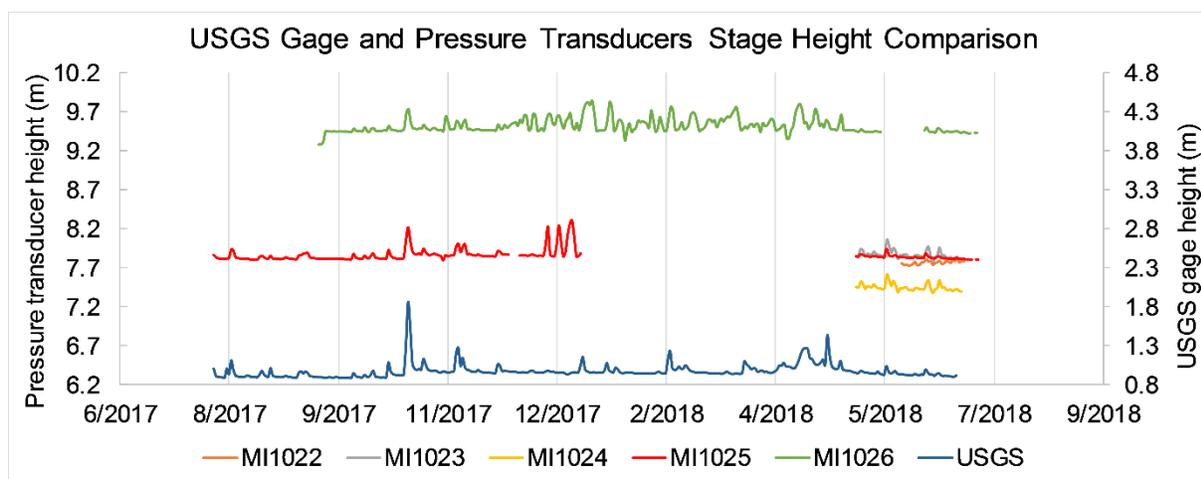


Figure 12. Comparison of USGS gage height and pressure transducer stage heights.

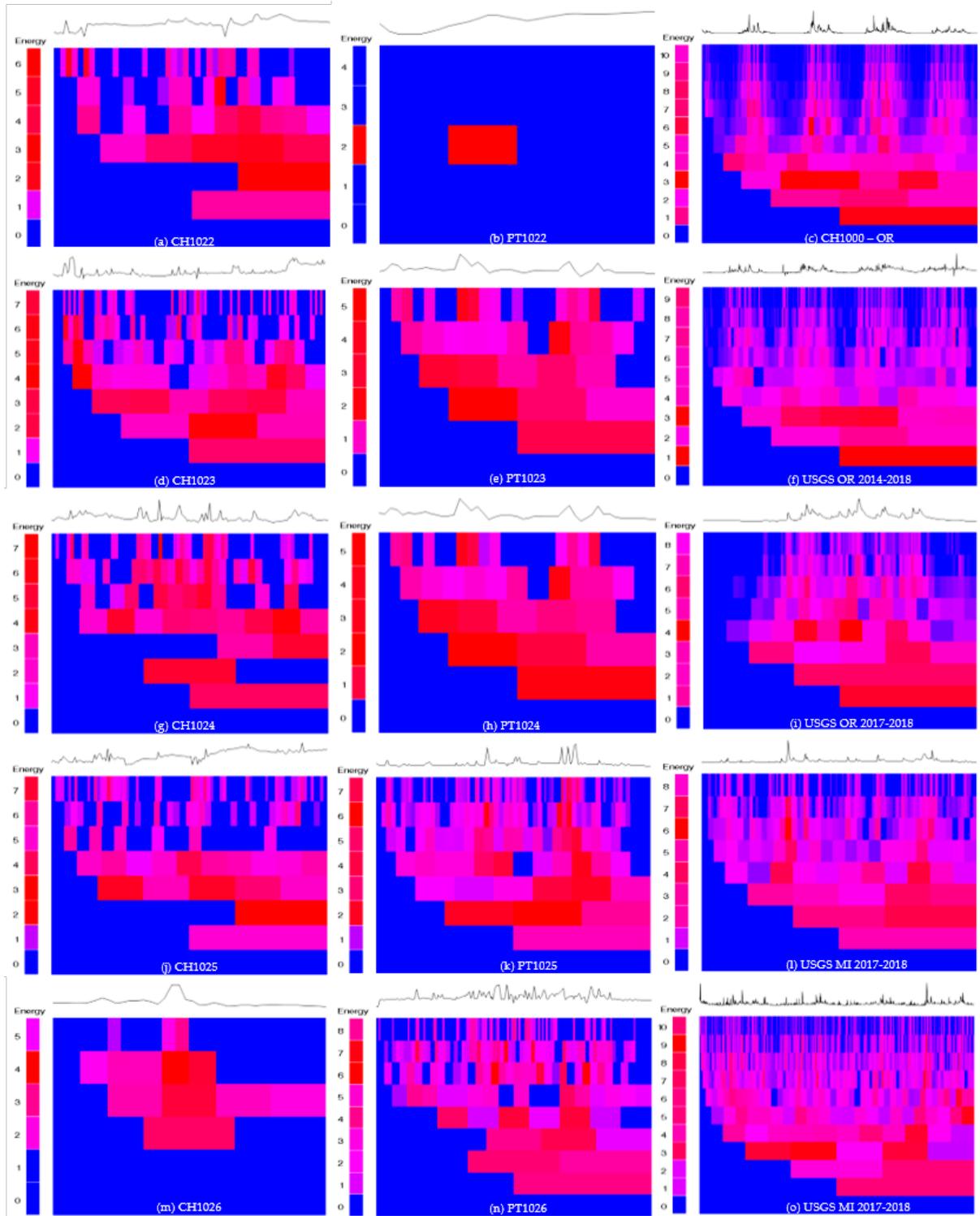


Figure 13. Scalogram images for SAS Wavelet Analysis. Left: CrowdHydrology sites CH1022-1026 (a, d, g, j, m). Middle: pressure transducer sites PT1022-1026 (b, e, h, k, n). Right: OR CH site (c), USGS OR gage 2014-2018 (f), USGS OR gage 2017-2018 (i), USGS MI gage 2017-2018 (l), USGS MI gage 2014-2018 (o).

6. Conclusion

There are many crowdsourced databases involving various data types such as hydrology, precipitation, and water quality (Appendix 2), covering the regions where no systematic monitoring existed previously. To improve the quality of citizen science data, a binary decision tree model has been generated to flag potentially erroneous data points. Optimal categorization rules were selected based on their performance for finding ‘wrong record’ (negative stage value) and ‘bias record’ (extreme value). The overall classification accuracy of the decision tree shows potential; however, the misclassification of natural extreme values can occur. The RMSE of the crowdsourced data obtained from our study is above the 0.0061 m level of the USGS gauging staff, but the percentage estimates are all very small indicating that the crowdsourced data has a high level of accuracy. RMSE may not be the best estimate of error for two observed datasets, and we should consider other methods of error estimation to determine uncertainty in crowdsourced datasets. Through further refinement of the decision tree, the use of high-accuracy citizen science data would fill NWM observational gaps and enhance output forecasts.

Acknowledgments: We would like to thank at the National Water Center: NOAA-Trey Flowers, Fernando Salas, Fred Ogden, and Ed Clark; USGS- Dave Blodgett and Martin Briggs; NCAR: Aubrey Dugger and Katelyn FitzGerald; CUAHSI: Jared Bales, Lauren Grimley, and Fernando Aristizabal. We would also like to thank Chris Lowry, Ben Ruddell, Sagy Cohen, John Brackins, and Azbina Rahman, Nishani Moragoda. Di Wu would like to thank her funding source CUAHSI, and her academic advisor Ruopu Li for his supports during this summer institution. Elizabeth Del Rosario would like to thank her funding sources CUAHSI, Harte Research Institute for Gulf of Mexico Studies and the NOAA EPP CCME program (National Oceanic and Atmospheric Administration, Office of Education Educational Partnership Program Award (NA16SEC4810009)). “Its contents are solely the responsibility of the award recipient and do not necessarily represent the official views of the U.S. Department of Commerce, National Oceanic and Atmospheric Administration. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the view of the U.S. Department of Commerce, National Oceanic and Atmospheric Administration.”

Supplementary Materials:

S.1. Literature reviewed crowdsourced projects

Table S.1. Literature review of crowdsourced water flow/water level projects.

Data source	Data description	Project objectives	Study area	Technique & algorithm	Complementary data	Reference
Water flow/levels						
YouTube	Flood movie taken by witnesses, recorded under non-ideal conditions	Gauging extreme floods for post-flood determination of stream discharges	the Saint-Julien torrent, France	Large-scale particle image velocimetry (LSPIV)		Le Boursicaud et al. 2016
YouTube	Videos of the cave recording suitable reference points like cave graffiti	Monitoring water level rise in cave	Dahl Hith, southeast of Riyadh, Saudi Arabia	Estimated the distances between the reference points and water level	Videos by own photographs taken during two site visits	Michelsen et al. 2016
Hydrological observations of social sensors used by citizens	Crowdsourced water level	Demonstrating data collected by citizens can complement traditional sensor networks and improve the accuracy of flood forecasts.	The Brue (UK), Sieve (Italy), Alzette (Luxembourg), and Bacchiglione (Italy) catchments	Lumped and semi-distributed hydrological models	Streamflow data from professional agencies	Mazzoleni et al. 2017
Flood Chasers Project	Videos and photos of flash floods in rivers recorded and shared by citizens	Estimating river flow velocity and discharge based on flood videos recorded by citizens	Argentina	LSPIV		Le Coz et al. 2016
The FloodScale project	Flood home movies from the public	Extracting surface flow velocities and discharges from flood videos	Ardèche river catchment, South-East France	Portable surface velocity radars; LSPIV	Discharge data from upstream gauging station	Le Coz et al. 2016

Table S.2. Literature review of crowdsourced hydrologic projects

Data source	Data description	Project objectives	Study area	Technique & algorithm	Complementary data	Reference
Hydrologic database generation						
Text message from passersby	Hydrologic measurements and the station number read by untrained passersby	Generating a crowdsourced database to collect and post hydrologic measurements	Buffalo, Cattaraugus, and Wiscoy Creek watersheds in Western New York, USA.	Object-oriented Python		Fienen et al. 2012; Lowry et al. 2013
Water clarity reading from volunteers	Disappearance / reappearance point of a Secchi disk	Expanding the state water-quality monitoring network to help determine the condition of Minnesota lakes and streams	Minnesota, USA			Minnesota Pollution Control Agency. 2014 (Accessed July 25 2014)
Paper data forms from trained volunteers	Coordinates of the beginning and end points of all surface water larger than 30 ft in length	Monitoring the extent of perennial surface flow in dryland regions	The San Pedro River, Arizona, USA	Wet/dry mapping	National Hydrography Dataset, 1:24,000-scale, US Geological Survey	Turner et al. 2011
The RiskScape project	Pictures of maximum flood levels from residents	Building flood hazard map to calculate flood risk and potential losses across the city	New Zealand		LiDAR; additional survey data	Le Coz et al. 2016
QLD FLOOD CRISIS MAP	A floods crowd-map posting flood-related information from citizen's email, text message, Twitter, or the website itself.	Using volunteered geographic information such as photographs and videos to assist in mapping the flood extents in regions where there was little or no mapping available	Queensland, Australia		RTK GPS point data	McDougall. 2011

Table S.3. Literature review of crowdsourced water quality project

Water Quality						
Images taken by test users mobile phone cameras	Pictures of tags located at different water depths with GPS location and timestamp	Monitoring water quality (Secchi depth and turbidity) using mobile phones and a small device designed for water quality measurements.	Finland	Automatic Secchi3000 depth analysis	Field water quality measurement	Toivanen et al. 2013

Table S.4. Literature review of crowdsourced data confidence estimation.

Data source	Data description	Project objectives	Study area	Technique & algorithm	Complementary data	Reference
Data confidence estimation						
Florida LAKEWATCH program	TP, TN, chlorophyll and Secchi disk data from volunteer samplers	Testing the reliability of volunteer collected water quality data	Florida, USA	Freezing for sample preservation; ANOVA; correlation coefficient	Professional measured data	Canfield et al. 2002
Waterwatch Victoria monitoring network	Turbidity, electrical conductivity (EC), pH and total phosphorus	Assessing confidence limits of community collected water quality data	Australia	T-tests; F-tests	Professional data from Victorian Water Quality Monitoring Network (VWQMN)	Nicholson et al 2002
Records from Secchi application on seafarers' smartphones	Secchi depth with GPS location, date and time	Estimating accuracy of ocean transparency data from seafarer in phytoplankton and climate research	Global scale	Ocean Color Index (OCI); inverse-squared distance method	MODIS Level-3 data; Bathymetry data from a global ocean bathymetry chart (0.1 latitude by 0.1 longitude)	Seafarers et al. 2017

S.2. Online crowdsourced data sources

Table S.5. Crowdsourced hydrology data.

Hydrology, water levels, stream flow									
Name	Website	Data Description	Collection Method	Location	Data Availability	Data Amount	Start Year	Region	Participants
Crowd Hydrology	www.crowdhydrology.com	water level measurements in streams and lakes	text message of stage height value from posted signs	established sites	map visualization of data points; freely accessible via download when point clicked	>16,400 measurements	2010	United States	registered volunteers, local groups
Crowd Water	www.crowdwater.ch/en/crowdwater-2/	water level, streamflow, soil moisture, flow conditions	photos to app for water level, measurements entered via app (CrowdWater SPOTTER ON)	citizen's location and established sites	photos shown on interactive map; freely accessible and displayed when photo clicked	~541 measurements	2016	Europe; North America; South America; Australia; global	registered users with app
EPA Office of Water	www.epa.gov/nps/nonpoint-source-volunteer-monitoring	watershed survey assessment, macro-invertebrates, water quality, stream flow, and ecological measurements	volunteers analyze, make full reports, and present data	volunteer's locations	visualization map displays monitoring programs https://aci.gov/monitoring/vm/program/s/vm_map.html	n/a	n/a	United States	trained volunteers
Fresh Water Watch	https://freshwaterwatch.thewaterhub.org/	water quality, flow, water level, ecological measurements	measurements uploaded from paper, web, or app based datasheet along with photo	volunteer's location and established sites	map visualization; freely accessible and displayed when data point clicked	>20,000 measurements	2012	global	organizations, school groups, volunteer groups, k-12 schools

Montana Ground-water Academy	www.umt.edu/groundwateracademy/	measurements of groundwater and groundwater-surface water interactions	data submitted online via datasheet at citsci.org	established sites	freely accessible via download www.citsci.org	316 measurements	2015	Montana, USA	registered member volunteers, high school students
Stream Flow Monitoring in Upstate New York	www.citsci.org/CWIS438/Browse/Project/Project_Info.php?ProjectID=673	stream water height measurements	data submitted online via datasheet at citsci.org	established sites	freely accessible via download www.citsci.org	~112 measurements	2015	Finger Lakes region of New York, USA	registered member volunteers
Stream Tracker	www.streamtracker.org/	record presence (water) or absence (dry) of flow	observations collected via mobile application and datasheets	established sites	freely accessible via download www.citsci.org	~6,157 measurements	2016	United States	registered member volunteers
Stream Tracker Boulder County	www.citsci.org/CWIS438/Browse/Project/Project_Info.php?ProjectID=2039	flow and water quality data measurements	data submitted online via datasheet at citsci.org	volunteer's location and established sites	freely accessible via download www.citsci.org	~152 measurements	2018	Boulder County, CO, USA	local schools
Water Pressures.org	www.waterpressures.org/	watershed assessment	2-3 minute video documentary films	citizen's location	videos freely accessible www.waterpressures.org/videos	n/a	2007	global	registered volunteers with recording devices
Watershed Education Network	www.montanawatershed.org/	stream bed assessments, macro-invertebrate, and water quality measurements	data submitted online via datasheet at citsci.org	established sites	some data freely accessible via download www.citsci.org	n/a	1996	Montana, USA	registered member volunteers, school programs, summer programs
What's your	http://arcgis.com/Pe4fp	water level reports, water depth	measurements entered	citizen's location	map visualization, freely accessible	>1000 observations	2012	United States	volunteers with

water level?			via online form		by clicking on each report			States; global	website access
WiKi Watershed	https://wikiwatershed.org/	water quality, atmospheric & water conditions measurements (depth, pressure, temperature, humidity)	sensor data imported to an interactive map-based web application, Monitor My Watershed	volunteer's sensor location	map visualization tools; data freely accessible http://data.envirodiy.org/browse/	167 sites	n/a	global	registered member volunteers with sensor

Table S.6. Crowdsourced precipitation, temperature, and tide level data.

Precipitation, temperature, and tide levels									
Name	Website	Data Description	Collection Method	Location	Data Availability	Data Amount	Start Year	Region	Participants
CoCoRaHS (Community Collaborative Rain, Hail and Snow Network)	www.cocorahs.org/	rain, hail, snow, soil moisture measurements	rain gauge, data entered via interactive web-site form	volunteer's locations	map visualization; freely accessible via download	>8000 measurements per day	2010	United States; Canada; Bahamas	registered volunteers, classroom education program, 4H
IceWatch USA™	www.natureabounds.org/IceWatch_USA.html	snow, precipitation, and ice cover measurements; wildlife activity observations	observation forms are reported via mail or email	volunteer established site	available to interested scientists	n/a	n/a	United States; Canada	registered volunteers
ISeeChange	www.iseechange.org/	flooding, agriculture, snow and ice, landscapes, coastal erosion, recharge and rain, high tides	pictures are uploaded to ISeeChange app	citizen's location	post freely accessible www.iseechange.org/investigations	>1000 observations	n/a	global	registered volunteers

		and sea level rise							
M-ping	https://mping.ou.edu	collects geotagged, time stamped reports of weather conditions near the surface	smart phone or mobile device data submission	volunteer's location	open access to view map of weather reports, must apply for access to data https://mping.ou.edu/static/mping/access.html	n/a	2012; 2016	United States; global	registered volunteers
Snow Tweet	www.snowtweet.org/	snow and ice and measured and mapped in near real time	tweet in snow depth, or via interactive web-site form	citizen's location	view on visualization tool, freely accessible via download link	n/a	2009	global	Twitter account holders
Hot Spots for Trout	www.citsci.org/CWIS438/Browse/Project/Project_Info.php?ProjectID=2147	water & air temperature measurements	data submitted online via datasheet at citsci.org	established sites	freely accessible via download www.citsci.org	~36 measurements	2018	Roaring Fork Water-shed, CO, USA	registered member volunteers
Globe Observer	https://observe.globe.gov/	temperature, precipitation, water quality, soil characterization, landcover, carbon cycling	measurements submitted via Globe Observer apps	citizen's location	interactive map with visualization tools	>76,700 measurements	1994	global	registered members with app

Table S.7. Crowdsourced water quality data.

Water quality, temperature									
Name	Website	Data Description	Collection Method	Location	Data Availability	Data Amount	Start Year	Region	Participants
EarthEcho International	www.monitorwater.org/	water quality and temperature measurements	online database entry at app.monitorwater.org	volunteer's location	interactive map visualization; freely accessible via download www.monitorwater.org/tools/reports	>109,000 measurements	2005, 2007 (map)	global	community & school organizations k-12 schools
Eye On Water.org	http://eyeonwater.org/	water color and clarity measurements	pictures submitted via EyeOnWater app	citizen's location	map visualization tools; data freely accessible www.citclops.eu/search/welcome.php	4,881 app (color); >280,000 historical (color)	2016, 1936 (historical)	global	registered members with app
Rios Saludables de Osa - Costa Rica	https://riosaludablesdeosa.org/	water quality and temperature measurements	data submitted online via datasheet at citsci.org	established sites	freely accessible via download www.citsci.org	~4,403 measurements	2014	Costa Rica	registered member volunteers
Stream Selfie	www.streamselfie.org/	water quality	photos uploaded at https://scistarter.com/data/streamselfie	citizen's location	photos shown on map location	>1,025 photos	n/a	United States; global	registered members of scistarter.com
The WQI Project	https://thewqiprject.org/	water quality	measurements entered via website	volunteer's location	map visualization; freely accessible and displayed when data point clicked	>100 measurements	2015	global	k-12 schools
Trout Unlimited Coldwater Conservation Corps Water Quality Monitoring	www.citsci.org/CWIS438/Browse/Project/Project	water quality, temperature, rain measurements	data submitted online via datasheet at citsci.org	established sites	freely accessible via download www.citsci.org	~45,478 measurements	2010	West Virginia, Virginia, USA	registered member volunteers

Utah Water Watch	www.citsci.org/CWIS438/Browse/Project/Project_Info.php?ProjectID=2043	stream and lake water quality, temperature, rain, macro-invertebrate measurements	data submitted online via datasheet at citsci.org	established sites	freely accessible via download www.citsci.org	~89,237 measurements	2002	Utah, USA	registered member volunteers
Water Trail Association	www.nywatertrail.org/water_quality.html	water quality, precipitation, tide measurements	samples are collected and taken to a lab for analysis	established sites	map visualization with data displayed; freely accessible spreadsheet download	>1,000 measurements	2011	New York City, USA	volunteers
Yojoa Watershed Measurements	www.citsci.org/CWIS438/Browse/Project/Project_Info.php?ProjectID=655	water quality, temperature, rain measurements	data submitted online via datasheet at citsci.org	established sites	access granted by current members; data download www.citsci.org	19,406 measurements	2015	Honduras Central America	registered member volunteers

References

1. Buytaert, W.; Zulkafli, Z.; Grainger, S.; Acosta, L.; Alemie, T.C.; Bastiaensen, J; ... Foggin, M. Citizen Science in hydrology and water resources: opportunities for knowledge generation, ecosystem service management, and sustainable development. *Frontiers in Earth Science* 2014, 2, 26.
2. [NOAA] National Oceanic and Atmospheric Association. The National Water Model. Available online: <http://water.noaa.gov/about/nwm> (accessed 27, June 2018)
3. [NOAA] National Oceanic and Atmospheric Association. National Water Model: improving NOAA's water prediction services. Office of Water Prediction. Published on August 2016. Available online: <http://water.noaa.gov/documents/wrn-national-water-model.pdf> (accessed 16, July 2018).
4. Downing, J.A.; Cole, J.J.; Duarte, C.M.; Middelburg, J.J.; Melack, J.M.; Prairie, Y.T.; Kortelainen, P.; Striegl, R.G.; McDowell, W.H.; Tranvik, L.J. Global abundance and size distribution of streams and rivers. *Inland Waters* 2012, 2(4), 229–236, <https://doi.org/10.5268/IW-2.4.502>.
5. Gauchat, G. Politicization of science in the public sphere: A study of public trust in the United States, 1974 to 2010. *American Socio. Review* 2012, 77(2) 167-187, DOI: <http://dx.doi.org/10.1177/0003122412438225>.
6. Law, E.; Gajos K.Z.; Wiggins, A.; Gray, M.; William, A. Crowdsourcing as a tool for research: implications of uncertainty. *Proceedings of the ACM* 2017, 1544-1561, DOI: <http://dx.doi.org/10.1145/2998181.2998197>.

7. Fielen, M.N.; Lowry, C.S. Social.Water-A crowdsourcing toll for environmental data acquisition. *Computers & Geosciences* 2012, 49, 164-169, DOI: <http://dx.doi.org/10.1016/j.cageo.2012.06.015>.
8. Lowry, C.S.; Fielen, M.N. CrowdHydrology: Crowdsourcing hydrologic data and engaging citizen scientists. *Ground Water* 2012, DOI: <http://dx.doi.org/10.1111/j.1745-6584.2012.00956.x>.
9. [USGS] United States Geological Survey. Water Data for the Nation. 2018. Available online: <https://waterdata.usgs.gov/nwis> (accessed 20, July 2018).
10. Dieter, C.A.; Linsey, K.S.; Caldwell, R.R.; Harris, M.A.; Ivahnenko, T.I.; Lovelace, J.K.; Maupin, M.A.; Barber, N.L. Estimated use of water in the United States county-level data for 2015 (ver. 2.0, June 2018): U.S. Geological Survey data release. 2018. Available online: <https://doi.org/10.5066/F7TB15V5> (accessed 6, July 2018).
11. Blodgett, D. NWM V1.2 Retrospective Data for CrowdHydrology Sites, HydroShare, 2018. Available online: <http://www.hydroshare.org/resource/38818751e8d9437c83b39a43589f6c7a> (accessed 6, July 2018).
12. Dickinson, J.L.; Shirk, J.; Bonter, D.; Bonney, R.; Crain, R.L.; Martin, J.; ... Purcell, K. The current state of citizen science as a tool for ecological research and public engagement. *Frontiers Ecology & Environment* 2012, 10(6), 291-297.
13. [NOAA] National Oceanic and Atmospheric Association. Citizen science and crowdsourcing. Available online: <http://www.noaa.gov/office-education/citizen-science-crowdsourcing> (accessed 27, June 2018).
14. McKinley, D.C.; Miller-Rushing, A.J.; Ballard, H.L.; Bonney, R.; Brown, H.; Cook-Patton, S.C.; ... Ryan, S. F. Citizen science can improve conservation science, natural resource management, and environmental protection. *Biological Conservation* 2017, 208, 15-28.
15. Johnson, M.F.; Hannah, C.; Acton, L.; Popovici, R.; Karanth, K. K.; Weinthal, E. Network environmentalism: Citizen scientists as agents for environmental advocacy. *Global Environmental Change* 2014, 29, 235-245.
16. Bowser, A.; Shanley, L. New visions in citizen science. *Woodrow Wilson Centre* 2013. Available online: <http://www.wilsoncenter.org/publication/new-visions-citizen-science> (accessed 27, June 2018).
17. Ries, L.; Oberhauser, K.A citizen army for science: quantifying the contributions of citizen scientists to our understanding of monarch butterfly biology. *BioScience* 2015, 65(4), 419-430.
18. Silvertown, J. A new dawn for citizen science. *Trends in Ecology & Evolution* 2009, 24(9), 467-471.
19. Miller-Rushing, A.; Primack, R.; Bonney, R. The history of public participation in ecological research. *Frontiers in Ecology and the Environment* 2012, 10(6), 285-290.
20. Sauer, J.R.; Link, W.A.; Fallon, J.E.; Pardieck, K.L.; Ziolkowski Jr, D.J. The North American breeding bird survey 1966–2011: summary analysis and species accounts. *North American Fauna* 2013, 79(79), 1-32.
21. Poelen, J.H.; Simons, J.D.; Mungall, C.J. Global biotic interactions: an open infrastructure to share and analyze species-interaction datasets. *Ecological Informatics* 2014, 24, 148-159.
22. Hemmi, A.; Graham, I. Hacker science versus closed science: Building environmental monitoring infrastructure. *Information. Communication & Society* 2014, 17(7), 830-842.
23. McCormick, S. After the cap: risk assessment, citizen science and disaster recovery. *Ecology and Society* 2012, 17(4).

24. Sullivan, B.L.; Wood, C.L.; Iliff, M.J.; Bonney, R.E.; Fink, D.; Kelling, S. eBird: A citizen-based bird observation network in the biological sciences. *Biological Conservation* 2009, 142(10), 2282-2292.
25. Le Coz, J.; Patalano, A; Collins, D.; Guillén, N.F.; García, C.M.; Smart, G.M.; ... Braud, I. Crowdsourced data for flood hydrology: Feedback from recent citizen science projects in Argentina, France and New Zealand. *Journal of Hydrology* 2016, 541, 766-777.
26. Toivanen, T.; Koponen, S.; Kotovirta, V.; Molinier, M.; Chengyuan, P. Water quality analysis using an inexpensive device and a mobile phone. *Environmental Systems Research* 2013, 2(1), 9.
27. Seafarers, S.D.; Lavender, S.; Beaugrand, G.; Outram, N.; Barlow, N.; Crotty, D.; ... Kirby, R. Seafarer citizen scientist ocean transparency data as a resource for phytoplankton and climate research. *PLoS one* 2017, 12(12), e0186092.
28. Canfield Jr, D.E.; Brown, C.D.; Bachmann, R.W.; Hoyer, M.V. Volunteer lake monitoring: testing the reliability of data collected by the Florida LAKEWATCH program. *Lake and Reservoir Management* 2012, 18(1), 1-9.
29. Nicholson, E.; Ryan, J.; Hodgkins, D. Community data-where does the value lie? Assessing confidence limits of community collected water quality data. *Water Science and Technology* 2002, 45(11), 193-200.
30. Le Boursicaud, R.; Pénard, L.; Hauet, A.; Thollet, F.; Le Coz, J. Gauging extreme floods on YouTube: application of LSPIV to home movies for the post-event determination of stream discharges. *Hydrological Processes* 2016, 30(1), 90-105.
31. Michelson, N.; Dirks, H.; Schulz, S.; Kempe, S.; Al-Saud, M.; Schüth, C. YouTube as a crowd-generated water level archive. *Science of the Total Environment* 2016, 568, 189-195.
32. Turner, D.S.; Richter, H.E. Wet/dry mapping: using citizen scientists to monitor the extent of perennial surface flow in dryland regions. *Environmental Management* 2011, 47(3), 497-505.
33. Kampf, S.; Strobl, B.; Hammond, J.; Anenberg, A.; Etter, S.; Martin, C.; Punttenney-Desmond, K.; Seibert, J.; van Meerveld, I. Testing the waters: Mobile apps for crowdsourced streamflow data. *Eos* 99, <https://doi.org/10.1029/2018EO096355>. Published on 12 April 2018.
34. Wahl, K.L.; Thomas, Jr., W.O.; Hirsch, R.M. Stream-Gaging Program of the U.S. Geological Survey. U.S. Geological Survey Circular 1123, Reston, Virginia, 1995. Available online: <https://pubs.usgs.gov/circ/circ1123/collection.html> (accessed 23, July 2018).
35. [USGS] United States Geological Survey. How Streamflow is Measured Part 3: The stage-discharge relation. 2016. Available online: <https://water.usgs.gov/edu/streamflow3.html> (accessed 20, July 2018).
36. Grace-Martin, K. Assessing the fit of regression models. *The Analysis Factor* 2018. Available online: <https://www.theanalysisfactor.com/assessing-the-fit-of-regression-models/> (assessed 21, July 2018).
37. Lemon, S.C.; Roy, J.; Clark, M.A.; Friedmann, P.D.; Rakowski, W. Classification and regression tree analysis in public health: methodological review and comparison with logistic regression. *Annals of behavioral medicine* 2003, 26(3), 172-181.
38. Friedl, M.A.; Brodley, C.E. Decision tree classification of land cover from remotely sensed data. *Remote sensing of environment* 1997, 61(3), 399-409.
39. Anyanwu, M.N.; Shiva, S.G. Comparative analysis of serial decision tree classification algorithms. *International Journal of Computer Science and Security* 2009, 3(3), 230-240.
40. Safavian, S.R.; Landgrebe, D. A survey of decision tree classifier methodology. *IEEE transactions on systems, man, and cybernetics* 1991, 21(3), 660-674.

41. Quinlan, J R. Generating production rules from decision trees. In *ijcai* (Vol. 87, pp. 304-307) 1987, August.
42. American Meteorological Society. Glossary of Meteorology. *American Meteorological Society* 1959, 45 Beacon St. Boston, MA, 638pp.
43. Merriam Webster's Medical Dictionary. 2016
44. Goutte, C.; Gaussier, E. A probabilistic interpretation of precision, recall and F-score, with implication for evaluation. In *European Conference on Information Retrieval* (pp. 345-359). March 2005. Springer, Berlin, Heidelberg.
45. Powers, D.M. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. *Journal of Machine Learning Technologies* 2011, 2(1), 37-63. Available online https://bioinfopublication.org/files/articles/2_1_1_JMLT.pdf (accessed 22, July 2018).
46. [NCSU] North Carolina State University. Appendix B: Percent Error and Percent Difference. 2010. Labs for colleges physics mechanics. 2nd ed. Available online: www.webassign.net/labsgraceperiod/ncsulcpmech2/appendices/appendixB/appendixB.html (assessed 20, July 2018).
47. Blake, S. Whole Course Items: Error and Uncertainty. Available online: <https://www.nhn.ou.edu/~johnson/Education/Juniorlab/Error-SigFig/SigFiginError-043.pdf> (accessed 17, July 2018).
48. Buchanan, T.J.; Somers, W.P. Techniques of Water-Resources Investigations of the United States Geological Survey. Chapter A7, Stage Measurement at Gaging Stations. Book 3, *Applications Of Hydraulics*. United States Government Printing Office, Washington: 1982. Available online: <file:///E:/NWC%20Project/References/report.pdf> (accessed 23, July 2018).
49. SAS Institute Inc. Wavelet Analysis. SAS/IML(R) 9.22 User's Guide. Available online: http://support.sas.com/documentation/cdl/en/imlug/63541/HTML/default/viewer.htm#waveletanalysis_toc.htm (accessed 20, July 2018)

Appendix

Summer Institute Theme Leaders



Jude A. Benavides, Ph.D.
Associate Professor
UT Rio Grande Valley



Fred Ogden, Ph.D., P.E., P.H.
Senior Scientist
National Water Center



Sarah Praskievicz, Ph.D.
Assistant Professor
University of Alabama



Ehab Meselhe, Ph.D., P.E.
Professor / VP for Engineering
Tulane/Water Institute of the
Gulf



Kyle Mandli, Ph.D.
Assistant Professor
Columbia University



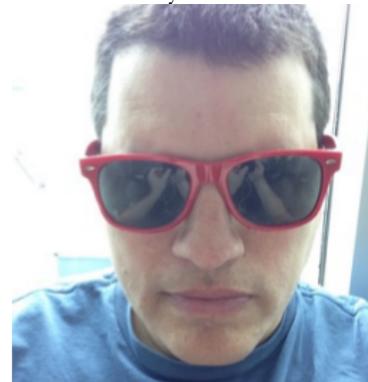
Sagy Cohen, Ph.D.
Associate Professor
University of Alabama



Chris Lowry, Ph.D.
Associate Professor
University of Buffalo



David Steward, Ph.D.
Professor
North Dakota State University



Joseph Hughes, Ph.D.
Hydrologist
USGS

Summer Institute Course Coordinators



Fernando Aristizabal, University of Florida
Lauren Grimley, University of Iowa

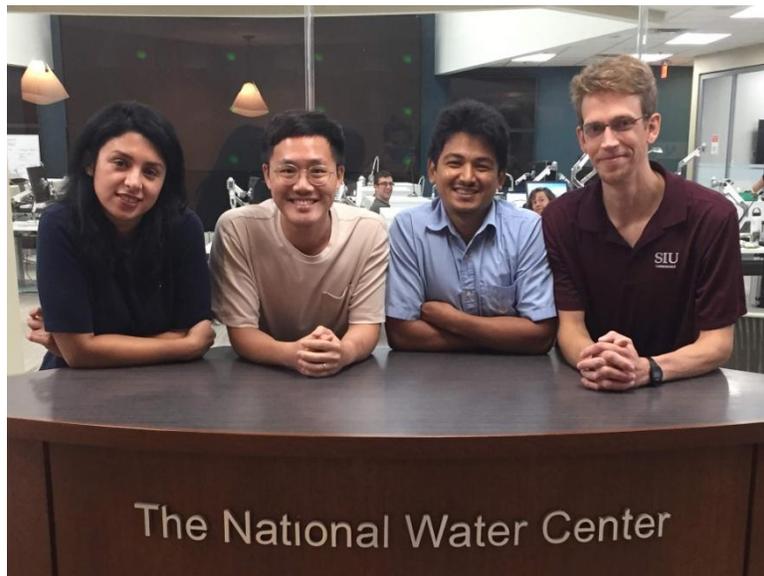
Summer Institute Research Fellows

Chapter 1: Evaluating the National Water Model Performance in Losing Streams
(Team Loss but Not Forgotten)



Holly Hutcheson, University of Georgia
Matt Thomas, University of Georgia
Liz Jachens, Oregon State University

Chapter 2: Evaluating Alternative Baseflow Estimations in the National Water Model for Improving Hydrograph Forecasting (Team The Bucket List)



Sama Memari, University of Alabama
Minki Hong, Texas A&M University
Ritesh KarKi, Auburn University
Joe Krienert, Southern Illinois University

Chapter 3: Discerning When to Initiate Hyper-Resolution Modeling in Low Gradient Watersheds through Dynamical Forcings (Team DAFT Sunk)



Andrew Goenner, Iowa State University
Tasnuva Rouf, George Mason University
Deeoa Gurung, University of Alabama
Francisco Perez, Michigan Technological University

Chapter 4: Sensitivity of Urban Flooding to Subsurface Storm Drainage Systems in Low-Gradient Watersheds (Team Linkin Pipes)



Parth Modi, Virginia Tech
Maggie Zarekarizi, Portland State University
Isha Deo, University of Texas at Austin
Jean Valle, University of Puerto Rico

Chapter 5: Using Dimensionless Scaling Parameters as Decision Metrics for Heterogenous Routing Schemes (Team FORTRAN Fanclub)



Maryam Lamjiri, University of California, San Diego
Kelly Flint, San Diego State University
Sean Matus, University of Illinois at Urbana-Champaign

Chapter 6: Improvements Upon Simplified Channel Geometry Assumptions for Continental-Scale Hydrological Modeling (Team Here's Looking at Euclid)



John Brackins, Tennessee Tech
Azbina Rahman, George Mason University
Nishani Moragoda, University of Alabama

Chapter 7: Exploration of Citizen Science Data and Potential Application to the National Water Model (Team Decision Tree)



Liz Del Rosario, Texas A&M University Corpus Christi
Di Wu, Southern Illinois University