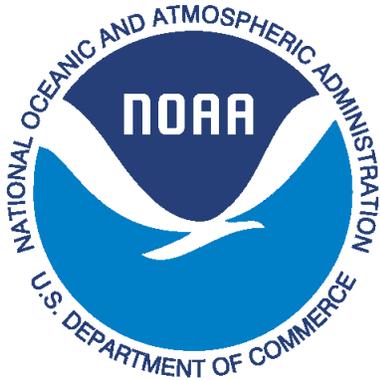


TECHNICAL REPORT 16

September 2019



OWP | OFFICE OF
WATER
PREDICTION



National Water Center
Innovators Program
Summer Institute Report 2019

National Water Center Innovators Program Summer Institute Report 2019

Editors:

Kelly Flint

Maryam Asgari Lamjiri

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. 16

DOI: <https://doi.org/10.4211/hs.096e7badabb44c9f8c29751098f83afa>

Version 1.02

September 2019

Suggested Citation:

Flint, K., M. Asgari Lamjiri, et al. (2019). National Water Center Innovators Program - Summer Institute, CUAHSI Technical Report, HydroShare, <http://www.hydroshare.org/resource/096e7badabb44c9f8c29751098f83afa>

Contents

Preface	3
Chapter 1	
Influence of Forcing Conditions on Total Water Level Prediction in Delaware Bay, USA	6
Chapter 2	
Model-Based Parametric Analysis of Total Water Prediction in Coastal Transition Zones of the US East and Gulf Coasts	20
Chapter 3	
Experimental One-way Coupling of TOPMODEL with NWM to Substitute Runoff Processes Delineation in a Headwater Catchment	32
Chapter 4	
A Study on Parsimonious Models in Catchments Generating Saturation Excess Runoff	45
Chapter 5	
National Water Model Dockerized Job Scheduler: A Reproducible Framework to Generate Parameter-based NWM Ensemble	57
Chapter 6	
A Visualization Workflow for Quantifying Parameter Sensitivities to Uncertainties for Hydrologic Models	66

Preface

The National Water Center Innovators Program: Summer Institute (NWC-SI) was first organized in 2015 by the National Water Center and Consortium of Universities for the Advancement of Hydrologic Science Inc. (CUAHSI). CUAHSI, a nonprofit organization of more than 130 member institutions, provides programs and services for advancing interdisciplinary water science. The National Water Center (NWC) is a National Oceanic and Atmospheric Administration (NOAA) facility that enables cooperation across federal agencies to deliver a new generation of water information and services to the nation. This partnership has allowed the NWC-SI to facilitate collaboration and ideation across disciplines and backgrounds, equipping existing and future scientists to create sustainable solutions with others, and for others.

Throughout the NWC-SI, graduate students (fellows) and senior academic faculty and federal scientists (theme leaders) have participated alongside NWC staff and other senior scientists to create and complete scientific projects geared toward improving U.S. water resources modeling, science, and services. Included in these is the National Water Model (NWM), a tool that more efficiently provides guidance for water resources related risks, which became operational in August 2016. NWC-SI projects have revolved around scientific and modeling challenges, prioritized by the NWC and categorized by 'themes'. Themes for NWC-SI 2019 were: 'Scaling of Hydrologic Processes'; 'Hydroinformatics', and 'Coupled Inland-Coastal Hydraulics'. During the first week of NWC-SI 2019, fellows were introduced to the themes, the NWC mission, and the NWC's current operational priorities for the NWM. Fellows and theme leaders participated in team formation activities, project brainstorming sessions, and data and project management training that included a multi-day, interactive introduction to GitHub, Jupyter, and Docker. These tools and training laid the foundation for team and project formation. Commencing the second week, fellows selected the themes around which they would frame their projects. Theme leaders provided theme-specific training that enabled fellows to develop the projects they worked on for the remaining five weeks. The following chapters provide an overview of NWC-SI 2019 participants and their projects.

Fellows

The 5th NWC-SI Fellows class consisted of 18 graduate students, three M.S. students, and 15 PhD candidates from 16 universities across the U.S. This cohort represented various academic backgrounds that included: civil, biological and environmental engineering; geography; environmental resource and policy; oceanography; coastal and marine science; computer science, and informatics. The diversity of fellows and perspectives is an integral and celebrated aspect of the NWC-SI.

Themes and Theme Leaders

The NWC-SI 2019 themes and theme leaders were:

- The 'Coupled Inland-Coastal Hydraulics' theme, led by Celso Ferreira of George Mason University. Additional technical support was provided by Patrick Burke of National Ocean Service and Roham Bakhtyar of the NWC.
- The 'Scaling of Hydrologic Processes' theme led by Fred Ogden of University Corporation for Atmospheric Research (UCAR)/National Water Center, and Hilary McMillan of San Diego State University.
- The 'Hydroinformatics' theme, led by David Blodgett of the United States Geological Survey (USGS) and Kyle Mandli of Columbia University. Additional technical support was provided by Nels Frazier of the National Water Center (NWC).

Project Summaries

The following provides a brief 2019 NWC-SI projects summary, presented in more detail in Chapters 1-6 of this report.

1. Projects Related to the ‘Coupled Inland-Coastal Hydraulics’ Theme

The project ‘Influence of Forcing Conditions on Total Water Level Prediction in Delaware Bay, USA’ analyzed the influence of dynamical forcing conditions (e.g., wind, waves, currents and river discharge) on total water level prediction, to understand spatial and temporal variations of total water level within the inland-coastal transition zone. Delft3D-FM and HEC-RAS, using different combinations of systematically varied input forcing conditions, were applied to model total water level in Delaware Bay during storm-surges produced by Hurricanes Isabel and Sandy. The project ‘Model-Based Parametric Analysis of Total Water Prediction in Coastal Transition Zones of the US East and Gulf Coasts’ investigated input model parameter contributions on output variability and possible parameter interactions, to provide insight for improved operational model calibration in coastal transition zones. A variety of idealized coastal zone domain geometries and physical input parameters were used to produce model output on which to perform targeted parametric and performance analyses. These two projects are described in Chapters 1 and 2, respectively.

2. Projects Related to the ‘Scaling of Hydrologic Processes’ Theme

The project ‘Experimental One-Way Coupling of TOPMODEL with NWM to Substitute Runoff Processes Delineation in A Headwater Catchment’ explored the potential to achieve a balance between computational efficiency and fidelity in hydrologic modeling, by integrating a simple topography-based hydrological model component into nationwide hydrologic modeling. NWM v1.2 and NWM’s LSM+TOPMODEL coupled version were used to model runoff generation in a headwater catchment, and their respective performances were evaluated and compared. The project ‘A Study on Parsimonious Models in Catchments Generating Saturation Excess Run-Off’ introduced a parsimonious rainfall-runoff model built on the assumptions of TOPMODEL, to provide insight on the role of saturation excess on runoff generation and to suggest areas where NWC can apply the model. Our model uses higher resolution topography and simple subsurface flow equations in a computationally efficient fashion to model overland flow. These two projects are described in more detail in Chapters 3 and 4, respectively.

3. Projects Related to the ‘Hydroinformatics’ Theme

Similar to the other two themes of 2019 NWC-SI, the hydroinformatic theme involved two projects. The project ‘National Water Model Dockerized Job Scheduler: A Reproducible Framework to Generate Parameter-Based NWM Ensemble’ focused on developing a generalized framework to efficiently generate performance-based NWM ensemble outputs. The framework used a pseudo-random Latin hypercube approach to scale channel parameters, utilized Docker to deploy and manage asynchronous NWM containers, and created a weighted average ensemble output based on a rank system. The project ‘A Visualization Workflow for Quantifying Parameter Sensitivities to Uncertainties for Hydrologic Models’ developed a reproducible workflow for evaluating parameter sensitivities and uncertainties, using the hydrologic modeling framework of the NOAA National Water Model. The workflow evaluated model output as a function of parameter choice, using both numerical and visualization techniques. Four case studies, generated from data provided by graduate student researchers at the 2019 NWC SI, demonstrate the workflow implementation. These two projects are described in more detail in Chapters 5 and 6, respectively.

Acknowledgements

It can be appreciated that an activity of this magnitude involves a great deal of organization. Jerad Bales and Danielle Tijerina of CUAHSI, and Pamela Harvey of the University of Alabama, organized institutional arrangements and also managed travel, housing, and living arrangements in Tuscaloosa, AL. Brent Hargis of the United States Army Corps of Engineers (USACE) provided team science and project management expertise during the first two weeks of the NWC-SI, equipping fellows with strategies to navigate challenges throughout the remaining program. Course Coordinators, Maryam Asgari Lamjiri and Kelly Flint, were instrumental in planning the first two weeks, coordinating throughout the NWC-SI, and providing on-site participant support. The contribution from our hosts at the University of Alabama is greatly appreciated.

A key to the NWC-SI success is the support it receives through voluntary collaboration from the academic community, and from various other partners. Since the first NWC-SI in 2015, more than 145 graduate students have experienced working together in NWC group research projects. Of equal importance to the technical progress participants make, are the friendships and professional networks formed that they carry 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 for the opportunity to contribute to the enhancement of water prediction for our nation.

Kelly Flint

Student Course Coordinator, NWC Innovators Program - Summer Institute 2019
Department of Civil, Construction & Environmental Engineering, San Diego State University

Maryam Asgari Lamjiri

Student Course Coordinator, NWC Innovators Program - Summer Institute 2019
Scripps Institution of Oceanography, University of California, San Diego

Chapter 1

Influence of Forcing Conditions on Total Water Level Prediction in Delaware Bay, USA

David F. Muñoz¹, Dongxiao Yin², and Jiannan Tian³

¹Department of Civil and Environmental Engineering, The University of Alabama;
dfmunoz1@crimson.ua.edu

²Department of Oceanography and Coastal Sciences, Louisiana State University; *dyin2@lsu.edu*

³Department of Computer Science, The University of Alabama; *jtian10@crimson.ua.edu*

Summer Institute Theme Advisor: Celso Ferreira, George Mason University, *cferrei3@gmu.edu*

Academic Advisors: Hamed Moftakhari, *The University of Alabama*; Zuo George Xue, *Louisiana State University*; Dingwen Tao, *The University of Alabama*

Abstract: Accurate forecasts of total water level (i.e., a combination of tides, surge, wave and freshwater components) are of vital importance to stakeholders and federal agencies that need to rapidly adopt strategies for potential flooding hazards. In this study, we analyze the influence of dynamical forcing conditions, wind, surge and river discharge, on total water level (TWL) prediction in Delaware Bay, USA. A systematic set of scenarios was generated in Delft3D-FM to quantify each component's contribution to modeled TWL for Hurricane Isabel (2003) and Hurricane Sandy (2012). In both hurricanes, storm surge-induced water level was the main contributor to TWL, followed by astronomical tide. Riverine discharge induced-water level is small compared to the other components. Analyses of TWL spatial variation and the temporal variation of prediction error suggest wind forcing plays a key role in TWL prediction, followed by river discharge. Inter-model comparison of Delft3D-FM and HEC-RAS was conducted to provide insight regarding the importance of model capabilities when modeling storm surges. Our results suggest the wind module of Delft3D-FM greatly improves TWL prediction, particularly at TWL peak. However, 2D HEC-RAS is a simpler alternative for modeling storm surge events when wind forcing is not relevant in the model domain.

1. Motivation

It is estimated that by 2050, 25% of the world's population will live in flood-prone areas [1]. In the United States, over half the population is settled in coastal regions [2], and 23 of the 25 most densely populated counties are located on the coast [3]. The North-Atlantic coasts are particularly vulnerable to coastal flooding caused by extreme sea levels during the Atlantic hurricane season [4, 5]. The National Water Center (NWC) has been providing water forecast services since 2017 in partnership with several federal agencies. Coupling the National Water Model (NWM) with appropriate ocean models is important in order to forecast compound flooding from storm surge and upland runoff. The complex interaction of dynamical forcing, among other factors (anthropogenic activities, land cover change, etc.), in Coastal Transition Zones (CTZs) impedes the NWM's ability to provide accurate Total Water Level (TWL) forecasts as compared to inland areas. A previous study of the Delaware River Basin explored the benefits of model coupling (i.e., 1D/2D Delft3D-FM and 1D HEC-RAS) on TWL forecasts in CTZs [6], however, the contribution of tides, river discharge and TWL surge and its spatial and temporal variation in the model domain, have not been analyzed. This study is based on the work of Maitaria et al., [7] with the 2D mesh and model inputs formerly used to simulate storm-surge events.

2. Objectives and Scope

2.1 Objectives

This study aimed to: 1) Investigate relevant forcing conditions influencing TWL prediction in the Delaware CTZ and 2) Evaluate the accuracy of TWL, simulated with two hydrodynamic models. For this purpose, a systematic set of scenarios was modeled in Delft3D-FM. TWL prediction accuracy was also evaluated through an inter-model comparison of both 1D and 2D HEC-RAS models and the Delft3D-FM model.

2.2 Study Area

Delaware Bay is on the east coast of the U.S., with a drainage basin of 35,000 km². The Delaware River and Schuylkill River contribute 58% and 15%, respectively, of freshwater discharge to the bay. During normal flow conditions, water motion in the bay is dominated by tidal currents. Data used in this study came from five stations managed by the National Oceanic and Atmospheric Administration (NOAA) and the United States Geological Survey (USGS) in the bay. The stations are in the CTZ (from the estuary mouth to the upstream part of the Delaware River), providing water level and meteorological records. Figure 1 and Table 1 show the five stations and their distance relative to Brandywine Shoal Light station.

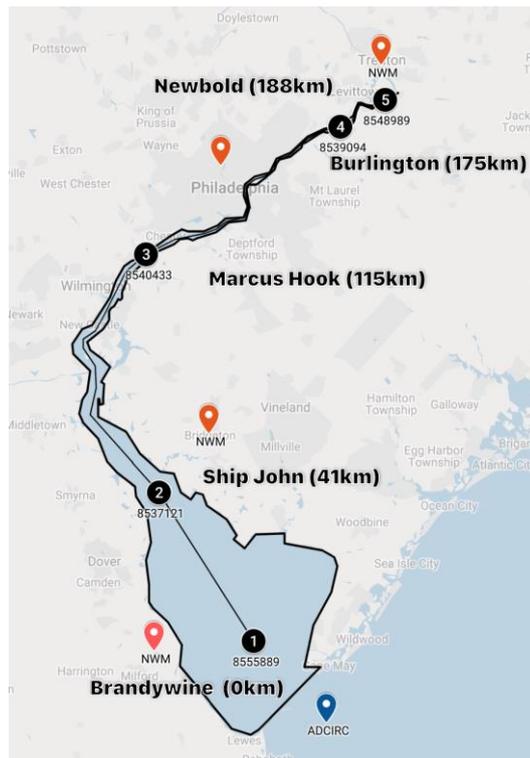


Figure 1. Location of Delaware Bay, USA.

Table 1. Selected stations, with location and distances from Brandywine Shoal Light, DE.

	Station Name	Station ID	Distance [km]
1	Brandywine Shoal Light, DE	8555889	0
2	Ship John Shoal, NJ	8537121	41
3	Marcus Hook, PA	8540433	115
4	Burlington, Delaware River, NJ	8539094	175
5	Newbold, PA	8548989	188

3. Previous Studies

The TWL in the coastal transition zone contains components induced by astronomical tides, storm surge, river discharge and waves. Numerous studies have used numerical methods to simulate TWL and evaluate the relative contribution of different components to it. By extending the 1D HEC-RAS model downstream, Mashriqui et al. [8] simulated water levels in tidal Potomac River with acceptable results. Serafin and Ruggiero [9] presented a simulation model to calculate various TWL components induced by waves, tides, and non-tidal residuals. Marsooli and Lin [10], by coupling ADCIRC with the SWAN model, studied storm surge water level components and surge-tide interaction along the U.S. East and Gulf Coasts during past tropical cyclones.

Previous efforts on inland to coastal coupling in the CTZ have focused on the Delaware River Basin [7]. A systematic comparison of 1D and 1D/2D hydrodynamic modeling revealed 1D HEC-RAS fails to reliably predict TWL during wind-dominated storm-surges. The 1D/2D coupling in DFLOW-FM, however, led to improved performance in both riverine and estuarine domains, due mainly to its wind-forcing module. While several studies have focused on model response sensitivities of physical processes [11–13], the contribution of forcing components to TWL, and its spatial and temporal variation in extreme storm-surge events, needs to be further explored. Additionally, comparisons between the 2D HEC-RAS model, 1D HEC-RAS and 1D/2D DFLOW-FM have not been conducted in Delaware Bay. In this study, relative contributions of storm surge, astronomical tide, and riverine discharge to TWL were studied for Hurricane Isabel and Hurricane Sandy. Finally, an inter-model comparison of HEC-RAS and Delft3D-FM aimed to add to findings of Maitaria et al. [7], in Delaware Bay.

4. Methodology

4.1. Model framework

To evaluate the influence of wind, surge and river discharge on TWL prediction, a systematic set of scenarios was devised and modeled in Delft3D-FM. Eleven scenarios were derived from different combinations of forcing conditions and input data sources (e.g., climate and oceanic models). Scenarios g33 and g10/g14 will subsequently be referred to as ‘base model’ and ‘reference model’, respectively (Table 2).

Table 2. Selected scenarios for various analyses in the Coastal Transition Zone of Delaware Bay.

Scenarios (Isabel/Sandy)	Description	Objective
g33	<i>Base model: Astronomical tide at the ocean boundary</i>	Individual contribution of forcing conditions on TWL
g34	Base model + River discharge	
g26/30	Base model + River discharge + Wind and atmospheric pressure obtained from CFSR*/HWRF**	
g10/g14	<i>Reference model: g26/30 + wave- induced water level</i>	Spatial and temporal variation of TWL
g2	Reference model - Wind and atmospheric pressure	
g9/g13	Reference model - River discharge	
g42/g46	Reference model - Surge	

* Climate Forecast System Reanalysis

** Hurricane Weather Research and Forecasting

4.2. Investigation of Relevant Forcing Conditions

4.2.1 Individual Contribution of Forcing Conditions to Total Water Level

A series of experiments was designed and conducted under the model framework described above, using the setup shown in Table 3. Experiments I1 and S1 calculated astronomical tide, while experiments I2 and S2 studied the influence of river discharge on water level. Under the assumption that surge-tide interaction is relatively small compared to the surge signal, experiments I3 and S3 quantified storm surge-induced water level variability when combined with experiments I2 and S2.

An intensive test study revealed surface wind and atmospheric pressure field datasets from Hurricane Weather Research and Forecast (HWRF) produced better model performance for Hurricane Sandy, while the same variables, provided by Climate Forecast System Reanalysis (CFSR), produced better model performance for Hurricane Isabel.

4.2.2 Spatial Variation of Total Water Level

Two approaches were followed to analyze TWL spatial variation in terms of Root Mean Squared Error (RMSE), bias, correlation coefficient (R^2), and model skill. The first approach compared the TWL simulated in Delft3D-FM with observations obtained from five main tidal/gauge stations in the study area (Table 1). The second approach compared TWL, simulated by a calibrated and validated Delft3D-FM model, to TWL simulated with scenarios g2, g9/g13 and g42/g46. The scenarios highlighted the influence of each forcing condition on TWL prediction (Table 2). This approach took advantage of TWL simulated over the entire model domain, enabling analysis at locations where observations were unavailable.

Table 3. Experiments setup for Hurricane Isabel and Hurricane Sandy

Experiments	River Boundary	Oceanic Boundary (water level)	Wind and pressure	Combination
<i>Experiment setup for Hurricane Isabel</i>				
I1 (g33)	N/A	Astronomical tide (AT)	N/A	(AT)
I2 (g34)	Discharge (D)	Astronomical tide (AT)	N/A	(RI) + (AT)
I3 (26)	Discharge(D)	Storm tide (ST)	Wind field Atmospheric pressure field	(RI) + (ST)
<i>Experiment setup for Hurricane Sandy</i>				
S1 (g33)	N/A	Astronomical tide (AT)	N/A	(AT)
S2 (g34)	Discharge (D)	Astronomical tide (AT)	N/A	(RI) + (AT)
S3 (g30)	Discharge(D)	Storm tide (ST)	Wind field Atmospheric pressure field	(RI) + (ST)

4.2.3 Temporal Variation of Total Water Level Error

TWL temporal variation was analyzed regarding error in prediction (i.e., scenarios vs. reference model) at three instances: 1) high water level (HWL) before the peak, 2) peak water level, and 3) HWL after the peak. These instances were considered, to study both over- and underestimation patterns in TWL prediction over the entire model domain. Timing of high and peak water levels of Hurricane Isabel and Hurricane Sandy were obtained from observations measured at Ship John Shoal station, located in the middle of the CTZ (Figure 1).

4.3 Computational Platform and Testbench

4.3.1 Research Platform

Conventional high-performance computing (HPC) employs parallelization to achieve speedups. Open Multi-Processing (OpenMP) [14] and Message Passing Interface (MPI) from which Open MPI [15] is predominately used, dominate real-world use scenarios. In comparison, research-oriented projects (e.g., our ongoing work) require heavy workloads utilizing GUI tools (i.e., Delft3D-FM GUI and HEC-RAS) to construct forcing conditions combinational models. Such workflows often go with back-and-forth customization and tweaking, which is distinct from operational HPC computations.

We constructed all exploratory experiments on local machines equipped with 4-core processors and 8 to 16 Gigabyte RAM. A Delft3D-FM model typically runs 15 minutes, with 4 cores in use, and 70 minutes on the server. It has a single core (partitioning is not enabled by default) that is qualitatively proportional to both the core number and the processor frequency. It is important to note Delft3D-FM supports the aforementioned MPI and OpenMP, but otherwise requires partitioning on datasets such as 2D mesh. Alternatively, HEC-RAS runs on the GUI-based personal workstation, which is not parallelizable.

4.3.2 Scalability and I/O Throttle

From an operational computing perspective we are aware of distinctions between, and capabilities of, different platforms and their corresponding computational expenses related to runtime, disk space, and overhead when parallelizing. While still seeking to fully automate workflow and parallelize tasks, we found it manageable to reproduce all results with minor changes each time (e.g., simulation duration, different inputs). This will benefit 1) our current topic of exploring parameter sensitivity, and 2) future work of similar model setups.

The straightforward implementation is to parallelize tasks, known as ‘embarrassingly parallelized’, and it is operationally desirable to run multiple models without diving into the delicate partitioning of jobs. To alleviate the I/O throttle, rather than delicately schedule parallelized tasks we sought to perform in-memory compression and decompression, a newly emerging technology. To achieve equivalent high-throughput I/O we chose a state-of-the-art Lossy Compressor SZ [16], already used by many scientist communities to perform proof-of-concept experiments. In the future, this methodology will be incorporated into our workflow, with potential modification of existing compute frameworks. Preliminary results are discussed in Section 5.3

4.3.3 Inter-model comparison

Maitaria et al. [7] previously analyzed TWL prediction accuracy in Delaware Bay based on 1D/2D DFLOW-FM and 1D HEC-RAS models. To complement their findings, a 2D HEC-RAS model was set up in this study with the same characteristics and model parameters (e.g., model domain, grid size, topo/bathymetry, roughness coefficients, etc.) of 1D/2D DFLOW-FM.

5. Results

5.1. Investigation on Relevant Forcing Conditions

5.1.1 Contribution of different components on Total Water Level

This discussion is based on results generated in the experiments (see section 4.2.1); the astronomical tide (AT), river discharge-induced water level (RI) and storm surge-induced water level (SSWL). For Hurricane Isabel, the gap between observed TWL and simulated AT stayed within 0.5 m until an extreme water level event occurred around September 18-19, 2012 expanding the difference by 1-2m (Figure 2, left). Compared to the AT, RI was always positive but small, which indicates Delaware Bay is a tidal current-dominated estuary during mean discharge conditions [17,18]. However, an abrupt increase in RI was found at September 24, corresponding to the sudden increase in river discharge erroneously simulated by the model. SSWL, described as water level variability induced by storm surge signal and surge-tide interaction, was represented by subtracting the water level simulated in I3 from that simulated in I2. SSWL variability was less than 0.6 m (smaller than the AT) during most of the simulation period before reaching peak around September 19, when a 1.5 m water level rise at Brandywine to 2.0 m at Newbold occurred, (mainly due to the funneling effect). Additionally, peak SSWL was higher than the observed TWL peak, indicating the storm surge did not coincide with the high tide water.

Model results for Hurricane Sandy are shown in Figure 2 (right). Similarly, AT contributed most to observed TWL during normal conditions. After storm surge approached the study domain (around October 28), SSWL became the most important water level component. River discharge as revealed during Hurricane Isabel had minor influence on TWL near the mouth of the bay, but caused 0.1-0.5 m variability of water levels upstream.

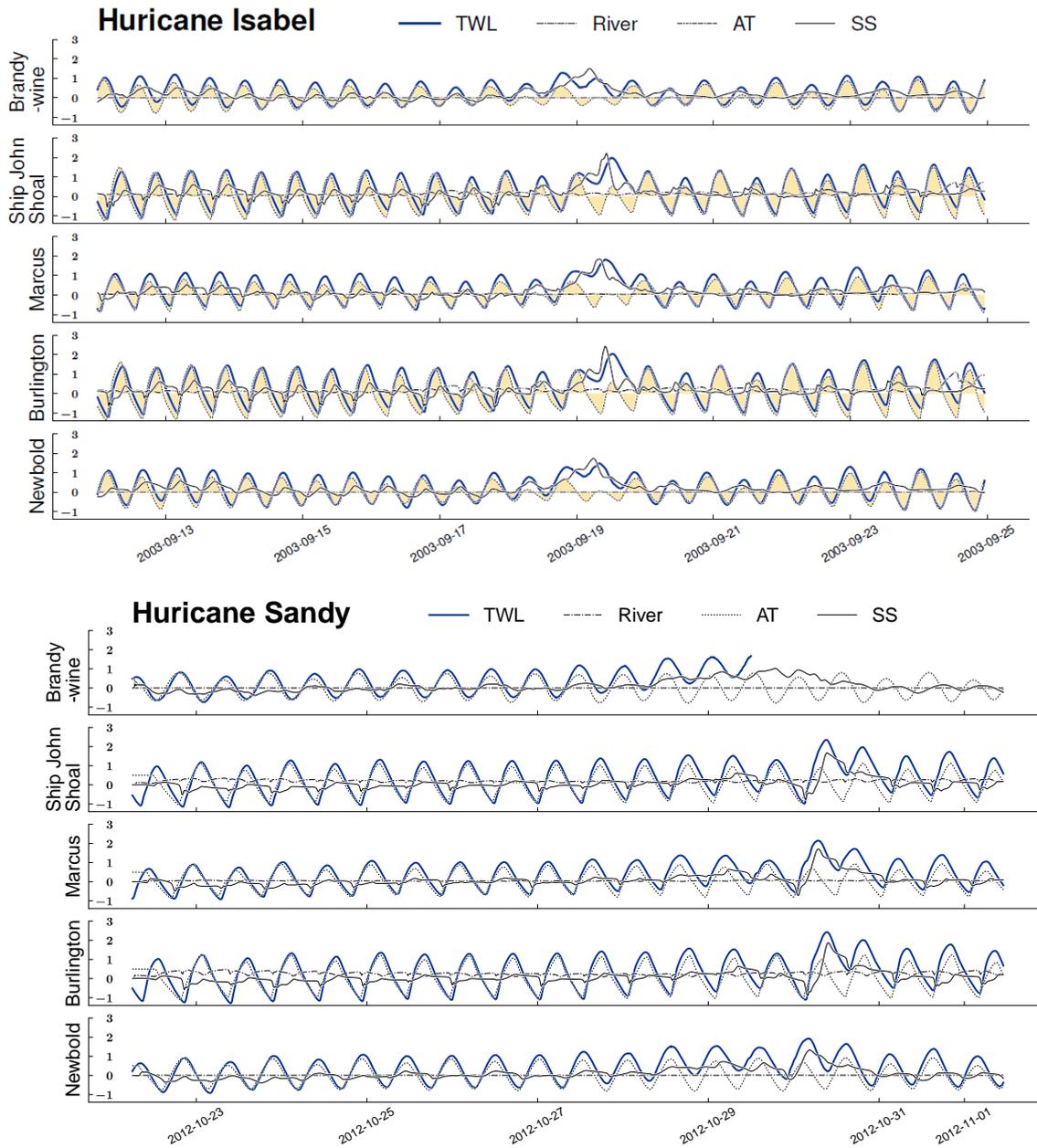


Figure 2. (Top) Hurricane Isabel and (Bottom) Hurricane Sandy. Simulated water level components time series (RI: dash-dotted line, AT: dotted line, SSWL: black solid line) and observed TWL (blue solid lines) at selected NOAA tidal/gauge stations. (top to bottom) Brandywine, Ship John Shoal, Marcus Hook, Burlington, and Newbold.

5.1.1a. Extreme water level condition

To further analyze the composition of TWL, two extreme water level events--observed peak water level and peak SSWL--were selected and all components extracted.

5.1.1b. Hurricane Isabel

For all selected gauges SSWL contributed most to TWL, with 59.12% on average, followed by AT with a 20.66% average, and finally RI with only 3% (Table 4; Figure 3, left). Following the longitudinal direction from Brandywine to Marcus, both the value of peak TWL peak and SSWL increased, while AT decreased. River discharge did not show noticeable variation. This indicates the increase at peak TWL was mainly due to the increase of SSWL, which resulted from the

funneling effect. From Marcus to Newbold, however, TWL still increased, despite decreased SSWL due to loss of energy as it propagated upstream. The TWL increase can be partly attributed to the RI increase but, more importantly, was due to the relatively higher AT with which the SSWL coincided. An additional conclusion is that during the storm surge-induced extreme water level event, the river discharge influence was small, regardless of location. Given this, river discharge consideration in the water level simulations did not influence the results; especially at gauges located downstream. Table 4 (SSWL Peak) and Figure 3 (right) show the relative components of water level at the peak event of storm surge-induced water level. At all selected gauges, SSWL was larger than observed TWL by approximately 0.5 m. Negative values of the AT indicate the storm surge peak coincided with the relatively low AT levels, resulting in the lower observed TWL compared to SSWL. Again, although the RI showed an increasing trend from downstream to upstream, its contribution was rather small.

Table 4. Water level components at selected gauge stations during peak event of non-SS (top) and SS- induced (bottom) water level for Hurricane Isabel (left) Sandy (right) respectively.

Peak Event	Station	Hurricane Isabel (2003)					Hurricane Sandy (2012)				
		Peak time	TWL [m]	Astronomical Tide [m]	Discharge [m]	Storm Surge [m]	Peak time	TWL [m]	Astronomical Tide [m]	Discharge [m]	Storm Surge [m]
TWL Peak	Brandywine	9/18 18:54	1.288	0.5732 (44.50)	0.0001 (0.01)	0.6175 (47.94)	N/A	N/A	N/A	N/A	N/A
	Ship John Shoal	9/19 7:48	1.487	0.0683 (4.59)	0.0052 (0.35)	1.1974 (80.52)	10/30 4:12	1.934	0.416 (21.52)	0.005 (0.25)	1.303 (67.39)
	Marcus	9/19 9:54	1.823	0.0346 (1.90)	0.047 (2.58)	1.2247 (67.18)	10/30 6:48	2.146	0.347 (16.15)	0.044 (2.06)	1.702 (79.32)
	Burlington	9/19 12:06	1.972	0.4926 (24.98)	0.1241 (6.29)	1.0881 (55.18)	10/30 9:12	2.354	0.474 (20.15)	0.147 (6.25)	1.652 (70.17)
	Newbold	9/19 12:30	2.038	0.5569 (27.33)	0.1173 (5.76)	0.913 (44.80)	10/30 9:18	2.428	0.533 (21.96)	0.191 (7.85)	1.760 (72.49)
SSWL Peak	Brandywine	9/19 4:30	0.794	-0.1518 (-19.12)	0.0012 (0.15)	1.5111 (190.31)	N/A	N/A	N/A	N/A	N/A
	Ship John Shoal	9/19 5:54	1.244	-0.093 (-7.48)	0.0026 (0.21)	1.7395 (139.83)	10/30 4:30	1.918	0.331 (17.26)	0.005 (0.28)	1.326 (69.15)
	Marcus	9/19 8:00	1.490	-0.2021 (-13.56)	0.0504 (3.38)	1.8494 (124.12)	10/30 6:54	2.146	0.320 (14.92)	0.049 (2.28)	1.708 (79.59)
	Burlington	9/19 10:24	1.589	-0.0471 (-2.96)	0.1103 (6.94)	2.2128 (139.26)	10/30 9:24	2.344	0.409 (17.45)	0.150 (6.41)	1.676 (71.51)
	Newbold	9/19 10:00	1.440	-0.2854 (-19.82)	0.1073 (7.45)	2.4299 (168.74)	10/30 9:42	2.406	0.345 (14.33)	0.235 (9.78)	1.874 (77.90)

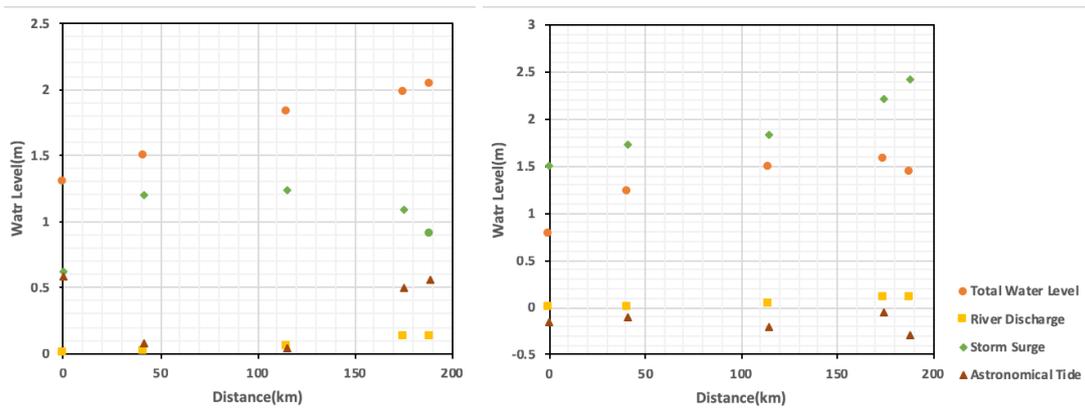


Figure 3. Water level components at selected NOAA gauges during: (left) TWL peak event, (right) SSWL peak event for Hurricane Isabel.

5.1.1c. Hurricane Sandy

During the TWL peak of Hurricane Sandy, SSWL contributed most to TWL with an average 72.34%, followed by AT with 19.95%, and RI with 4.1% (Table 4 and Figure 4, left). Similar to Hurricane Isabel, SSWL for Hurricane Sandy first increased along the longitudinal direction to the upstream, and then decreased due to energy losses while propagating.

During the peak SSWL event of Hurricane Sandy, as shown in Table 4 and Figure 4 (right), SSWL was smaller than the TWL at all selected gauges, indicating the storm surge signal coincided with relatively high astronomical tide (positive compared to the vertical datum), differing from the AT results for Hurricane Isabel. But similar to Hurricane Isabel, the river discharge contribution to TWL, although increased in upstream direction, was rather small compared to other components.

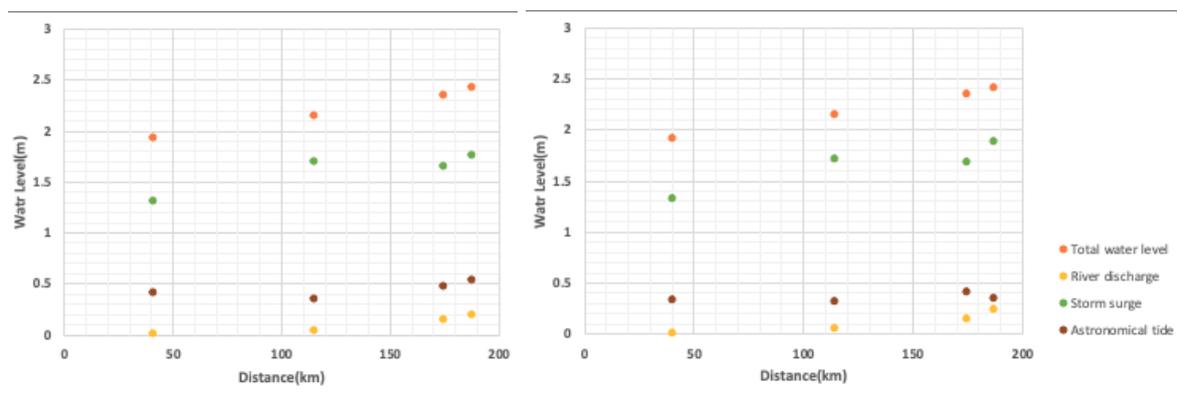


Figure 4. Water level components at selected NOAA gauges during **(left)** TWL peak event, and **(right)** SSWL peak event for Hurricane Sandy.

5.1.2 Spatial Variation of Total Water Level

The ‘trends’ of R^2 and Model Skill of Hurricane Isabel were similar to Hurricane Sandy (Figure 5). The highest correlations and skills came from the reference model (g10/g14), followed by scenarios simulated without river discharge (g9/g13), wind and atmospheric pressure (g2) and surge (g42/g46). The RMSE trends suggest the largest error deviations in TWL prediction occurred at distances greater than 100 km (Marcus, Burlington and Newbold gauge stations). The RMSE at those stations was above 0.20 m for both hurricanes, while the lowest RMSE came from simulations of the reference model (g10/g14), with values ranging from 0.11 to 0.32 m and 0.25 to 0.31 m, respectively. In concordance with R^2 and Model Skill, the RMSE gradually increased when using scenarios g9/g13, g2 and g42/g46. Bias analysis, however, leads to different results for Isabel and Sandy. When modeling Isabel, scenarios g10 and g2 overestimated TWL with maximum values of 0.12 m, while scenarios g9 and 42 underestimated TWL, with maximum negative values of 0.10 and 0.25 m, respectively. For Sandy, the four scenarios underestimated TWL with maximum negative values of 0.33 m. However, scenarios g2 and g14 overestimated TWL in the most upstream gauge stations, Burlington and Newbold, with maximum values of 0.12 and 0.07 m, respectively. Overall, the most favorable RMSE, BIAS, R^2 , and Model Skill were observed in scenarios that do not include surge forcing in simulations.

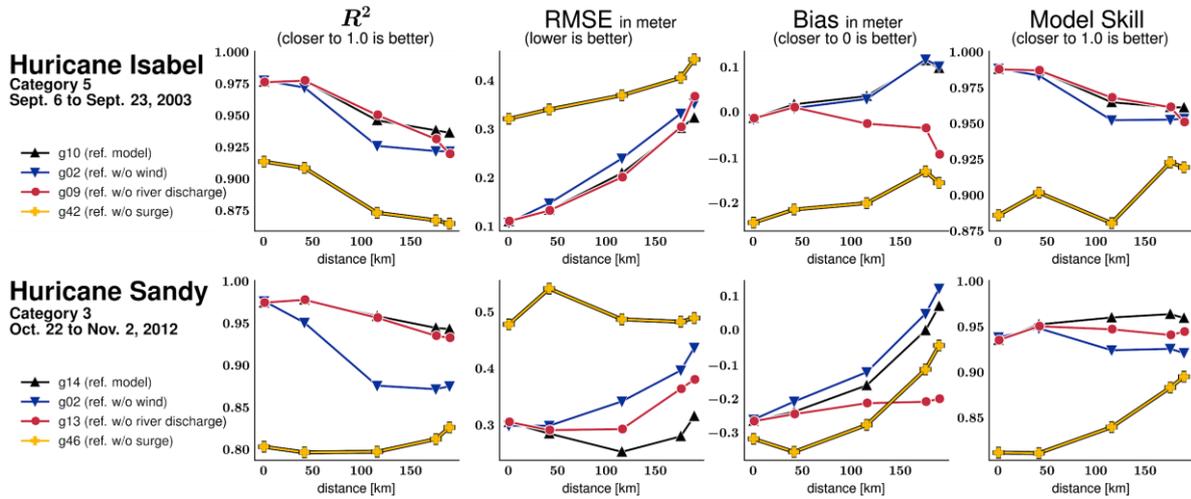


Figure 5. Spatial variation of R^2 , RMSE [m], Bias [m] and Model Skill calculated with observations and scenarios derived from the reference model.

The influence of Isabel’s wind was more evident in upstream parts of the Delaware River (figure 5), starting at Philadelphia (distances > 150 km) where the RMSE varies from 0.20 to 0.40 m (Figure 6, top). Conversely, Sandy’s wind showed a gradual increase of RMSE with low values at the estuary mouth (RMSE < 0.05 m), moderate values in the estuary from Brandywine station to Ship John station (RMSE up to 0.20 m) and high values along the river from Marcus station to Newbold station (RMSE up to 0.30 m). The influence of river discharge on RMSE was evident in the upstream parts of the Delaware River in both hurricanes. Moderate-to-high RMSE variation (0.20 to 0.40 m) was observed from Burlington station to Newbold station, while low RMSE (< 0.05 m) was observed from the estuary mouth to the Delaware River mouth (at Ship John station). As was expected, the scenarios that ignored surge forcing exhibit the highest RMSE, with values above 0.30 m over the entire model domain.

5.1.3 Total Water Level Temporal Variation

Wind influence was further investigated at three different instances around the TWL peak, as indicated in Figure 6 (bottom). Before Isabel’s TWL peak, an underestimation in TWL was evident at the estuary mouth, and from the river mouth up to Newbold station, with values ranging from -0.20 to -0.40 m. From Brandywine station to Ship John station (in the middle of the estuary), a TWL overestimation was observed, with values ranging from 0.20 to 0.40 m. At TWL peak, the TWL underestimation increased from Marcus station to Burlington station, with values ranging from -0.50 to -0.70 m. Similarly, a TWL overestimation was observed from the middle of the estuary to Delaware City, with values ranging from 0.20 to 0.40 m. After TWL peak, the TWL underestimation decreased from Marcus station to Burlington station, with values ranging from -0.30 to -0.50 m. Also, TWL overestimation decreased from the middle of the estuary to Delaware City with maximum values of 0.30 m.

The wind influence was more evident at the three instances for Sandy than for Isabel. For Sandy, there was a transition from TWL overestimation over the entire model domain, passing to both over- and underestimation at TWL peak and ending with underestimation after the peak was observed. Before TWL peaks, a relatively large TWL overestimation was evident along the river, with values ranging from 0.40 to 0.70 m. At TWL peak, the model domain was divided into three zones characterized by: 1) a TWL overestimation at the southwestern part of the estuary, with values ranging from 0.20 to 0.40 m; 2) a TWL underestimation in the middle-eastern part of the

estuary, continuing to Delaware city, with values ranging from -0.20 to -0.80 m, and 3) a TWL overestimation from Marcus station to Newbold station, with values ranging from 0.20 to 1 m. After the TWL peak, a relatively large TWL underestimation was evident along the Delaware River, with values ranging from -0.40 to 0.60 m.

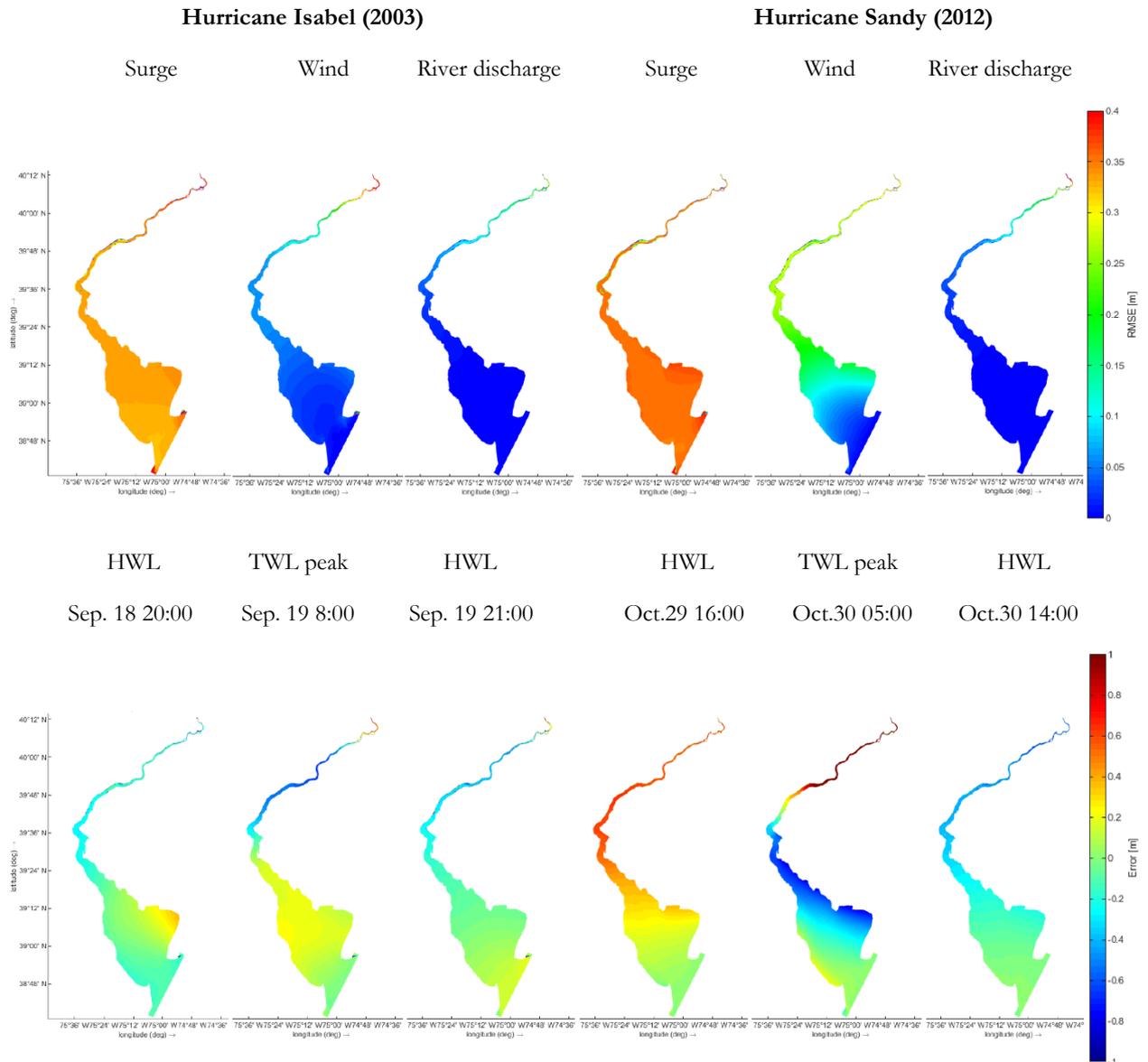


Figure 6. (Top) Spatial variation of TWL in terms of RMSE [m]. The spatial variation was calculated with the reference model and scenarios without wind, river discharge and surge, respectively. (Bottom) Temporal variation of TWL error [m]. The error was calculated as the difference between the reference model and scenarios without wind at three instances: high water level before peak, at peak water level, and high-water level after peak, respectively.

5.2 Computational Platform and Testbench

5.2.1 Task Parallelism Preliminary Result

To emulate the I/O throttling, we used Hurricane Isabel dataset [19], with its 12 subsets of 96 Megabytes (1,152 Megabytes in total) burst loaded into the memory. As shown in Table 5, parallelized I/O is saturated even with only four tasks issued simultaneously (average time to load does not change accordingly); it takes proportional time to input file sizes. As the model

size increases and the (task) parallelism grows, each issued task must wait for: 1) the initial and compulsory read and final write to permanent storage rather than memory, and 2) sporadic read and write that is specific to each model type.

Table 5. I/O throttle when parallelizing reading.

	par. 4	par. 8	par. 16	par. 32
parallel I/O	45.5	89.4	179.1	355.8
(average)	11.375	11.175	11.19	11.18

5.2.2 Inter-model comparison

Since HEC-RAS cannot simulate wind and pressure fields over the model domain, the reference model (g10) was modified to exclude those two fields (g2), hence allowing a model inter-comparison. When comparing g2 to 2D HEC-RAS in Hurricane Isabel, the four statistical measures present the largest differences at Newbold station, with maximum absolute values of 0.11, 0.25 m, 0.29 m and 0.13, respectively (Figure 7). For Hurricane Sandy, the largest differences in these measures also occur at Newbold station, with maximum absolute values of 0.05, 0.07 m, 0.02 m and 0.03, respectively. When comparing g2 to 1D HEC-RAS in Isabel, the largest differences of R^2 , RMSE and Model Skill also occur at Newbold station, with maximum absolute values of 0.29, 0.29 m and 0.57, respectively. In terms of Bias, however, the largest difference for Isabel occurs at Brandywine station, with maximum absolute value of 0.06 m. For Sandy, the largest differences of R^2 , RMSE and Model Skill again occur at Newbold station with maximum absolute values of 0.07, 0.13 m and 0.30, respectively. Similarly, the largest difference in Bias occurs at Brandywine station with maximum absolute value of 0.09 m.

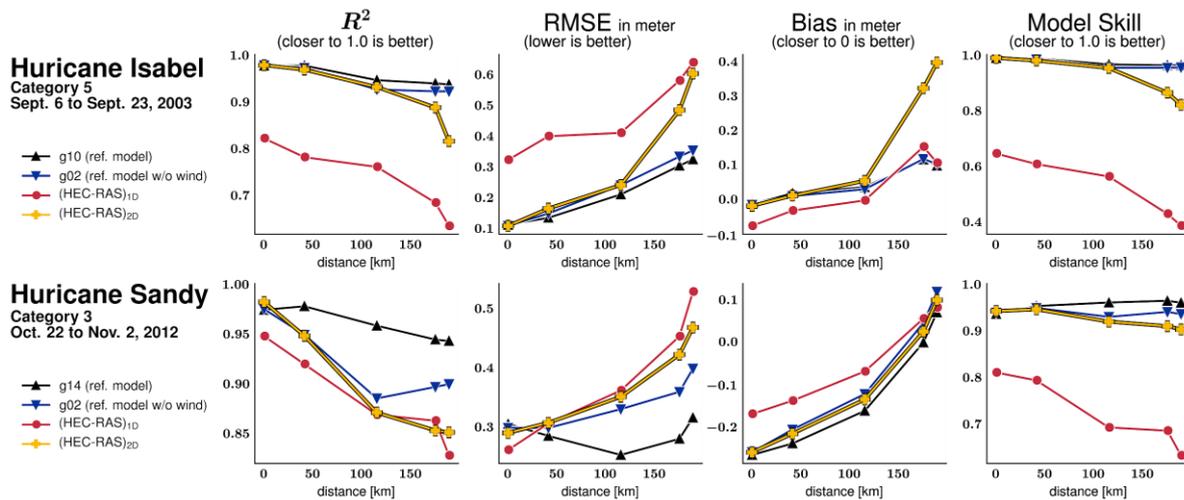


Figure 7. Inter-model comparison between Delft3D-FM and both 1D and 2D HEC-RAS models for Hurricane Isabel and Sandy. Accuracy in Total Water Level Prediction (between observations and simulations) is shown for 5 main stations. Inter-model comparison is presented in terms of R^2 , RMSE, Bias and Model Skill.

6. Conclusion

This study evaluated the contribution of three main components (tides, surge, and river discharge) to TWL, using a set of scenarios in a model framework with validated model input. Analyses of the three components and their respective contribution to TWL reveal surge-induced water level contributes most to peak water level, followed by the astronomical tide. River discharge, under mean flow conditions, had little influence on model results. The spatial and temporal TWL analysis indicated wind forcing plays a key role in TWL prediction, followed by river discharge. These results were more evident in Hurricane Sandy, where TWL underestimation and overestimation was observed in three instances of TWL peak. The wind module of Delft3D-FM significantly improved TWL prediction, particularly at peak water level in both hurricanes. Nevertheless, 2D HEC-RAS proved an simpler alternative for modeling storm-surge events when wind forcing is not relevant in the model domain.

Although the proposed model framework was not devised to fully represent water level induced by wave and current processes, future work might complement this study by coupling a wave model (e.g. SWAN and/or ADCIRC) with Delft3D-FM. Similarly, a rigorous analysis of storm surge-induced water level is suggested for future studies, to fully decompose the tide surge interaction.

1. Aerts, J.C.J.H.; Botzen, W.J.W.; Emanuel, K.; Lin, N.; Moel, H. de; Michel-Kerjan, E.O. Evaluating Flood Resilience Strategies for Coastal Megacities. *Science* **2014**, *344*, 473–475.
 2. Moftakhari, H.R.; AghaKouchak, A.; Sanders, B.F.; Feldman, D.L.; Sweet, W.; Matthew, R.A.; Luke, A. Increased nuisance flooding along the coasts of the United States due to sea level rise: Past and future. *Geophysical Research Letters* **2015**, *42*, 9846–9852.
 3. Milne, G.A.; Gehrels, W.R.; Hughes, C.W.; Tamisiea, M.E. Identifying the causes of sea-level change. *Nature Geoscience* **2009**, *2*, 471–478.
 4. Scavia, D.; Field, J.C.; Boesch, D.F.; Buddemeier, R.W.; Burkett, V.; Cayan, D.R.; Fogarty, M.; Harwell, M.A.; Howarth, R.W.; Mason, C.; et al. Climate change impacts on U.S. Coastal and Marine Ecosystems. *Estuaries* **2002**, *25*, 149–164.
 5. Titus, J.; Anderson, K. *Coastal Sensitivity to Sea-level Rise: A Focus on the Mid-Atlantic Region*; Government Printing Office, 2009; ISBN 978-0-16-083086-0.
 6. Muis, S.; Lin, N.; Verlaan, M.; Winsemius, H.C.; Ward, P.J.; Aerts, J.C.J.H. Spatiotemporal patterns of extreme sea levels along the western North-Atlantic coasts. *Scientific Reports* **2019**, *9*, 3391.
 7. Maitaria, K.; Bakhtyar, R.; Trimble, B.; Velissariou, P.; Flowers, T.; Mashriqui, H. 1D and 2D Hydrodynamic Modeling of Riverine-Estuary System under Extreme Storms: A Case Study of Delaware Bay/River Basin. *AGU Fall Meeting Abstracts* **2018**, *41*.
 8. Mashriqui, H.; Reed, S.; Aschwanden, C.; Seminar, O.H.D. Toward Modeling of river-estuary-ocean interactions to enhance operational river forecasting in the NOAA National Weather Service. In Proceedings of the Proceedings of the 2nd Joint Federal Interagency Conference, Las Vegas, NV, USA; 2010; Vol. 27.
 9. Serafin, K.A.; Ruggiero, P. Simulating extreme total water levels using a time-dependent, extreme value approach. *Journal of Geophysical Research: Oceans* **2014**, *119*, 6305–6329.
 10. Marsooli, R.; Lin, N. Numerical Modeling of Historical Storm Tides and Waves and Their Interactions Along the U.S. East and Gulf Coasts. *Journal of Geophysical Research: Oceans* **2018**, *123*, 3844–3874.
 11. Resio, D.T.; Westerink, J.J. Modeling the physics of storm surges. *Physics Today* **2008**, *61*, 33.
 12. Ferreira, C.M.; Irish, J.L.; Olivera, F. Uncertainty in hurricane surge simulation due to land cover specification. *Journal of Geophysical Research: Oceans* **2014**, *119*, 1812–1827.
 13. Garzon, J.L.; Ferreira, C.M. Storm Surge Modeling in Large Estuaries: Sensitivity Analyses to Parameters and Physical Processes in the Chesapeake Bay. *Journal of Marine Science and Engineering* **2016**, *4*, 45.
 14. OpenMP Architecture Review Board. Home Available online: <https://www.openmp.org/> (accessed on Jul 23, 2019).
 15. Open MPI Project Open MPI: Open Source High Performance Computing Available online: <https://www.open-mpi.org/> (accessed on Jul 23, 2019).
 16. Tao, D.; Di, S.; Chen, Z.; Cappello, F. Significantly Improving Lossy Compression for Scientific Data Sets Based on Multidimensional Prediction and Error-Controlled Quantization. In Proceedings of the 2017 IEEE International Parallel and Distributed Processing Symposium (IPDPS); 2017; pp. 1129–1139.
 17. Sharp, J.H.; Cifuentes, L.A.; Coffin, R.B.; Pennock, J.R.; Wong, K.-C. The influence of river variability on the circulation, chemistry, and microbiology of the Delaware Estuary. *Estuaries* **1986**, *9*, 261–269.
 18. Walters, R.A. A model study of tidal and residual flow in Delaware Bay and River. *Journal of Geophysical Research: Oceans* **1997**, *102*, 12689–12704.
 19. IEEE Visualization 2004 Contest: Data Set IEEE Visualization 2004 Contest: Data Set Available online: <http://vis.computer.org/vis2004contest/data.html> (accessed on Jul 23, 2019).
-

Chapter 2

Model-Based Parametric Analysis of Total Water Prediction in Coastal Transition Zones of the US East and Gulf Coasts

Taher Chegini¹, Gustavo Coelho², and John Ratcliff³

¹University of Houston; tchegini@uh.edu

²George Mason University; gcoelho@gmu.edu

³University of North Carolina at Chapel Hill; john.ratcliff@unc.edu

Summer Institute Theme Advisors: Celso Ferreira, George Mason University, cferrei3@gmu.edu; Kyle Mandli, Columbia University, kyle.mandli@columbia.edu; Patrick Burke, pat.burke@noaa.gov

Academic Advisors: Hong-yi Li, *University of Houston*; Celso Ferreira, *George Mason University*; Rick Luettich, *University of North Carolina at Chapel Hill*

Abstract: Vulnerability to extreme weather events and flood inundation motivates development of an operational total water prediction modeling solution in coastal transition zones. Coastal transition zones are defined in this study as the regions between the coast and upland reaches where tides no longer influence water level variation. Numerical modeling of these zones in operational settings is challenging, due to the complexity of coastal-inland dynamics and the computational expense of the effort. To provide insights for tackling issues associated with coupling coastal-inland models, this study investigated coastal transition zones by applying idealized domains. This approach facilitated in-depth analysis of relevant contributing phenomena through performing many simulations with different model input parameters, and then investigating output variability. To do this, three classes of idealized domains were generated, based on assessing the sizes and shapes of estuaries that have socioeconomic importance along the US East Coast and Gulf Coast. These domain classes include: a direct river-to-coast connection, a triangular-shaped bay, and a trapezoidal-shaped bay. Physical input parameter variations included river discharge, bed roughness, and tidal range. The extent of coastal influence in each domain was computed based on the reach of tidal signal propagation. Results demonstrated the damping effect of a bay on tidal signal propagation through the upland. In general, the presence of estuaries in a domain decreased the model's sensitivity to input parameters. Moreover, a non-linear relationship with variation in bed roughness and tidal signal was evident. The results quantified sensitivities of model outputs to studied input parameters and provide insights into dominant processes in coastal transition zones. These findings show computational efficiency of coupled coastal-inland modeling can be increased by focusing on the most influential parameters, and have an estimation of their contribution to output variability.

1. Motivation

More than 40 percent of the U.S. population lives in coastal shoreline counties, with U.S. coastal and ocean economies producing over \$8 trillion dollars in gross domestic product [1]. While these coastal areas continue to attract more residents and drive a significant portion of the nation's economy, they are extremely vulnerable to natural disasters. More than half the disaster costs incurred in the U.S. since 1980 can be attributed to damages resulting from tropical cyclones, with most damage resulting from storm surge and flooding effects [2]. Additionally, potential threats for compound flooding in coastal zones may be increasing, due to sea level rise

and growing probabilities for joint occurrences of heavy precipitation and storm surge events [3].

The initiative to enhance the National Water Model (NWM) to include total water prediction in coastal transition zones is underway, but coupling the NWM to existing coastal hydrodynamic models in operational settings presents challenges. In particular, the model resolution required to achieve accurate solutions in dynamic coastal areas incurs considerable computational cost to operate.

The initiative to couple the NWM with a coastal model involves extending coastal models further inland, to selected data exchange points with the NWM. This creates a larger coastal domain to resolve and even more computational expenses. With the goal of an operational system that can produce real-time forecasts, understanding physical processes and their contribution is essential to accurate and efficient modeling.

One approach for determining potential model simplification is to perform parametric analysis of model outputs. For this purpose, a set of parameters and geometries of interest that can potentially affect coastal area dynamics, were selected. Then, to form the parameter and geometry spaces, a range of possible values were estimated based on available data and information. Subsequently, a set of simulations were performed according to the determined parameter and geometry spaces. This process may produce insights regarding contributions of individual inputs toward model output variability, conclusions regarding input parameter significance, and guidance to apply toward subsequent studies [4].

2. Objectives and Scope

This study's objective was to quantify contributions of the most influential input model parameters for coastal transition zones hydrodynamic modeling. The results provide insights for improving the efficiency of coupled inland-coastal modeling in operational settings. The scope of this study was defined based on the following questions:

1. Can the tidal signal amplitude threshold be evaluated to enable identification of the reach of coastal zone influence?
2. Can investigation of relevant physical processes contributing to total water prediction provide guidance toward model calibration and efficiency?

Delineation of tidal signal thresholds could influence decisions on data exchange locations between inland and coastal models, an important consideration when optimizing the balance between accuracy and performance. In addition, insights into the respective significance of various physical parameters will help guide model calibration and efforts to achieve computational efficiency, by improving estimation of only relevant inputs or features.

Due to the broad scope of this study, use of complex hydrodynamic models such as D-Flow Flexible Mesh (D-Flow FM) module in Delft3D FM, in combination with complex geometries of actual coastal areas, was not feasible. Therefore, for this study a series of idealized domains were defined, based on classification of coastal areas along the US East and Gulf coasts. This classification, and all relevant physical input parameters and boundary conditions, were determined based on a preliminary study of US East and Gulf coasts regions, details that are discussed in Section 4. Three classes of idealized domains were defined: 1) direct connection of a river to a coast; 2) a triangular-shaped bay, and 3) a trapezoidal-shaped bay. Within these classes, three categories were defined to account for geometric variations that include adding sinuosity to the river, and adding a barrier to enclose the mouth of the bay. In addition to geometric variations, physical input parameters were varied: bed roughness, river discharge, and

water elevation at the ocean boundary. After running simulations, statistical analyses were performed on the results, based on water levels and velocities at two observation points and extracted tidal constituents from a cross-section along the middle of the computational domain.

3. Previous Studies

While many published works feature case studies that examine model sensitivities to input parameters, a literature survey showed few results for studies with a broad scope regarding various coastal areas, with similar objectives to this study. Akbar et al. [5] examined the effects of variations in bottom friction on water levels during a landfalling hurricane in the Gulf of Mexico. Due to energy dissipation at the seafloor, study results showed increases in water levels with decreases in Manning's roughness, and vice versa [6]. Bunya et al. [7] and Dietrich et al. [8] both showed the importance of spatially varied bottom friction on water level prediction accuracy. In the current study, bottom friction was constant over the whole domain.

Calero Quesada et al. [9] showed that river discharge can influence tidal propagation through estuaries [9]. Since part of our objective was to delineate a tidal signal threshold, it was important to examine tidal signal response to variations in discharge rates. Maskell et al. [10] pointed out simplification of domain geometries can allow for efficient computation, enabling many simulations to be performed while focusing on the significance of relevant physical processes [10]. Geometric variations may play a significant role in inundation processes, particularly with the interaction between river and storm surge components [10]. Alebregtse and de Swart [11] used a semi-analytical model to study the effect of river discharge on tidal wave propagation. They showed that when increasing river discharge, the amplitude of semi-diurnal sea surface elevation decreased, while velocity was not affected.

4. Methodology

The transition zones between coasts and rivers, where tides have significant influence on water level variation, were the regions of interest for this study. Initially, to see the extension of tidal environments, all active United States Geological Survey (USGS) water level stations over rivers and bays on the Gulf and East Coasts where tidal signals can be observed in the records, were identified (Figure 1a). With knowledge of coastal regions where tidal signals are present, the entire coast was inspected via satellite images to identify the most common types of geometries that could best represent the region (Figure 1b).

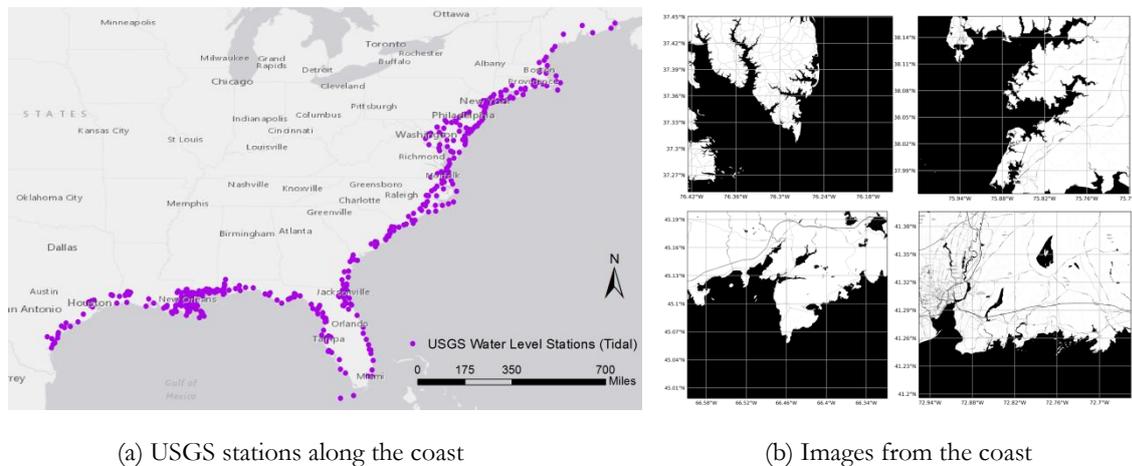


Figure 1. (a) USGS Water Level stations with tidal signal influence and (b) examples of satellite images extracted over the coastline

Regarding US coast heterogeneity, this study aimed to translate these complexities in a simple way, to produce results that can be extended to the entire country. 420 coastline images were gathered and analyzed. The analysis showed that most bays follow a triangular or trapezoidal shape. Thus, the coastal transition zones were categorized into three classes. (Figure 4a): direct river-to-coast connection (C1), triangular-shape bay (C2) and trapezoidal-shape bay (C3). Then, a sample of 46 bays and rivers was considered for parameterizing the geometries, based on four measurements: river width (W_r), upstream bay width (W_t), downstream bay width (W_b), and bay length (L_b). Three geometry ratios were defined as references to build a relationship with the magnitude of processes contributing to the water level and to be replicable in other domains (Figure 2). These ratios are as follows:

$$R_{lb} = \frac{L_b}{W_b}, R_{br} = \frac{W_b}{W_r}, R_{bt} = \frac{W_b}{W_t}$$

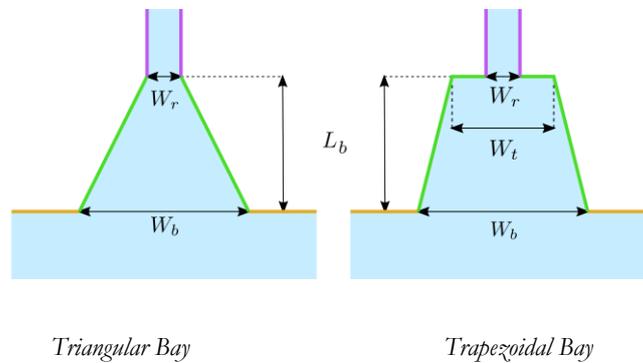


Figure 2. Geometry measurements used for classification and idealized domains design

Based on statistical analysis of the 46 sample bays and rivers the mean and median of R_{br} are 10.1 and 8.1 respectively, and uniformly distributed over the coast, with higher values appearing above a latitude of 37° (Figure 3 a). Moreover, the median of R_{lb} is 2.2 and higher values are observed for triangular shapes located between latitudes 32° and 38° (Figure 3 b).

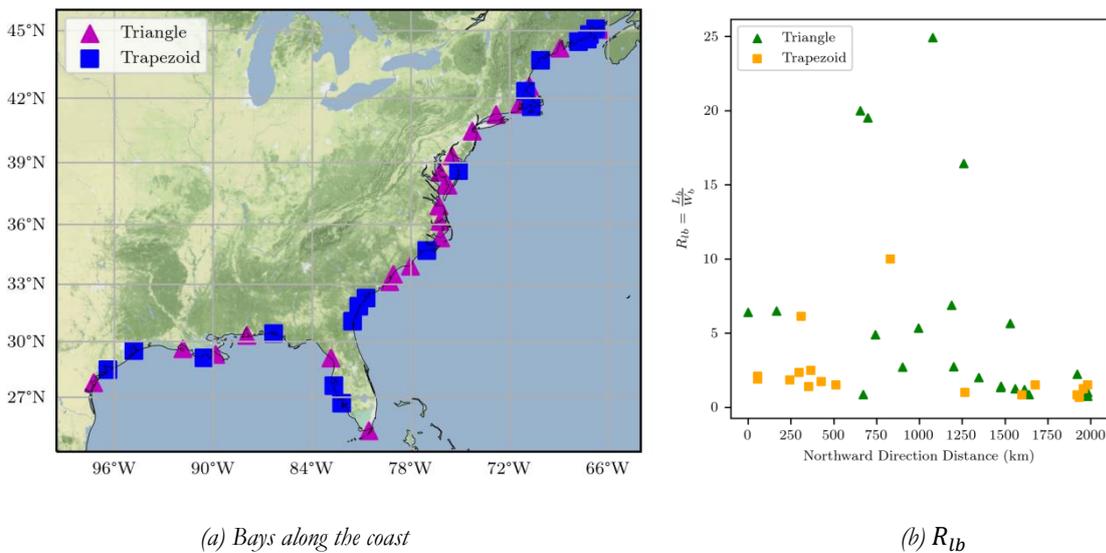


Figure 3. (a) The bays considered for classification R_{br} and (b) distribution of triangular and trapezoidal bays along the coast based on R_{lb}

Additionally, to account for geometric variations within the classes of idealized domains, three categories were defined: (A) straight line river; (B) sinuous river, and (C) bay with island barriers (Figure 4 a). Regarding forcings, the models were studied using variations of two main boundary conditions: river discharge at the upstream, and water level time series at the ocean boundary. The ocean region for all models has a 100 km length (perpendicular to the coast) and 50 km width. Bay width (W_b) was adopted as an independent parameter to scale the idealized models. For all configurations with bays (classes 2 and 3), W_b was set to 10 km, based on the sampling mean and median (Table 1). Additionally, within each class one reference scenario was defined for performing statistical analyses of the results. For the reference case, bay length (L_b) was set to 20 km based on the median value of R_{lb} , and river width (W_r) to 1 km, using the sampling mean of R_{br} . For category B, a river with sinuosity (SI) of 1.45 was studied. For category C, barrier openings that enclose the mouth of the bays were set to 2 km, equivalent to 20% of bay width (Figure 4 b).

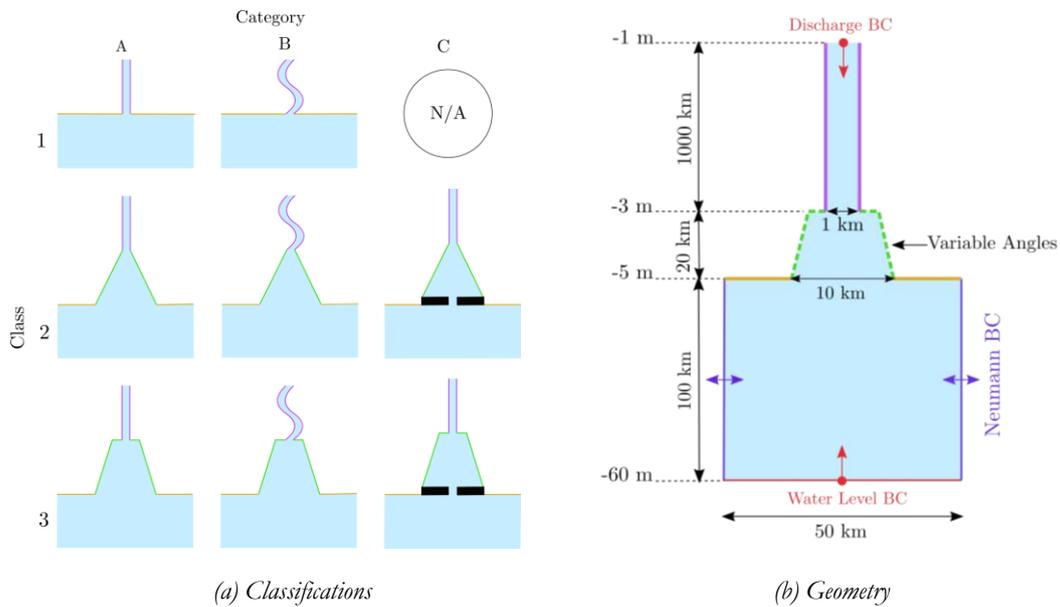


Figure 4. (a) Idealized domains classes and categories and (b) reference configuration with boundary conditions

Table 1. Summary of geometry measurements statistics

Statistic	W_b	R_{bt}	R_{br}	R_{lb}
Sample size	46	26	46	46
Mean	11.9 km	17.2	10.1	4.7
Standard deviation	11.3 km	23.2	14.	5.7
Minimum	0.4 km	0.23	0.75	0.67
Median	8.9 km	1.7	8.1	2.2
Maximum	49 km	79	90	25

A set of scenarios was chosen for the simulations and a procedure was established for each simulation, based on the scenarios. The procedure consisted of changing only one parameter per scenario, allowing quantification of input parameter contribution to output results. Simulation

scenarios are shown in Table 2. These were defined based on uniform Manning's roughness at the bed-level, discharge in the river upstream and water level time series in the ocean boundary (Figure 4 b). Two water level time series at the ocean boundary were used: predicted tides (P), and observed water level during a storm surge event (S).

Table 2. *Set of simulation scenarios*

Scenario	Roughness (R)	Discharge (D) [m ³ /s]	Tide (T)
Ref	0.025	0	P
R20	0.020	0	P
R30	0.030	0	P
S	0.025	0	S
D200	0.025	200	P
D500	0.025	500	P

Tides vary significantly over the study region, with lower amplitudes on the Gulf Coast than the East Coast. Selection of a location as a water level data source was based on NOAA station proximity to a significant number of samples of higher tide amplitude during an extreme event. This made it easier to identify changes among the classes, the categories and the simulation scenarios. Thus, at the ocean boundary water levels from NOAA station at Atlantic City, NJ (39°21.4' N, Lon = -74°25.1', ID = 8534720) during Hurricane Sandy (Oct 15 – Nov 10, 2012) were imposed.

Moreover, results analyses were performed based on velocity and water level data extracted from two observation points and a cross section along the domain's middle (Figure 5). To identify the limit of coastal zone influence in upland reaches, the tidal signal amplitude was extracted for the M₂ (principal lunar semidiurnal) tide constituent based on water levels along the cross section of each simulation. The tidal vanishing point was defined as the distance at which the rolling average of the M₂ amplitude over a window size of 20 km (5 observation points) reaches below 0.05 m.

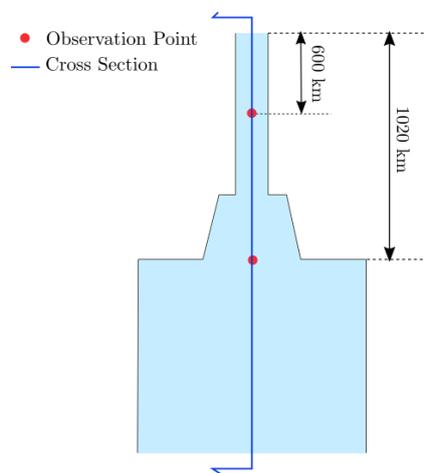


Figure 5. *The location of observation points and the cross section in the domain*

The process contributions were evaluated based on the Root Mean Squared Error (RMSE) as a measure of deviation from the reference case. RMSE were computed over the entire simulation period in the two observation points. All simulations were conducted using the hydrodynamic module D-Flow Flexible Mesh (D-Flow FM) of the Delft3D Flexible Mesh Suite developed by Deltare [12], with unstructured grids configured to perform 2D simulations.

5. Results

This section contains two subsections, to present study results according to each proposed research question area: (5.1) tidal signal amplitude in upland reaches and (5.2) relevant physical processes contributing to total water prediction.

5.1. Tidal signal amplitude in upland reaches

Analysis was performed on category A for all domain classes (1, 2 and 3) to find bias from corresponding reference simulations for propagation distance of the M_2 tidal signal upstream from the mouth of the river (Figure 6). The R30 scenario, with an increase in Manning's roughness, resulted in decreased M_2 signal propagation for all three classes, with a higher decrease for class 1 than for classes 2 and 3. Scenario R20, with a lower Manning's roughness, produced increases for all 3 classes although class 1 again appeared to show additional sensitivity, with a more substantial increase than classes 2 and 3. Scenario D200, featuring an upward adjustment of river discharge rate, created very slight changes across all classes. Increasing the river discharge for the D500 scenario produced a noticeable difference in reducing propagation for all three classes, with class 1 experiencing a slightly larger decrease. Scenario TSE, the simulation that includes storm surge, produced M_2 signal propagation distance reduction for class 1, with classes 2 and 3 showing minimal change.

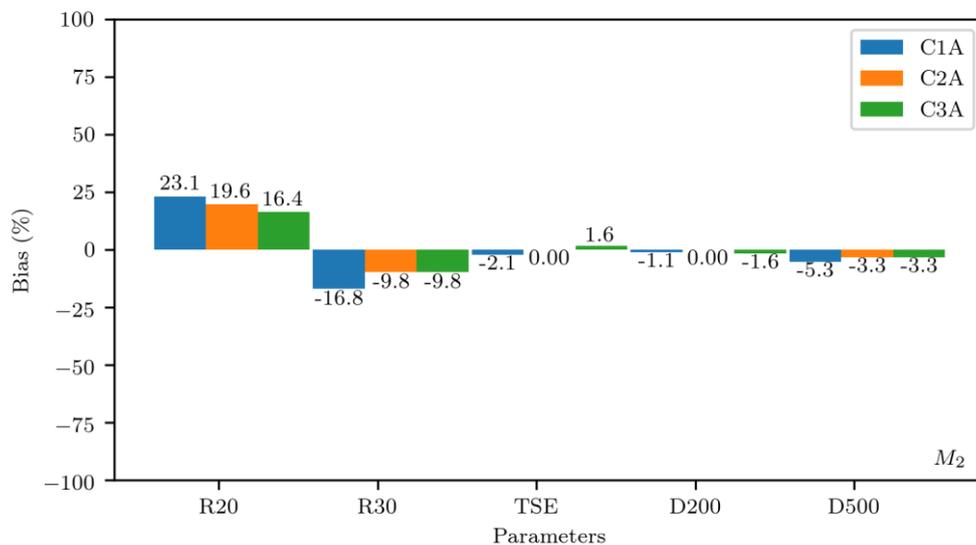


Figure 6. Tidal signal vanishing point percent bias when comparing variations in Manning's roughness, river discharge rate, and storm surge across geometric classes to respective reference simulations

The results showed that changes in selected input parameters may present indications of geometric influence on tidal signal propagation. Regarding results for using variations in Manning's roughness, one thing to note is the apparent increased sensitivity for class 1, which showed both a larger increase in signal distance with a lower Manning's roughness setting, and a larger decrease than classes 2 and 3 when applying a higher roughness setting. This may be explained as an attenuating role played by the bay, as the tidal signal damping was more noticeable in classes 2 and 3 than in class 1. A small difference in propagation was seen between

the class 2 and class 3, triangular versus trapezoidal bays, in the R20 scenario, possibly displaying a funneling effect with the convergence of the triangular bay. Length and depth of the bays could be important as part of this investigation, and adjustments to those domain dimensions could create different impacts in the results [13]. Also, it is important to note the non-linear relationship that can be seen when varying Manning's roughness. Increased sensitivity was apparent when adjusting roughness to lower levels, indicating underestimation of this input parameter may create larger output differences than overestimation.

Domain geometry also may be a factor in results of varying river discharge rate in the model. Class 1 again displayed sensitivity to variations in discharge, and was in line with previous research. This shows a relationship between tidal signal and discharge based on quadratic bed friction, so increasing discharge should produce less tidal signal propagation [14]. Classes 2 and 3 presented smaller changes with river discharge increases. A possible explanation for the difference between classes 2 and 3 and class 1 would again be the bay region damping effect.

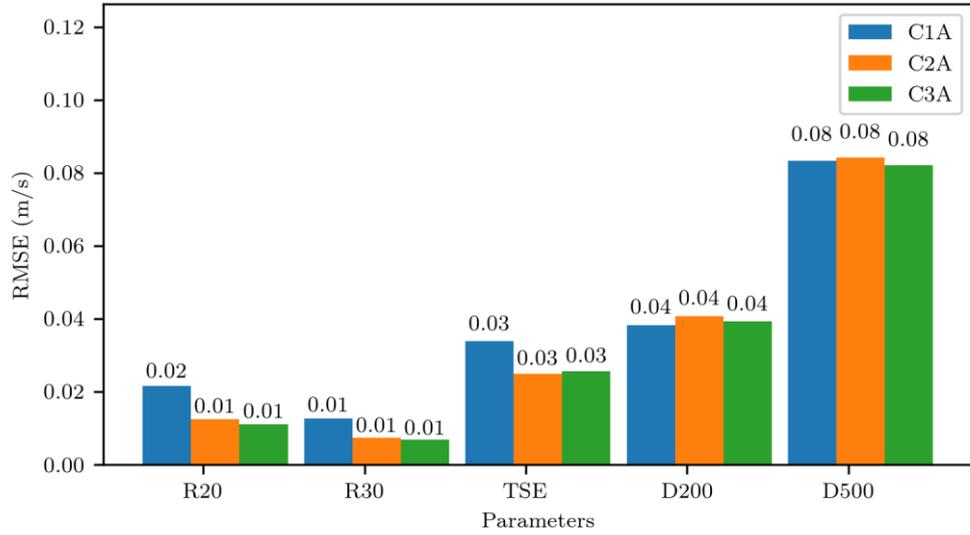
5.2. Relevant physical processes contributing to total water prediction

RMSE was used to quantify contributions of each selected input parameter on water level (Figure 7a) and velocity (Figure 7b) outputs at observation points 600 km downstream from the upper river boundary. No consequential water level changes or velocity outputs were seen at these river locations when bed roughness was increased or decreased. River discharge changes had a greater impact on both values of all three classes, as both the D200 and D500 scenarios showed sizable deviation from the reference scenario in water levels when compared to the reference simulations, and smaller errors with velocities.

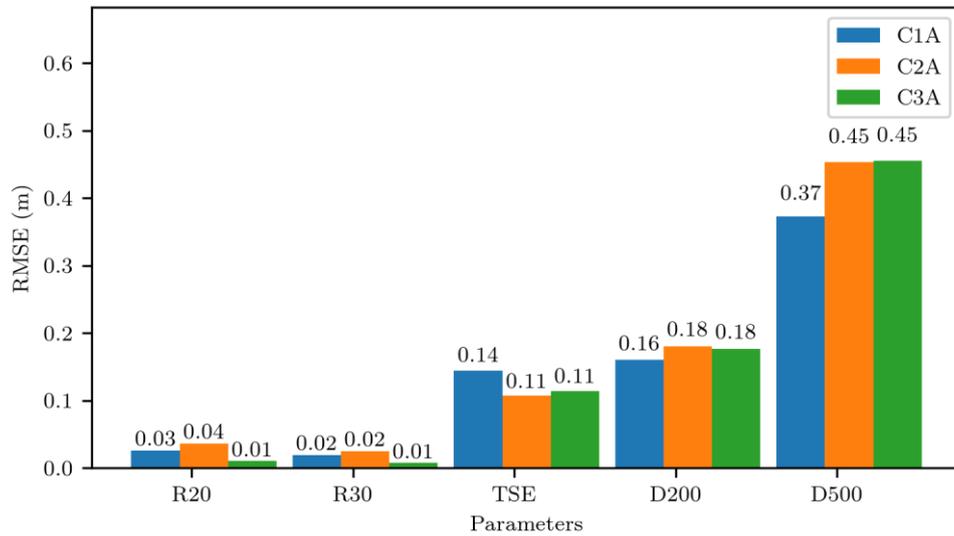
A potential reason for limited error for both the R20 and R30 scenarios is the distance upstream of observation points. The tidal signal may already have diminished enough that observation points were beyond the tidal influence limit, so little difference was detected for either velocity or water level from the reference. Results for water levels and velocities for D200 and D500 met expectations that increases would be seen in both values as discharge rates were raised.

The comparison of a straight river (Category A) to a sinuous river (Category B) did not indicate a notable water level change (less than 0.01 m) or velocity change (less than 0.01 m/s) RMSE for observation points 60 km upstream from the river mouth.

When analyzing island barrier impact (Category C), RMSE at the transition point between ocean and bay was 0.04 m in water level and 0.17 m/s in velocity for triangular bays (Class 2). For rectangular bay (Class 3), the RMSE is even higher, 0.08 m water level and 0.30 m/s in velocity. At an observation point in the river, 333 km upstream of its mouth, the bay shape had little effect on water level (< 0.01 m) and velocity (< 0.01 m/s) RMSE (Table 3).



(a) water level



(b) velocity

Figure 7. RMSE of (a) water level and (b) velocity at 600 km downstream of the river for different scenarios

Table 3. RMSE comparison between categories (reference scenario)

Class	C2A vs. C2C		C3A vs. C3C	
	Water Level	Velocity	Water Level	Velocity
River	< 0.1 m	< 0.01cm/s	< 0.01 m	< 0.01 m/s
Bay	0.04 m	0.17 m/s	0.08 m	0.30 m/s

The insights obtained from all the analyses were summarized into a parameter space (Table 4) and a geometry space (Table 5). The parameter space demonstrates parameter

contribution for each class. Based on parameter space, discharge contribution to water level is approximately 60% for all classes. Thus, a more accurate estimation of discharge at the upstream of a computational domain can have considerable impact on the accuracy of water level prediction. However, the geometry space shows triangular-shaped bays are more sensitive to roughness while direct river-coast connection geometry is more sensitive to discharge. Also, all three geometry classes are affected by the storm surge almost equally.

Table 4. *Weighted percentage of the parameter space for water level in class*

Class	Roughness	Surge	Discharge	Sum
C1 (River)	5	35	60	100
C2 (Triangular)	6	34	60	100
C3 (Trapezoidal)	2	36	62	100

Table 5. *Weighted percentage of the geometry space for water level for each parameter*

Parameter	C1 (River)	C2 (Triangular)	C3 (Trapezoidal)	Sum
Roughness	39	45	16	100
Surge	30	35	35	100
Discharge	38	31	31	100

6. Conclusion

The objective of this work was to quantify the contribution of input parameters in hydrodynamic modeling of coastal transition zones. Three main findings will contribute to the improvement of total water prediction in the NWM:

1. A method to classify coastal transition zones was proposed, that facilitates correlation of results found in idealized models to real world scenarios. This method introduces one independent variable, W_b , and three dependent variables for generating an idealized geometry. Based on our preliminary study of the US East and Gulf coasts, a range of values were defined for dependent variables. Only a subset of these ranges was considered in this study to keep the work scope manageable, yet informative. Further investigation on unexplored ranges would help expand the applicability of the results to real world scenarios.
2. The tidal signal influence threshold plays a significant role in defining the model upland domain. An accurate estimation of this point can help reduce computation cost and increase accuracy. The vanishing point of a tidal signal was determined based on the principal lunar semidiurnal tidal constituent, M_2 as it is the strongest constituent of semidiurnal regimes. Our investigation revealed Manning's roughness is the most significant parameter for determining the vanishing point. Moreover, the results showed that when Manning's

roughness is underestimated, change in the tidal influence limit is larger than when Manning's roughness is overestimated.

3. Investigation of the different domain configurations and scenarios demonstrated discharge is an important contributor in upland reaches, but not relevant in the bay. The sinuous river analysis did not show changes when compared to the straight river, but further investigation is required, as the sinuous river length was not large enough in our study to allow full investigation parameters. Additionally, it was shown that island barriers enclosing bays influenced water level and velocity over the bays, but not over the rivers.

Insights provided in this work would potentially assist coastal modelers with directing available resources toward the most significant parameters. The guidelines introduced quantify the uncertainties of each parameter and can be used to estimate the total uncertainty of predictions under various scenarios.

Supplementary Materials: The resources developed during this work are available on the GitHub repository: https://github.com/taataam/SI_2019_Coastal.

References

1. NOAA Office for Coastal Management, Digital Coast, Socioeconomic Data Summary (2017). Retrieved July 10, 2019 from <https://coast.noaa.gov/data/digitalcoast/pdf/socioeconomic-data-summary.pdf>
2. NOAA National Centers for Environmental Information (NCEI) U.S. Billion-Dollar Weather and Climate Disasters (2019). Retrieved July 10, 2019 from <https://www.ncdc.noaa.gov/billions/>
3. Chan, K., Saltelli, A., & Tarantola, S. (1997, January). Sensitivity analysis of model output: variance-based methods make the difference. In Winter Simulation Conference (pp. 261-268).
4. Wahl, T., Jain, S., Bender, J., Meyers, S. D., & Luther, M. E. (2015). Increasing risk of compound flooding from storm surge and rainfall for major US cities. *Nature Climate Change*, 5(12), 1093.
5. Akbar, M., Kanjanda, S., & Musinguzi, A. (2017). Effect of bottom friction, wind drag coefficient, and meteorological forcing in hindcast of hurricane rita storm surge using SWAN+ ADCIRC model. *Journal of Marine Science and Engineering*, 5(3), 38.
6. Kerr, P. C., Donahue, A. S., Westerink, J. J., Luettich Jr, R. A., Zheng, L. Y., Weisberg, R. H., ... & Roland, A. (2013). US IOOS coastal and ocean modeling testbed: Inter-model evaluation of tides, waves, and hurricane surge in the Gulf of Mexico. *Journal of Geophysical Research: Oceans*, 118(10), 5129-5172.
7. Bunya, S., Dietrich, J. C., Westerink, J. J., Ebersole, B. A., Smith, J. M., Atkinson, J. H., ... & Cardone, V. J. (2010). A high-resolution coupled riverine flow, tide, wind, wind wave, and storm surge model for southern Louisiana and Mississippi. Part I: Model development and validation. *Monthly weather review*, 138(2), 345-377.
8. Dietrich, J. C., Westerink, J. J., Kennedy, A. B., Smith, J. M., Jensen, R. E., Zijlema, M., ... & Cardone, V. J. (2011). Hurricane Gustav (2008) waves and storm surge: hindcast, synoptic analysis, and validation in Southern Louisiana. *Monthly Weather Review*, 139(8), 2488-2522.
9. Quesada, M. C. C., García-Lafuente, J., Garell, E., Cabello, J. D., Martins, F., & Navas, J. M. (2019). Effects of tidal and river discharge forcings on tidal propagation along the Guadiana Estuary. *Journal of sea research*.
10. Maskell, J., Horsburgh, K., Lewis, M., & Bates, P. (2013). Investigating river–surge interaction in idealised estuaries. *Journal of Coastal Research*, 30(2), 248-259.

11. Alebregtse, N. C., & de Swart, H. E. (2016). Effect of river discharge and geometry on tides and net water transport in an estuarine network, an idealized model applied to the Yangtze Estuary. *Continental Shelf Research*, 123, 29–49. doi:10.1016/j.csr.2016.03.028
12. Deltares (2019). D-Flow Flexible Mesh. D-Flow FM in Delta Shell. User Manual. Version 1.5.0, January 18, 2019.
13. van Rijn, L. C. (2011). Analytical and numerical analysis of tides and salinities in estuaries; part I: tidal wave propagation in convergent estuaries. *Ocean dynamics*, 61(11), 1719-1741.
14. Moftakhari, H. R., Jay, D. A., Talke, S. A., Kukulka, T., & Bromirski, P. D. (2013). A novel approach to flow estimation in tidal rivers. *Water Resources Research*, 49(8), 4817-4832.

Chapter 3

Experimental One-way Coupling of TOPMODEL with NWM to Substitute Runoff Processes Delineation in a Headwater Catchment

DongHyun Kim¹, Amina Naliaka², and Zhipeng Zhu³

¹San Diego State University/UC, Santa Barbara; dkim8398@sdsu.edu/donghyunkim@ucsb.edu

²Southern Illinois University; amina.naliaka@siu.edu

³University of Cincinnati; zhu4@mail.uc.edu

Summer Institute Theme Advisors: Hilary McMillan, San Diego State University, hmcmillan@sdsu.edu; Fred Ogden, National Water Center, fred.ogden@noaa.gov

Academic Advisors: Hilary McMillan, *San Diego State University*; Ruopu Li, *Southern Illinois University*; Patrick Ray, *University of Cincinnati*

Abstract: In the continental US, 53% of streams [38] can be classified as headwaters. The National Water Model (NWM) applies a uniform model structure for hydrological processes and spatial scale over the whole domain. But it may not always be accurate and time-efficient in all watersheds, and can cause higher uncertainty in headwater or small catchments. This work aimed to seek a balance between computational efficiency and hydrologic fidelity in headwater catchments by substituting different conceptualizations of subsurface processes in the NWM. For this, TOPMODEL, a simple physically-based model that incorporates topographical details, was loosely one-way coupled with NWM's land surface model (LSM; Noah-MP). NWM options and parameters were modified to produce output that is usable by TOPMODEL, with minimal post-processing. To evaluate model performance, streamflow simulated by the NWM and TOPMODEL-coupled version were compared to observed streamflow at Elder Creek United States Geological Survey (USGS) gage (11475560) in Eel River Critical Zone Observatory. The study area receives very little snow and has deeply weathered soil and bedrock that makes subsurface flows a dominant runoff. TOPMODEL-coupled version (PBIAS 1.67, NSE 0.75, RSR 0.50) outperformed the NWM (PBIAS 16.807, NSE 0.31, RSR 0.83) and was a better predictor of flood peak and water availability in the study area. Running the modified NWM's LSM did not decrease computation time of the coupled model system. However, a significant benefit of the coupled model was that rapid parameter optimization of TOPMODEL could be performed to calibrate the model without re-running the LSM. The result of the experimental coupling of TOPMODEL with NWM's LSM indicated this type of partial simplification is achievable, and can improve streamflow prediction. It also supports the idea of an adaptable hydrological modeling framework for both spatial and processes representation, to fulfill needs for varying environment and region.

1. Motivation

The National Water Model (NWM) is a modeling framework that aims to provide accurate streamflow forecasts over the continental United States. It provides complementary hydrologic guidance on National Weather Service (NWS) river forecast locations, and covers under-served locations[1]. The current NWM version is based on Weather Research and Forecasting Hydrologic model (WRF-Hydro) v5.0, developed and supported by the National Center for Atmospheric Research (NCAR). WRF-Hydro is coupled to the NOAH-MP land surface model [39] so the NWM configuration represents surficial, terrain routing and channel routing, and also

subsurface processes that involve several physics options based on grid-cell spatial representation. As it covers an expansive spatial area, the NWM requires considerable computation to run in a large domain [2], [3].

At least 53% [38] of reaches and watershed areas in the US can be classified as headwater, with 34% of stream length showing moderate-high (2 - 4%) gradient [4]. The NWM framework aims to provide comprehensive coverage, up to and including smaller watersheds and reaches. However, because the NWM applies uniform model structure over an entire domain, with resolutions that may ignore heterogeneities in parameters and process representation, uncertainties may occur in headwater catchments. The spatial resolution of the terrain routing grid (250m) is coarse enough to disregard detailed topographical influence on hydrological processes that may be significant in smaller catchments [5], [6]. The NWM core, WRF-Hydro, only requires setting the resolution of terrain routing grid as an integer divisor of Land Surface Model (LSM) grid. Therefore, applying a higher resolution routing grid for headwater catchments may solve the problem. However, that will result in tremendous computational burden when done in the whole CONUS.

Meanwhile, subsurface NWM processes are mostly decided by implicit parameter, not by watershed topography. Subsurface lateral flow routing in shallow soil considers topography and saturation of soil layers in the LSM; however, the other processes are determined by parameters that require calibration. GWBUCKPARM file controls groundwater reservoir storage and baseflow rate. It also controls slope index parameter, a linear reservoir coefficient, and controls underground runoff from bottom soil toward groundwater reservoir, affecting subsurface lateral flows [2], [7].

We suspected the NWM's current structural and spatial representation may not be ideal for representing hydrological processes in headwater or small catchments. In short, its approach may be neither time-efficient nor accurate. This awareness evolved into inspiration to alter the spatial representation and subsurface-flow processes of the current NWM.

2. Objectives and Scope

This work had two objectives. The first was to test different conceptualizations of subsurface processes and spatial representation incorporating details from topographical data. The second was to explore the possibility of achieving a balance between computational efficiency and hydrologic fidelity in nation-wide modeling, by coupling an existing model component with the NWM. For this, we examined the applicability of coupling, and compared streamflow prediction performance of the original NWM to a one-way coupled version in a headwater catchment.

The criteria applied to select a coupling model included: 1) topography oriented; 2) structurally simple; 3) physically reasonable subsurface processes representation and dynamic soil moisture contents consideration, and 4) low computational demand. We selected the topography-based hydrological model (TOPMODEL) for the loose one-way coupling with the NWM. TOPMODEL proposes a combined groundwater pathway that respects topography and allows soil porosity to change with depth, and has been successfully predicting flows in numerous watersheds. The fact that TOPMODEL is often held as a benchmark for expanding, or newly formulating delineation of the relationship between the water table and runoff processes, was also considered [8], [9]. The aim of watershed hydrology is to understand and model emergent watershed behavior as some hydrologic responses are more apparent at catchment scale. Given this, topography-based TOPMODEL may also provide some advantage over the grid-based NWM in utilizing domain knowledge about watersheds.

3. Previous Studies

3.1 TOPMODEL concept

TOPMODEL, initially developed by Beven and Kirkby in 1979 [10], is considered a set of conceptual tools for modeling hydrological processes, specifically surface and subsurface contributing area dynamics, in a relatively simple way [11] (Figure 1.). It has evolved to include a simple theory of hydrological similarity of points in a catchment, determined by deriving a Topographic Wetness Index (TWI) from a Digital Elevation Model (DEM) [12]. The TWI divides a catchment into classes of hydrological similarity and thus simplifies TOPMODEL by reducing data and the number of model parameters required in simulations [13], [14]. The simplified structure of TOPMODEL makes it a combination of computational and parametric efficiency, with clear foundations in physics [15][16].

TOPMODEL's ability to aptly represent impacts of terrain on hydrological processes (soil moisture and groundwater movement) using digital terrain model data, has led to its popularity in application to different LSMs. It has been successfully integrated with the Simplified Simple Biosphere Model (SSiB) [17], the soil-vegetation-atmosphere transfer (SVAT) model, and the Interactions between Soil, Biosphere, and Atmosphere (ISBA) model [18-20]. These integrations have shown similar conclusions: that integrating models improved discharge prediction and reduced the number of parameters requiring calibration.

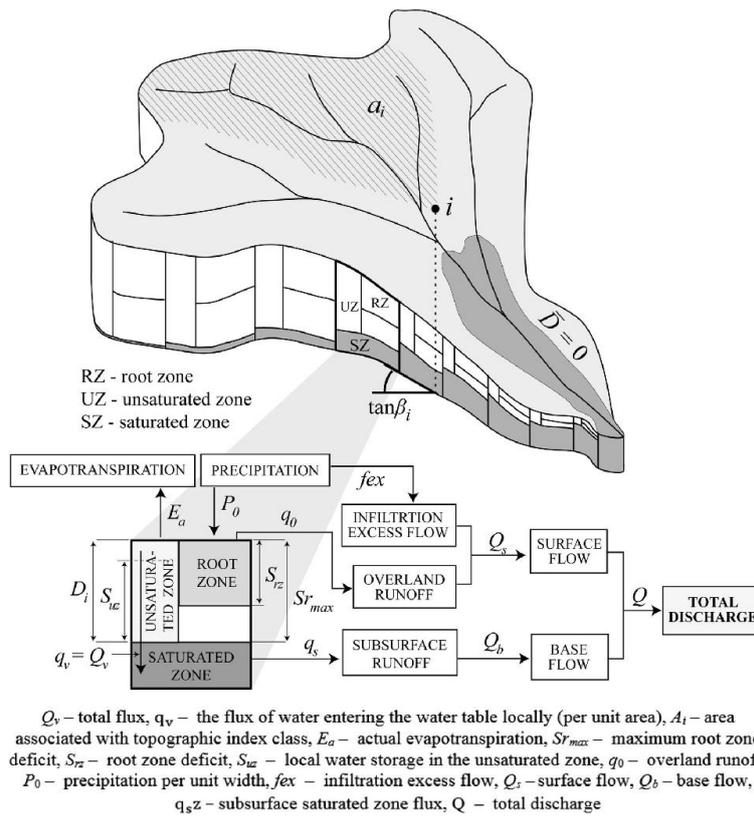


Figure 1. TOPMODEL concept [21]

3.2 Subsurface routing and underground runoff in WRF-Hydro

The NWM system core is WRF-Hydro coupled with the Noah-MP LSM to simulate land surface, terrain and channel routing, and subsurface processes [1]. WRF-Hydro solves the Boussinesq equation for saturated subsurface lateral flow, adding exfiltration from fully saturated cells to infiltration from the LSM [7]. Subsurface processes are simplified through the following assumptions; 1) Soil layers have a uniform depth of two meters, and 2) Water table depth is determined in Noah-MP according to saturation of the soil layer nearest the surface. Additionally, baseflow is depicted using a conceptual nonlinear reservoir (referred to as an exponential bucket model in the WRF-Hydro documentation) for each catchment [2], [22]. The lateral flow/subsurface routing in the saturated zone of the 2-m soil column and a groundwater bucket model can be evaluated for its contribution in base flow to river discharge, stored in output files such as CHRTOUT (Streamflow on the 2D high resolution routing grid), RTOUT (Terrain routing variables on the 2D, 250m routing grid), and GWOUT (Groundwater output variables) [23]. The relationship between soil layer subsurface flow and underground runoff, to the groundwater bucket reservoir, is determined by a parameter called SLOPE or slope index, a linear reservoir coefficient.

4. Methodology

4.1 Workflow overview

The workflow in Figure 2 (below), examines applicability of loose one-way coupling, and compares performance of the NWM v1.2 to the NWM LSM (Noah-MP)+TOPMODEL version in a headwater catchment. Spin up and model run for the NWM was done with its original configuration, with no alteration or calibration. Then, a modified NWM was run to produce LSM output that can be reformatted as hourly water flux input for TOPMODEL. Topographical data was used to calculate indices, and was converted to distribution function as required by TOPMODEL. TOPMODEL ran with numerous random sampled parameter-sets; an optimized parameter-set was chosen to compare the model performance with the NWM. In short, TOPMODEL replaced the NWM's runoff calculation and flow-routing schemes, using output data from Noah-MP (LSM). Then performance of the original NWM and TOPMODEL coupled version were compared.

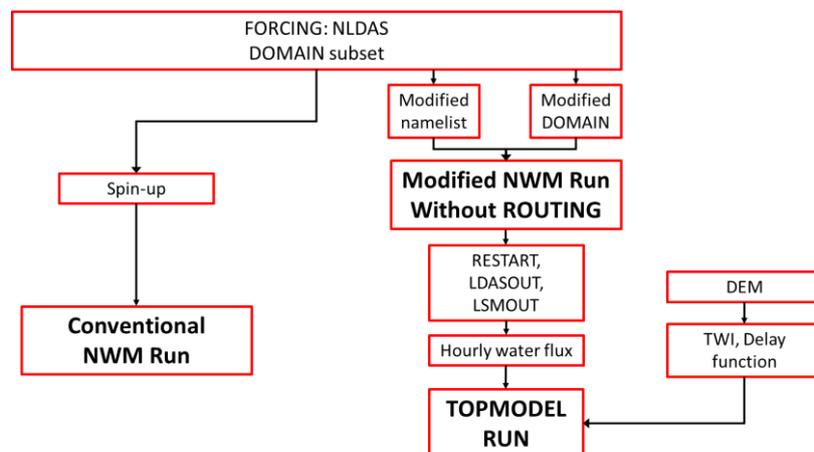


Figure 2. Workflow overview for one-way coupling of TOPMODEL and NWM's LSM

4.2 Land Surface Model output from NWM configuration

The National Water Model v1.2.2 [1] was used for this study, with the study area domain file subset using CUAHSI Domain Subsetter (<http://subset.cuahsi.org/>) [24]. The domain covers Angelo Coast Range Reserve, or 'Elder Creek-South Fork Eel River' watershed area that falls in HUC12 (hydrologic unit code) 180101060103. The North American Land Data Assimilation System-2 (NLDAS-2) from the National Aeronautics and Space Administration (NASA) Earthdata repository (<https://earthdata.nasa.gov/>) was used for meteorological forcing input. Model spin-up was done with repetitive model runs, using the first six weeks of the simulation period to emulate watershed wetting from dry condition (09/29/2017 - 11/09/2017). The actual NWM model run was done from October 1st, 2017 to April 30th, 2018 (Figure 3)

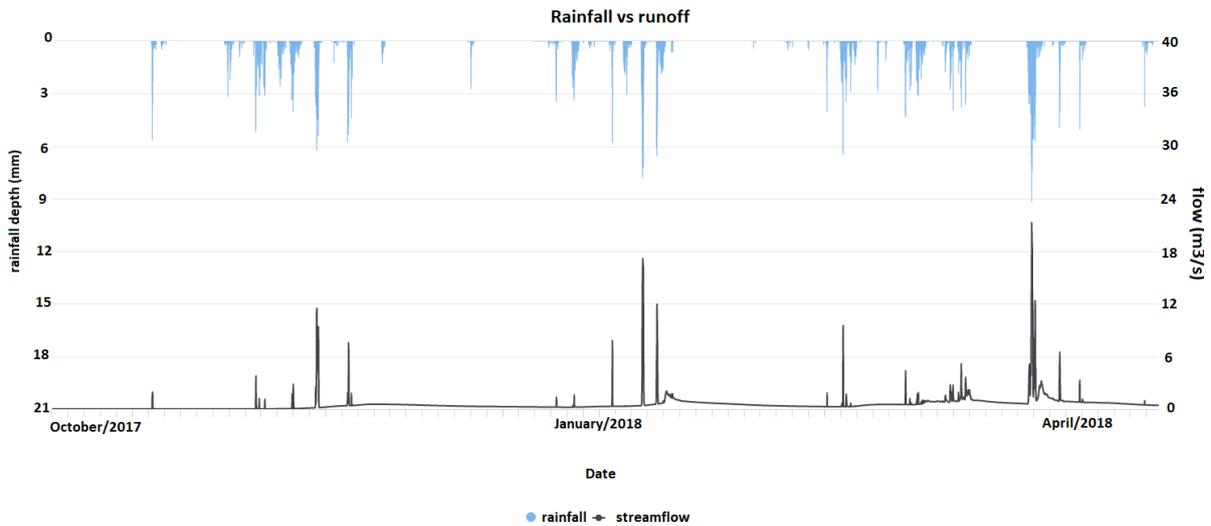


Figure 3. Rainfall and observed streamflow at USGS gage 11475560 "Elder Creek."

Simulation of the original NWM was done with no parameter or run option alteration, to examine its current performance on headwater catchments. However, to produce LSM output to feed TOPMODEL some soil/groundwater-related parameter and model run options were adjusted. Surface and subsurface flow routing were deactivated to prevent LSM output from being disturbed by the routing scheme. The slope index parameter in the 'soil_properties.nc' file was changed to 1, to prevent soil being overly saturated due to deactivated subsurface lateral flow. Doing this allowed free soil water gravitational drain in the bottom boundary toward the aquifer. Although the routing module was deactivated, this did not significantly lessen time spent running the model. We assumed this was because of modified NWM's setting to write RESTART file every hour. This was inevitable, as RESTART output file includes essential variables related to the internal state of the LSM.

LSM output from the modified NWM was reformatted for TOPMODEL. Target variables that were initially scattered through multiple output files, such as LDASOUT (Land surface model output), LSMOUT (Land surface diagnostic output) and RESTART (Land surface model internal state output), were extracted and calculated into basin averaged. Then, any accumulated values were disaggregated and calculated into hourly water flux between land surface and atmosphere. From LSM output variables, the water flux equation established the following:

$$(PRCP - \Delta CANWAT - \Delta SNEQV) - ET = (RUNOFF + \Delta SMC) \quad (\text{Equation 1})$$

PRCP is precipitation(mm); *CANWAT* is canopy water storage(mm); *SNEQV* is snow water storage(kg/m²); *ET* is calculated total evapotranspiration(mm); *RUNOFF* includes surface runoff and underground runoff to groundwater bucket, and SMC is total volumetric soil moisture contents. LSM output variables on the left of Equation 1 are used as TOPMODEL input; water that reaches the ground-surface and evapotranspiration. From this input, hydrologic components from the right side can be calculated by TOPMODEL. R package 'rwrflhydro', distributed by NCAR, was applied extensively to handle Network Common Data Form (NetCDF) output data of WRF-Hydro [25].

4.3 TOPMODEL

TOPMODEL itself has been claimed as a conceptual approach tool by the original paper [10], [12], however, multiple efforts have developed it into an accessible hydrological modeling tool [26], [27]. In this research, we used implementation version R of TOPMODEL [28]. R.TOPMODEL offers a topographical data process function to calculate topographic index, but topography was partially processed by TauDEM [29].

$$TWI = \ln \quad \text{(Equation 2)}$$

The following topographic index distribution was used for input to TOPMODEL (Figure 4).

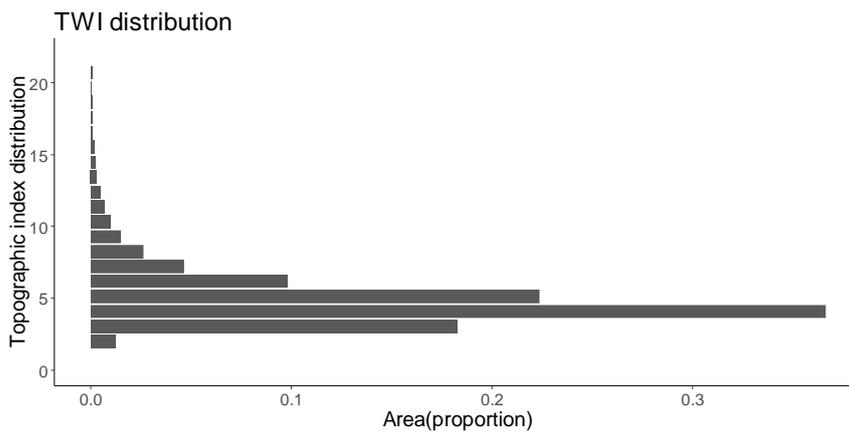


Figure 4. *Topographic index distribution in the study area*

Reasonable parameter bounds were set to random sample parameter sets (Table 1) [31]–[33]. Surface hydraulic conductivity (*k*₀) was the exception, setting the upper bound very high (100m/h), to allow the model to simulate the case where no infiltration excess occurs at all. TOPMODEL ran 50,000 times within 5 minutes on a laptop environment (Intel Core i7-8750H). That showed this type of model coupling may benefit in lower time demands for parameter calibration/optimization process, however, there is no guarantee the same will apply for a different type of ‘tighter’ coupling. The best parameter set with highest NSE was recorded and used to compare the simulated streamflow in time series.

Table 1. Range of reasonable TOPMODEL parameter ranges [31]–[33]

Parameter	Min	Max	Description	Sensitivity[31]	Reference
Qso (m/h)	0.0005	0.00075	Initial subsurface flow per unit area	Insensitive	Abou-shanab et al., 2011
	0	0.0001			Cho et al., 2019
lnTe ($\ln(\text{m}^2/\text{h})$)	-3.3	-1.5	Areal average of soil surface transmissivity	More sensitive, peak flow	Abou-shanab et al., 2011
	-7	10			Cho et al., 2019
	0.5	1.5			Nan et al., 2011
m	0.054	0.085	Scaling parameter for transmissivity decline	Highly sensitive, base flow	Abou-shanab et al., 2011
	0.001	0.25			Cho et al., 2019
	0.0001	0.03			Nan et al., 2011
Sr0 (m)	0.00001	0.01	Initial root zone storage deficit	Insensitive	Abou-shanab et al., 2011
	0	0.01			Cho et al., 2019
Srmax (m)	0.02	0.1	Maximum root zone storage deficit	Sensitive, physically based	Abou-shanab et al., 2011
	0.005	0.08			Cho et al., 2019
	0	0.5			Nan et al., 2011
Td (h/m)	10	40	Unsaturated zone time delay	Less Sensitive	Abou-shanab et al., 2011
	0.001	40			Cho et al., 2019
	0.001	50			Nan et al., 2011
vch (m/h)	2300	4000	Channel routing velocity	Sensitive, time to peak	Abou-shanab et al., 2011
	50	2000			Cho et al., 2019
	500	9000			Nan et al., 2011
vr (m/h)	50	2000	Surface routing velocity	Time to peak	Cho et al., 2019
	500	5000			Nan et al., 2011
k0 (m/h)	0.0011	0.1	Surface hydraulic conductivity	Physically based, infiltration excess	Abou-shanab et al., 2011
	0.0001	0.2			Cho et al., 2019
psi (m)	0.11	0.25	Wetting front suction; not included in r.TOPMODEL	Insensitive, Green-Ampt parameter	Abou-shanab et al., 2011
	0.01	0.5			Cho et al., 2019
dTheta	0.25	0.36	Water content change in wetting front; not included in r.TOPMODEL	Insensitive, Green-Ampt parameter	Abou-shanab et al., 2011
	0.01	0.6			Cho et al., 2019

4.4 Study area

Well-studied and data-rich headwater catchment was chosen as a study site, to test model coupling validity and to consider future study extension. This study focused on Angelo Coast Range Reserve (39°43'47" N, 123°38'34" W), an area within the Eel River Critical Zone Observatory. It is one of the University of California's Natural Reserves and is located on a tributary of the South Fork of Eel River in Northern California. The modeling was done for the sub-catchment of the USGS 'Elder Creek' stream gage site (11475560). The study site has a steep hillslope (average 32°) that drains to the Elder Creek catchment (approximately 17-km²), at 392 m above sea level [34]. The area's climate typically features warm, dry summers and cold, wet winters, with annual average rainfall of 2042 mm, and very little snow. Most precipitation occurs between October and May [35]

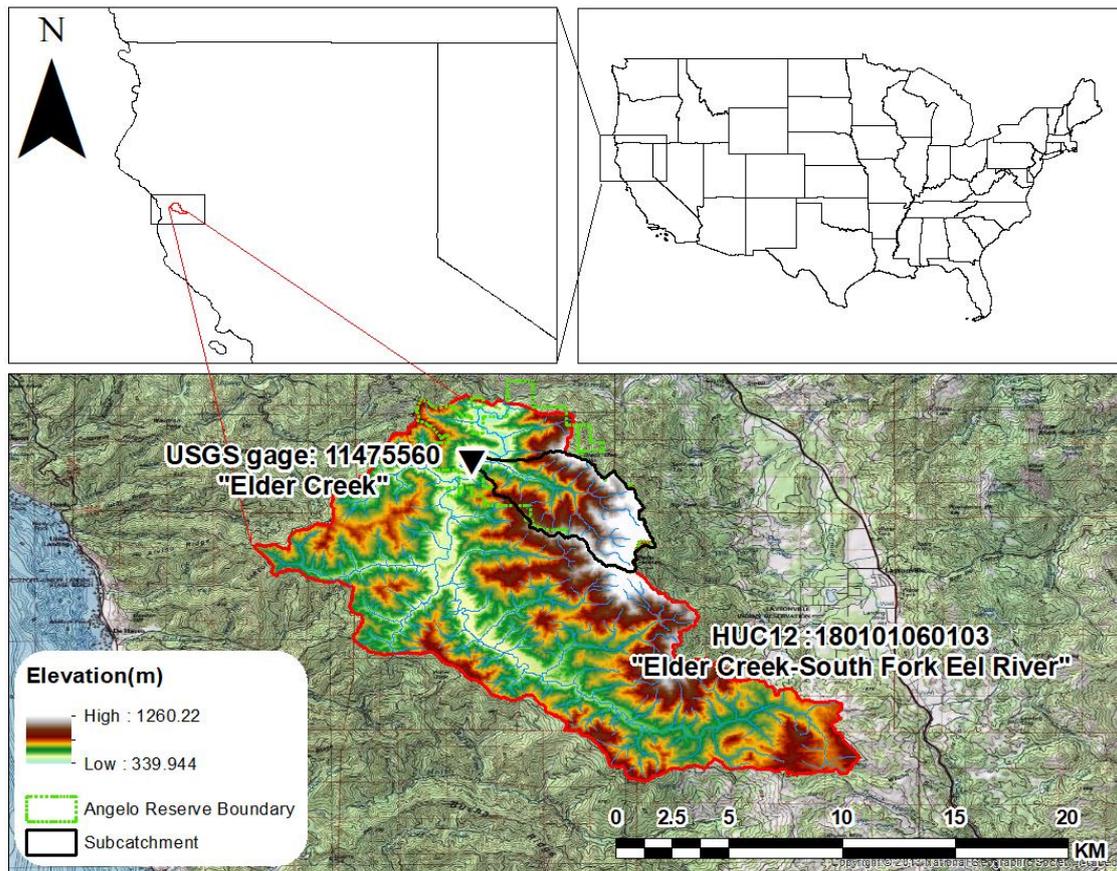


Figure 5. Study area: Elder Creek-South Fork Eel River watershed in California

The underlying geology of the study site consists of the Franciscan Formation Coastal Belt. This area is described as well-bedded, little sheared, and containing locally highly-folded mudstone-rich turbidities with interbeds and lenses of sandstone and conglomerate [36]. Thick weathered bedrock (5-25 meters) is overlaid with a thin soil layer (0.5–0.75 meters). Most of the site is underlain by almost vertically-dipping argillite that strikes approximately parallel to the hillslope gradient. Groundwater levels indicate flow is directed across, rather than along, the bedding orientation. Pervasive fracturing likely prevents the bedding from controlling groundwater flow direction. Along the eastern divide, a sandstone interbed is exposed at the surface [36].

5. Results

In this section, we compared the model performance of simulated original NWM streamflow, and TOPMODEL that received the NWM's LSM output. Due to time constraint, streamflow was the only criterion applied for model evaluation. We used Percent Bias (PBIAS), Nash-Sutcliffe Efficiency (NSE), root mean square error to the standard (RSR) as a statistical score. The flow duration curve, one of the simplest hydrologic signatures, was generated to briefly inspect the model's capability to predict water availability in the watershed.

For hourly streamflow, the NWM has shown a notable difference in hydrograph from the observed value (Figure 6). It returns poor statistical scores at PBIAS, 16.80%, NSE, 0.31, and RSR, 0.83. Most critically, the NWM consistently failed to simulate the magnitude of peak flow, predicting it to be significantly smaller than it actually was. According to the flow duration curve, the NWM prediction for water availability during low flow season is better. However, the slope and overall flow duration curve shape are considerably 'off' from observed values for the remaining 80% of the time (Figure 7).

In contrast, the TOPMODEL coupled version showed high similarity on hydrograph, except during low flow conditions (Figure 6). Similarly, statistical scores greatly improved compared to the NWM at PBIAS 1.67(%), NSE 0.75, RSR 0.50. Overall, flow duration curve shape and slope are almost identical between models for 75% of the simulated period, but the TOPMODEL coupled version overestimated water availability during low flow season. During the parameter optimization process, the TOPMODEL was shown as highly sensitive to parameter $\ln T_e$, and moderately sensitive to qs_0 and m parameter (Figure 8). Order of sensitivity shows some gap between previous research[25], implying parameter sensitivity trend may differ by cases (Table 1).

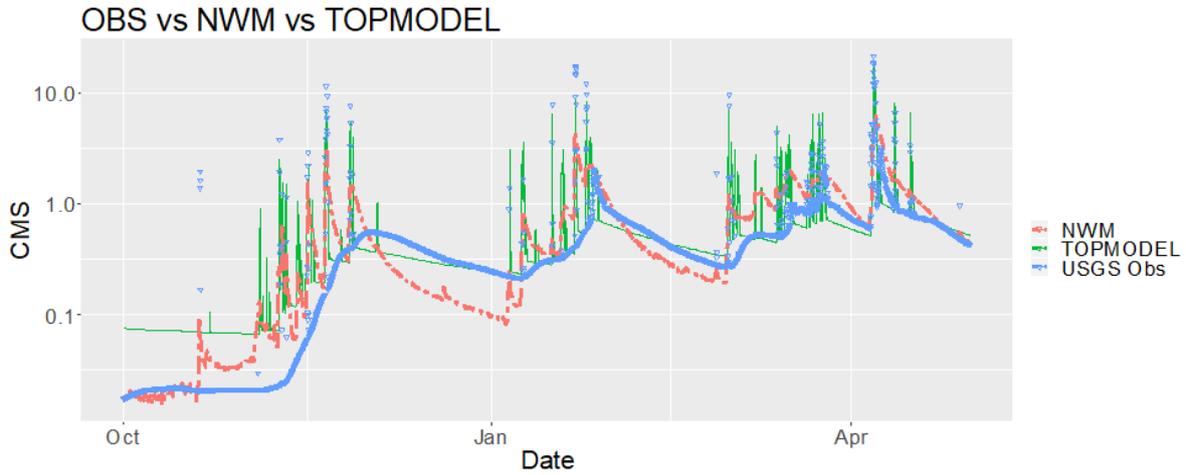


Figure 6. Streamflow: NWM vs TOPMODEL vs USGS observed

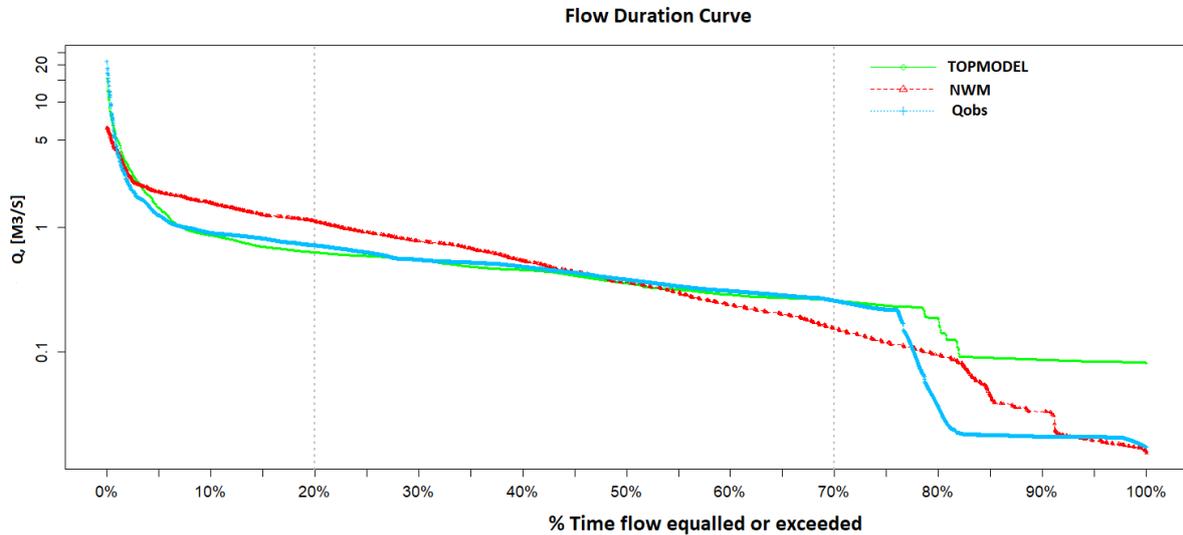


Figure 7. Flow duration curve: NWM vs. TOPMODEL vs. Observed

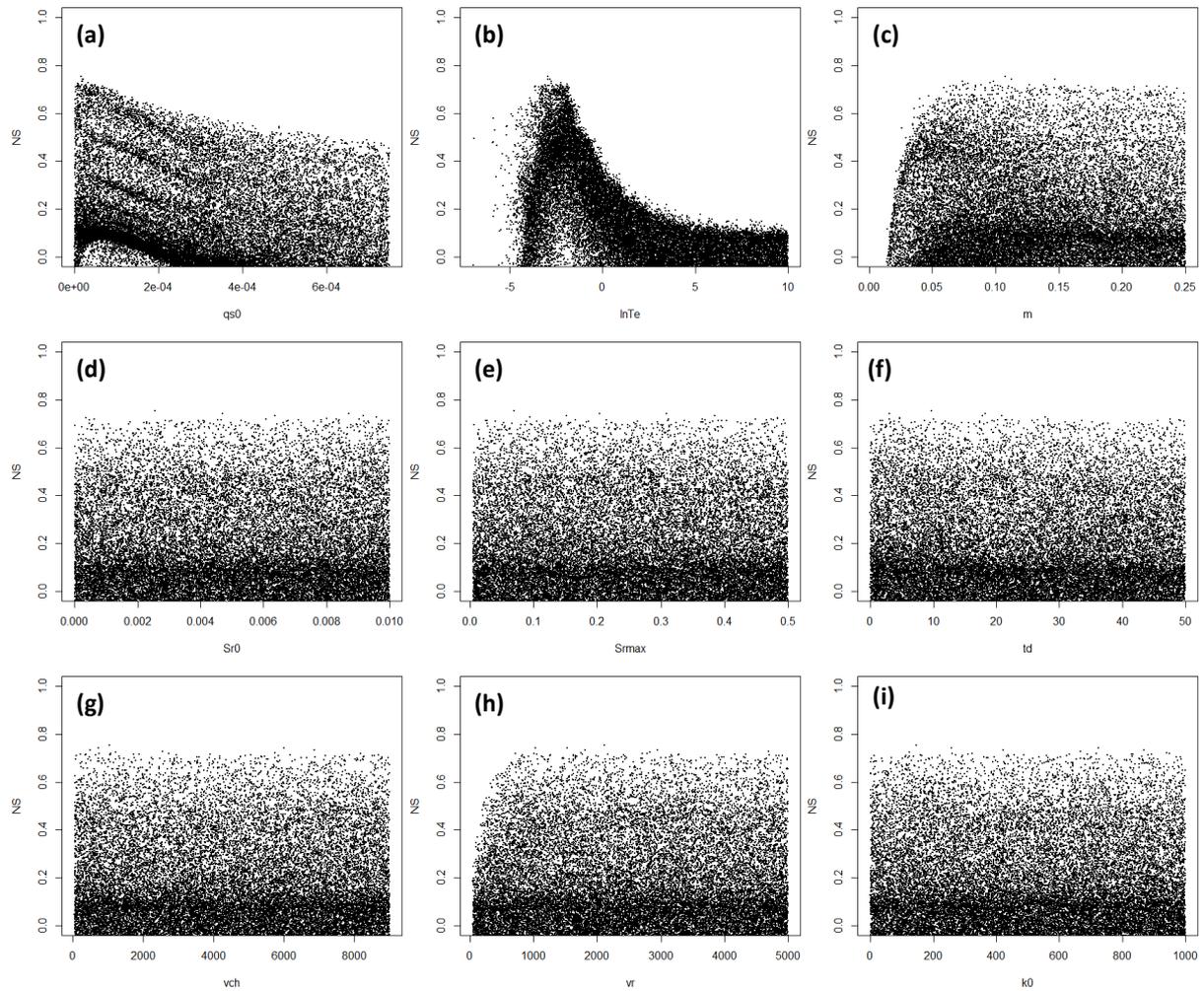


Figure 7. Parameter sensitivity. The x axis represents TOPMODEL parameters, y axis represents the Nash-Sutcliffe Efficiency. (a): qs_0 , (b): $lnTe$, (c): m , (d): Sr_0 , (e): Sr_{max} , (f): td , (g): vch , (h): vr , (i): k_0

6. Conclusion

This experiment showed a simple, one-way coupling of TOPMODEL with NWM's LSM outperformed the NWM v1.2, for simulating both flood peak magnitudes and water availability via the flow duration curve. Such a result could be due to different degrees of detailed topographic data usage, runoff and flow processes representation, parameter calibration process, or a combination of these factors. The one-way TOPMODEL and NWM coupling is structurally less complex, compared to WRF-Hydro coupling, and is capable of carrying more accurate detail. The detachment of distributed LSM's output from NWM's other modules is relatively simple and can benefit TOPMODEL. Running modified NWM's LSM does not reduce runtime, but the parameter calibration process of TOPMODEL is very time efficient.

This work had some limitations. One issue arose from the type of coupling done between TOPMODEL and the LSM. It does not allow models to update the internal state reciprocally, and this resulted in soil moisture being considered twice when calculating evapotranspiration (ET). This can result in underestimation of ET. However, it can be easily solved by bypassing the ET calculation scheme of TOPMODEL. Another limitation was that the LSM's output was averaged for the area of LSM cells that falls in the watershed boundary, not for the exact watershed area with area weighting. This inaccurate zonal averaging is attributable to

complication and discrepancy in the coordinate reference system of WRF-Hydro grids in NetCDF format.

Coupling of TOPMODEL was a test case to show that even a simplistic loose coupling of an existing lumped or semi-distributed model with LSM can lead to a better result. A modular approach to allow multiple physics options (e.g. the Structure for Unifying Multiple Modeling Alternatives or SUMMA [37]) is expected to be one future direction that hydrologic modeling may take. From the study results, we assumed it was reasonable to consider a modular approach in structural level. The national scale hydrologic modeling framework that can adopt different model structures and spatial representation based on specific needs, instead of a single, uniform structure over the whole domain, could have some benefit over current NWM in streamflow prediction.

We suggest further research be conducted to assess the impact that coupling between hydrologic models of different complexities can have on model performance. This may involve different levels of integration or coupling in hydrologic modeling: 1) a loose one-way coupling based on external data sharing without structural reformation; 2) a one-way coupling that manages the overlapping internal structure or components of the models, 3) a tight two- or multiple-way coupling with different models communicating and updating internal state through an overlapping structure or component. Would there be a significant gap between model performance due to different applications? How important is the numerical inaccuracy that can occur from coupling practice, when coupling a parameter-parsimonious model? Answering such questions could provide insight into the development of a coupling approach for models with a different structure.

The successful application of TOPMODEL in this work shows that partial simplification of NWM structure at headwater catchments can be achievable and desirable, which could allow the model to be more realistic and computationally efficient.

Supplementary Materials:

Experimental 1-way TOPMODEL-NWM coupling, HydroShare,
<http://www.hydroshare.org/resource/aa8fb91200294550a26f1cb10aeccac9optional>. (currently private).

References

1. "The National Water Model," *Office of Water Prediction*, 2019. .
2. D. J. Gochis, W. Yu, and D. Yates, "The WRF-Hydro model technical description and user's guide, version 3.0.," *NCAR Tech. Doc.*, no. May, p. 120, 2015.
3. Z. G. Xue *et al.*, "Modeling hydroclimatic change in Southwest Louisiana rivers," *Water (Switzerland)*, vol. 10, no. 5, pp. 1–20, 2018.
4. R. A. McManamay and C. R. Derolph, "Data descriptor: A stream classification system for the conterminous United States," *Sci. Data*, vol. 6, pp. 1–18, 2019.
5. M. Stieglitz, D. Rind, J. Famiglietti, and C. Rosenzweig, "An Efficient Approach to Modeling the Topographic Control of Surface Hydrology for Regional and Global Climate Modeling," *J. Clim.*, vol. 10, pp. 1–20, 1996.
6. W. Zhao and A. Li, "A Review on Land Surface Processes Modelling over Complex Terrain," *Adv. Meteorol.*, vol. 2015, 2015.
7. D. J. Gochis *et al.*, "WRF-Hydro V5 Technical Description Originally Created : Updated : WRF-Hydro V5 Technical Description," *NCAR Tech. Note*, 2018.
8. B. Van Den Hurk, M. Best, P. Dirmeyer, A. Pitman, J. Polcher, and J. Santanello, "Acceleration of land surface model development over a decade of glass," *Bull. Am. Meteorol. Soc.*, vol. 92, no. 12, pp. 1593–1600, 2011.

9. M. P. Clark *et al.*, “Improving the representation of hydrologic processes in Earth System Models,” *Water Resour. Res.*, no. 4, pp. 1–27, 2015.
10. K. J. Beven and M. J. Kirkby, “A physically based, variable contributing area model of basin hydrology,” *Hydrol. Sci. Bull.*, vol. 24, no. 1, pp. 43–69, 1979.
11. K. Beven, “TOPMODEL: a critique,” *Hydrol. Process.*, vol. 11, no. 9, pp. 1069–1085, 1997.
12. K. J. Beven and M. J. Kirkby, “Towards a Simple Physically-based, Variable Contributing Area Model of Catchment Hydrology,” 1976.
13. M. J. Kirkby and D. Weyman, “Measurements of contributing area in very small drainage basins,” *Dep. Geogr. Univ. Bristol*, vol. No. 3, no. Seminar Paper Series B, 1974.
14. M. J. Kirkby, “Hydrograph Modeling Strategies,” *Process. Phys. Hum. Geogr.*, pp. 69–90, 1975.
15. K. Beven, R. Lamb, R. Quinn, R. Romanowicz, and J. Freer, “Topmodel,” in *Computer Models of Watershed Hydrology*, V. P. Singh, Ed. Highlands Ranch, Colorado: Water Resources Publications, 1995, pp. 627–668.
16. K. Beven, “Infiltration into a class of vertically non-uniform soils,” *Hydrol. Sci. J.*, vol. 29, no. 4, pp. 425–434, 1984.
17. H. P. Deng and S. F. Sun, “Incorporation of TOPMODEL into land surface model SSiB and numerically testing the effects of the corporation at basin scale,” *Sci. China Earth Sci.*, vol. 55, no. 10, pp. 1731–1741, 2012.
18. B. Decharme and H. Douville, “Introduction of a sub-grid hydrology in the ISBA land surface model,” *Clim. Dyn.*, vol. 26, no. 1, pp. 65–78, 2006.
19. L. Bouilloud *et al.*, “Coupling the ISBA Land Surface Model and the TOPMODEL Hydrological Model for Mediterranean Flash-Flood Forecasting: Description, Calibration, and Validation,” *J. Hydrometeorol.*, vol. 11, no. 2, pp. 315–333, 2010.
20. B. Vincendon *et al.*, “Benefit of coupling the ISBA land surface model with a TOPMODEL hydrological model version dedicated to Mediterranean flash-floods,” *J. Hydrol.*, vol. 394, no. 1–2, pp. 256–266, 2010.
21. J. Jeziorska and T. Niedzielski, “Applicability of TOPMODEL in the mountainous catchments in the upper Nysa Klodzka river basin (SW Poland),” *Acta Geophys.*, vol. 66, no. 2, pp. 203–222, 2018.
22. H. K. Senatore, Alphonse, Giuseppe Mendicino, David J. Gochis, Wei Yu, David N. Yates, “Fully coupled atmosphere-hydrology simulations for the central Mediterranean: Impact of enhanced hydrological parameterization for short and long time scales,” vol. 7, no. 4, pp. 1693–1715, 2004.
23. J. Arnault *et al.*, “Role of Runoff–Infiltration Partitioning and Resolved Overland Flow on Land–Atmosphere Feedbacks: A Case Study with the WRF-Hydro Coupled Modeling System for West Africa,” *J. Hydrometeorol.*, vol. 17, no. 5, pp. 1489–1516, 2015.
24. T. Castronova and Danielle Tijerina, “CUAHSI/nwm_subsetting: CUAHSI Subsetter,” Jul. 2019.
25. J. L. McCreight, A. Dugger, and D. J. Gochis, “The Rwrhydro R Package,” 2015. .
26. H. Cho, “GIS Hydrological Modeling System by Using Programming Interface of GRASS,” Kyungpook National University, 2000.
27. O. Conrad, “SAGA-GIS Module Library Documentation (v2.1.3) Module TOPMODEL,” 2003. .
28. W. Buytaert, “Implementation of the Hydrological Model TOPMODEL in R,” 2018.
29. D. G. Tarboton, “Terrain analysis using digital elevation models (TauDEM),” *Utah State Univ. Logan*, 2005.
30. P. Campling, A. Gobin, K. Beven, and J. Feyen, “Rainfall-runoff modelling of a humid tropical catchment: The TOPMODEL approach,” *Hydrol. Process.*, vol. 16, no. 2, pp. 231–253, 2002.
31. A. Sigdel, R. Jha, D. Bhatta, R. V. Abou-Shanab, Reda A.I., Sapireddy, and B.-H. Jeon, “Applicability of TOPMODEL in the Catchments of Nepal: Bagmati River Basin,” *Geosystem Eng.*, vol. 14, no. 4, pp. 181–190, 2011.
32. H. Cho, J. Park, and D. Kim, “Evaluation of four GLUE likelihood measures and behavior of large parameter samples in ISPSO-GLUE for TOPMODEL,” *Water (Switzerland)*, vol. 11, no. 3, 2019.
33. Z. Nan, L. Shu, Y. Zhao, X. Li, and Y. Ding, “Integrated modeling environment and a preliminary application on the Heihe River Basin, China,” *Sci. China Technol. Sci.*, vol. 54, no. 8, pp. 2145–2156, 2011.
34. “Angelo Coast Range Reserve University of California Natural Reserve System,” 2019.
35. “PRISM Climate Group,” *Oregon State University*, 2017.

36. R. J. Mclaughlin *et al.*, “Geology of the Cape Mendocino, Eureka, Garberville, and Southwestern part of the Hayfork 30 x 60 Minute Quadrangles and Adjacent Offshore Area, Northern California,” *Misc. F. Stud. MF-2336*, pp. 2000–2000, 2000.
37. M. P. Clark *et al.*, “The structure for unifying multiple modeling alternatives (SUMMA), Technical Description, NCAR Technical Notes” *NCAR / TN 514+STR*, p. 54 pp., 2015.
38. Nadeau, T.; Rains, M. C. Hydrological connectivity between headwater streams and downstream waters : how science can inform policy. **2007**, *43* (1). <https://doi.org/10.1111/j.1752-1688.2007.00010.x>.
39. Niu, G. Y.; Yang, Z. L.; Mitchell, K. E.; Chen, F.; Ek, M. B.; Barlage, M.; Kumar, A.; Manning, K.; Niyogi, D.; Rosero, E.; et al. The Community Noah Land Surface Model with Multiparameterization Options (Noah-MP): 1. Model Description and Evaluation with Local-Scale Measurements. *J. Geophys. Res. Atmos.* **2011**. <https://doi.org/10.1029/2010JD015139>

Chapter 4

A Study on Parsimonious Models in Catchments Generating Saturation Excess Runoff

Federico Antolini¹, Stefany Baron², Brittany Lopez Barreto³ and Ishrat Jahan Dollan⁴

¹University of Iowa; federico-antolini@uiowa.edu

²Virginia Polytechnic and State University; stefbaron@vt.edu

³University of California, Merced; blopezbarreto@ucmerced.edu

⁴George Mason University; idollan@gmu.edu,

Summer Institute Theme Advisors: Fred Ogden, National Water Center, fred.ogden@noaa.gov;
Hilary McMillan, San Diego State University, hmcmillan@sdsu.edu

Academic Advisors: Eric Tate, *University of Iowa*; Randel Dymond *Virginia Polytechnic and State University*; Erin Hestir, *University of California, Merced*; Viviana Maggioni, *George Mason University*

Abstract: Conceptual hydrology models, e.g. TOPMODEL, represent an attractive alternative to distributed models, especially due to their simplicity, parsimony and computational efficiency. However, they often lack physical concepts and are unable to capture relevant hydrologic phenomena in certain areas. This study examined surface runoff generation in the Northeastern US, specifically the Hubbard Brook, NH, and Sleepers River, VT, watersheds, where saturation excess is the dominant runoff generation mechanism. We compared an established TOPMODEL version that uses watershed topography to predict subsurface and overland flow, against an alternate model we created that also accounts for soil and groundwater characteristics. Results showed TOPMODEL tends to over-predict peak flows from intense rainfalls and under-predict the others. The alternate model greatly over-predicts peaks, although this should be attributed to imperfect calibration of the delay function for overland flow routing. We used water volume error to assess model performance corresponding to peak flow events, and found errors vary across events but are, in general, comparable between the two models. The study found conceptual models can reproduce saturation excess runoff relatively quickly and with sufficient accuracy. When applied to small watersheds, model results can be passed to larger basins, overcoming the reliance of fully distributed models on coarse grids (> 50 m resolution) for hydrologic simulation.

1. Motivation

Hydrologic models serve, among other purposes, to predict stream flows in areas of interest during storm events. The ability to do this accurately is of utmost importance for hazard emergency management and watershed planning purposes. Computational efficiency is also an important characteristic, as models must be able to take advantage of the high quantity and quality of data currently available, to minimize use of computational resources. Conceptual models have the potential to be accurate and efficient at the same time. They are built around a few hydrologic concepts, need limited amount of data input and are parsimonious, since they use a minimal number of parameters. While parsimony results in easier calibration and fewer uncertainties, their simplicity may not allow the capture of all relevant hydrologic processes, thus limiting their applicability across scales, climates and landscapes.

The National Water Model (NWM) focuses on hydrologic predictions at the continental scale and uses grids of coarse resolution; 1 km for land surface and 250 m for routing [1]. Such resolution is, in most cases, too high to effectively model headwaters catchments, which represent the majority of modeled catchment. Additionally, the NWM has shown problems when modeling saturation excess runoff, which in some regions is the main runoff generation mechanism. This has repercussions on streamflow prediction.

2. Objectives and Scope

This study focused on saturation excess overland flow (Dunne runoff) and investigated the ability of two parsimonious models of reproducing streamflow values through rainfall-runoff transformation. This approach can be adapted and incorporated into the NWM, to improve its performance, taking full advantage of high-resolution data available (topographic and geologic) in a computational efficient fashion. This study did not attempt to explicitly model other fundamental hydrological processes, such as interception, evapotranspiration, snowmelt, or streamflow routing, which are already part of the NWM. Instead, the objective was to study possible saturation excess runoff formulations that can plug into the existing NWM framework.

Our work takes TOPMODEL as the conceptual base for the parsimonious models. A comprehensive description of TOPMODEL and its applications are provided in the following section. We used an R package implementing TOPMODEL as the first parsimonious model [2]. We developed a second parsimonious model (in the following, alternate model) based on TOPMODEL's primary assumptions, but also using physical equations for subsurface flow and groundwater table oscillations to characterize a watershed. Since the relationship between soil moisture and overland flow is crucial to quantify saturation excess runoff, we explicitly introduced the physics beneath this process in the alternate model.

We applied the TOPMODEL R version and the alternate model in two study areas in the Northeast U.S with topographic and geologic characteristics that make saturation excess the main mechanism of runoff generation. The results provide insight into the usefulness and accuracy of modeling saturation excess runoff using minimal input data and the least number of parameters. An accurate parsimonious model would greatly improve NWM predictions quality across different scales, even in small- and medium-sized watersheds, and could possibly benefit the NWM infrastructure by minimizing computational resource requirements.

Finally, we acknowledge the importance of reproducibility and continuous improvement in hydrology. Hence, all datasets used, and code written for this study are available for access through a GitHub and a Hydroshare repository. Both links are provided at the end of this report.

3. Background and Previous Studies

Hydrological processes in a catchment are dynamic and heterogeneous. Given a lack of measurements of state variables and catchment attributes, a minimal number of parameters is preferable for representing hydrological connectivity and catchment response. Detailed process modeling would require more calibrated parameters, and that would increase model uncertainty.

TOPMODEL simulates hydrological fluxes that rely on topographic information (catchment area, local slope, and topographic wetness index). Beven et al. [3] consider TOPMODEL as a set of conceptual tools that reproduce the dynamics of both surface- and subsurface-contributing behavior, in a semi-distributed way. The Topographic Wetness Index (TWI) represents the tendency to develop saturated conditions in the catchment. Areas with higher TWI values represent greater runoff generation. The index derives from the upstream area that drains through the unit contour at a designated point, and the slope of the local ground surface. A simple approximate relationship between catchment storage deficit and lateral transmissivity is

developed within the TOPMODEL framework. These relationship attributes form basic physical equations for simulating a catchment response. Given its semi-distributed functional framework, it bridges the gap between complex distributed process models and simple lumped concepts [4]. R's TOPMODEL package converts topographic effects into a distribution of classes while simulating runoff in the outlet.

TOPMODEL relies on three basic assumptions [3]: 1) steady-state condition for a saturated zone; 2) local hydraulic gradient of a saturated zone approximated by local surface topographic slope, $\tan \beta$, 3) an exponential transmissivity decay function with depth (or equivalent moisture deficit). TOPMODEL simulates three layers of the soil column (root, unsaturated and saturated zones) as three interconnected reservoirs. The parameters TOPMODEL deals with include: initial subsurface flow; transmissivity decay rate in the soil profile; hydraulic conductivity at the surface, and others explained fully by Buytaert [2].

Fractional saturated areas in a catchment and their correlation with soil moisture are key concepts for surface runoff formulation [5]. Conceptually, when precipitation falls over those saturated grids, it converts into surface runoff. With no precipitation, the groundwater table naturally lowers and soil water flows laterally down to the outlet where subsurface flow discharges to a stream. Lateral flow is only modeled in the saturated part of the soil since unsaturated flow takes more time to drain. More recent implementations address heterogeneity of soil moisture due to topography variation. For example, Gedney and Cox [6] use information of sub-grid topography variation and the height of the mean water table to predict low land saturation.

This study's alternate model proposes a simple relationship between saturated areas and soil water content, when saturation excess is the dominant type of runoff. While it is based on TOPMODEL assumptions, the alternate formulation can run without the TWI, to build a functional relationship. As more rain falls onto the watershed, the relationship generates a saturation excess runoff simulation. Thus, a characteristic diagram of saturated areas and volumetric content in soil can replace the topographic index of hydrological similarity in a catchment. This watershed characterization best applies to small-scale catchment (less than 10 km²), a scale most distributed hydrological models have difficulty capturing. Finally, the alternate model replaces the routing scheme by simply calculating overland flow volume generated at an hourly time step. This is potentially useful since a daily time step, with response time typically in hours, cannot capture the dynamics of a small catchment,

4. Methodology

4.1 Overview

In this study we ran hydrologic simulations using the TOPMODEL (R package) established version and an alternate TOPMODEL-based model we developed, that incorporates soil parameters. Simulated hydrographs were then compared to observed hydrographs, to evaluate the performance and the weaknesses of models.

4.2 Study Sites

4.2.1 Hubbard Brook

The Hubbard Brook watershed in North Woodstock, New Hampshire is an extensively studied watershed [7]. The availability of a longstanding hydrologic data set made it an ideal choice for a study site. Hubbard Brook forest has nine individual subcatchments, as shown in Figure 8. Subcatchment #7 (HB7) served as our preliminary study site. We chose HB7 because it is a substantially natural area where no artificial practice has ever been introduced for research purpose (e.g. clear cut of vegetation). HB7 size is approximately 0.75 km².

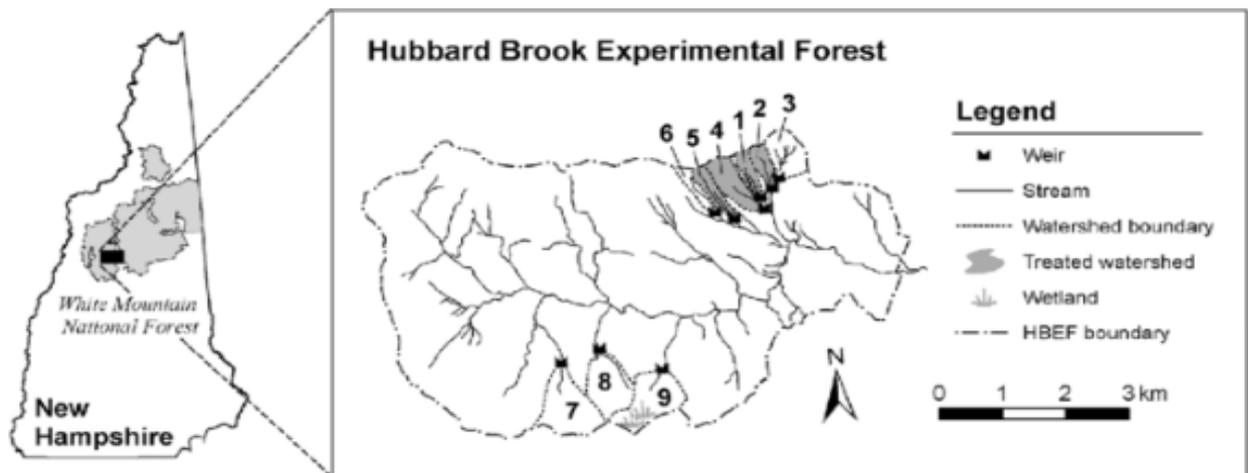


Figure 8. Hubbard Brook Watershed [8]

4.2.2 Sleepers River

The Sleepers River Research Watershed, in Vermont, has been an active hydrologic research site since 1959. Here, Dunne and Black (1970) determined controls of saturation-excess overland flow on streamflow generation. Specifically, we chose watershed W-3 as study area (Figure 9), covering an area of approximately 9 km².

Glaciers that covered New England thousands of years ago shaped the present landscape and impacted its hydrological patterns. Most of the watershed is covered by 1-4 meters of glacial till, leading to high buffered streamflow due to calcite weathering within the till. The climate is continental, with an average temperature of 6 degrees Celsius and average annual precipitation of 1.1 meters, of which approximately 20-30% is from snow [9].

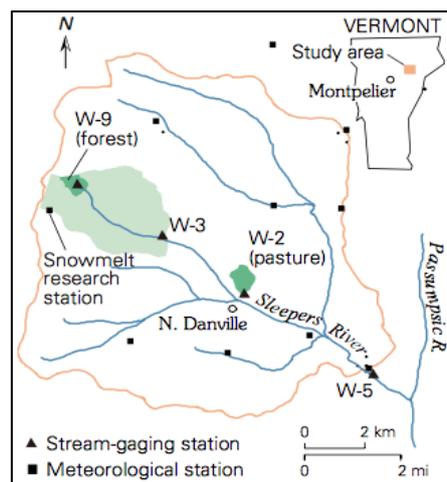


Figure 9. Sleepers River Watershed

4.3 TOPMODEL (R Package)

The study areas described in section 4.2 were modeled first with TOPMODEL. The main TOPMODEL function in R requires input of five variables: 1) a set of ten parameters; 2) a TWI data frame representing TWI values distribution for the Digital Elevation Model (DEM) of the

catchment; 3) a delay data frame representing channel routing through the watershed, and 4&5) precipitation and potential evapotranspiration values. For a more thorough description of these five input variables, see Buytaert [2].

The TOPMODEL R package has functions that can calculate TWI and delay data frames, using only the DEM of the catchment. Precipitation values were obtained from the Analysis of Record for Calibration (AORC) [10] and converted to meters of precipitation per hour. Flow data was obtained from the USGS and converted to cubic meters per hour per unit area. Evapotranspiration (ET) data was not available and, therefore was estimated using the water balance equation. We used the water balance method to find the average rate of ET for catchment W-3 was approximately 3 mm d^{-1} . With this information an ET vector was created that ranged from 2 to 5 mm d^{-1} , with colder months having lower values and hotter months having higher values. For future modeling, we recommend using independent measurements, or estimates of ET.

Parameters of TOPMODEL were estimated based on the range of values found in literature for sub-catchments W-3 of Sleepers River and Hubbard Brook. We also provided parameter ranges to a Hydroinformatics team of Fellows at the 2019 Summer Institute. The Fellows created 22,000 sets of parameters sampled with pseudo-random techniques [11], and performed a Sobol sensitivity analysis [12-13]. A full explanation of this sensitivity analysis can be found in their report, *A Visualization Workflow for Quantifying Parameter Sensitivities to Uncertainties for Hydrologic Models*. Through their sensitivity analysis, it was learned the parameters of initial subsurface flow (q_{s0}), rate of decline of transmissivity in the soil profile (m), and maximum root zone storage deficit (sr_{max}) are the most sensitive. Numerous experiments with different values for each parameter were performed to evaluate how each would affect simulated flow. One interesting conclusion from this analysis was that different months of the year should be calibrated with different values of q_{s0} . Therefore, TOPMODEL is not recommended for modeling a months-long time period. This study focused on a limited period of 2 weeks for HB7 (June 24th – July 7th, 2014) and 6 weeks for Sleepers River (April 30th – June 12th, 2017). These date ranges were chosen so snowmelt process could be considered exhausted and would not interfere with rainfall-runoff transformation.

4.4 Alternate Model

Our alternate model incorporated TOPMODEL assumptions to characterize the hydrologic behavior of a watershed, and applied a synthetic relationship to rainfall-runoff transformation. Three important assumptions of the alternate model were that in a given watershed: 1) the soil was highly porous, and thus contribution of infiltration excess (Hortonian runoff) to overland flow was negligible; 2) the unconfined aquifer thickness was spatially uniform across the watershed and limited to a few meters (<5), and the groundwater table and ground surface had the same slope, and 3) water transmissivity decayed exponentially as soil moisture deficit increased.

The alternate model was made of two distinct parts, or algorithms, as illustrated in Figure 3. Part 1 characterized the watershed, with a relationship between the proportion of saturated area and total soil water content or, equivalently, soil water deficit. Because saturation excess was generated by watershed areas susceptible to becoming completely saturated and hydrologically connected to the outlet, the amount of runoff based on estimated soil water content at the watershed level could be quantified. To accomplish this, the algorithm identified DEM flow direction based on the prevalent slope in the D8 scheme [14]. The algorithm started from an ideal condition of complete watershed saturation and simulated natural groundwater lowering and subsurface flow to the outlet. At each time step, it identified the saturated cells, calculated the ratio of saturated area over total area, and recorded the water amount stored in the aquifer,

in both unsaturated and saturated layers. Hence, corresponding values of saturated area ratio and normalized soil water content built a relationship useful for modeling saturation excess in an actual hydrologic model. The pseudocode in Algorithm 1 (below) summarizes the steps of Part 1.

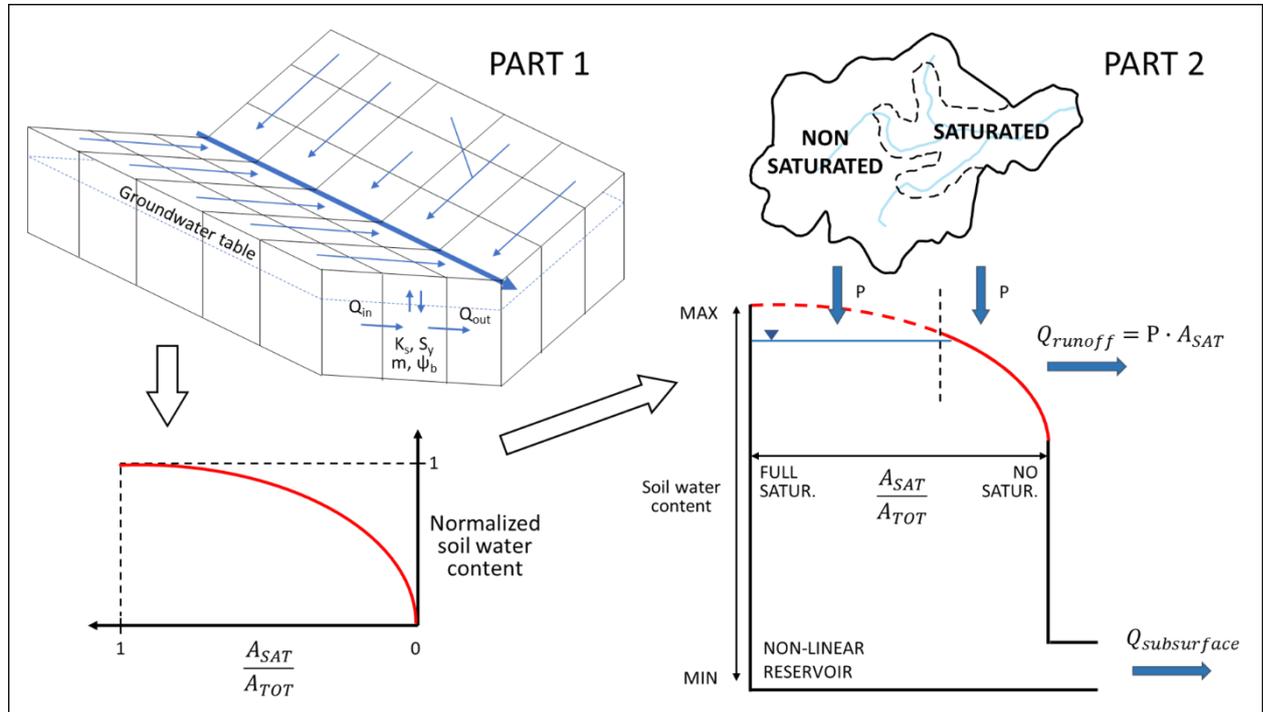


Figure 3. Alternate model overview. Part 1 simulated watershed drainage from an initial condition of complete saturation, to generate the relationship between saturated area and soil moisture. The relationship was used to model the non-linear reservoir in the rainfall-runoff transformation in Part 2.

Differently from Part 1, Part 2 moved away from a distributed model and treated the watershed as one reservoir that behaved according to the relationship defined in Part 1. In fact, Part 2 was a lump type model. It took precipitation and evapotranspiration as input and applied the mass balance equation to determine subsurface storage and overland flow at the watershed outlet. Subsurface flow was modeled with a non-linear function of the storage ratio:

$$Q_{subsurface} = b \cdot A \cdot \left(\frac{S}{S_{max}} \right)^n \quad (1)$$

where A was the watershed area, S the soil water storage, n (>1) a decay rate parameter and b the subsurface flow per unit area at full soil saturation. We calculated b reversing Equation 1 and used the streamflow record from the end of April, when snow was melted and the soil supposedly saturated. We estimated n with a trial-and-error approach, to better approximate the observed hydrograph. Pseudocode in Algorithm 2 (below) summarizes rainfall-runoff transformation algorithm in Part 2. The key concept was that soil water storage and surface runoff (eq. 1) were strictly related and follow the relationship from Part 1 (fig. 3).

5. Results

5.1 Hubbard Brook subcatchment #7

Figure 4 (below) shows observed and modeled streamflow from June 24th through July 7, 2014 in HB7. Both TOPMODEL and the alternate model reproduced fairly well the peak flow size for the main storm event that occurred during this time period. The time to peak for the models, however, was a few hours delayed from the observed. There were minor discharge increases around July 4th that were not observed, despite a 1-hour rain of significant magnitude. Overall, modeled hydrographs were very close to observed hydrographs and the models behaved quite satisfactorily in this study area. Still, the problem remains of how reliable precipitation forcing data are, with a very small catchment. While in a large watershed errors in precipitation values and/or timing might even out, and overall mitigate, similar errors in a small catchment may lead to significant errors in streamflow prediction, without possibility of amendment.

5.2 Sleepers River

Figure 5 (below) shows observed and modeled streamflow from April 30th to June 12th, 2017 in W-3 watershed of Sleepers River. Visually, from the graphs neither model performed as well here when compared to HB7. TOPMODEL overpredicted most peak flows and did not do a good job capturing peak events duration. The tendency was to overpredict peak flows from intense rainfall events and underpredict peak flows from weaker rainfalls. The alternate model performed well in recognizing flow peaks, including minor ones. Observed and modeled peaks aligned, and the alternate model did not generate more peaks than were observed. However, peaks were greatly overestimated at the beginning of the simulation and were shorter in time. Two weeks in the simulation peaks were better reproduced, with the last event, around June 6th, underestimated. The alternate model responded immediately to precipitation inputs and peaks had a sharp shape, especially on recession limbs. One limitation of the alternate model was the absence of a delay function, normally introduced in lumped models to account for water travel time from the most hydrologically distant areas, to the outlet [15].

The alternate model was qualitatively valid as it could predict when a discharge peak occurred but was not reliable for quantification of peak magnitude. Nash-Sutcliffe Efficiency and peak average error are typical ways to assess performance of a hydrologic model [16]. However, Figure 5 already shows high systematic discrepancies between modeled and observed streamflow at major peaks, and we expect the abovementioned indicators indicate a mediocre performance of that model. Many factors affected model overprediction, including: lack of a delay function; effective correspondence of model hypotheses to reality; imperfect model calibration; rainfall data quality, and the amount and variability of evapotranspiration through the season.

Hubbard Brook subcatchment 7, NH

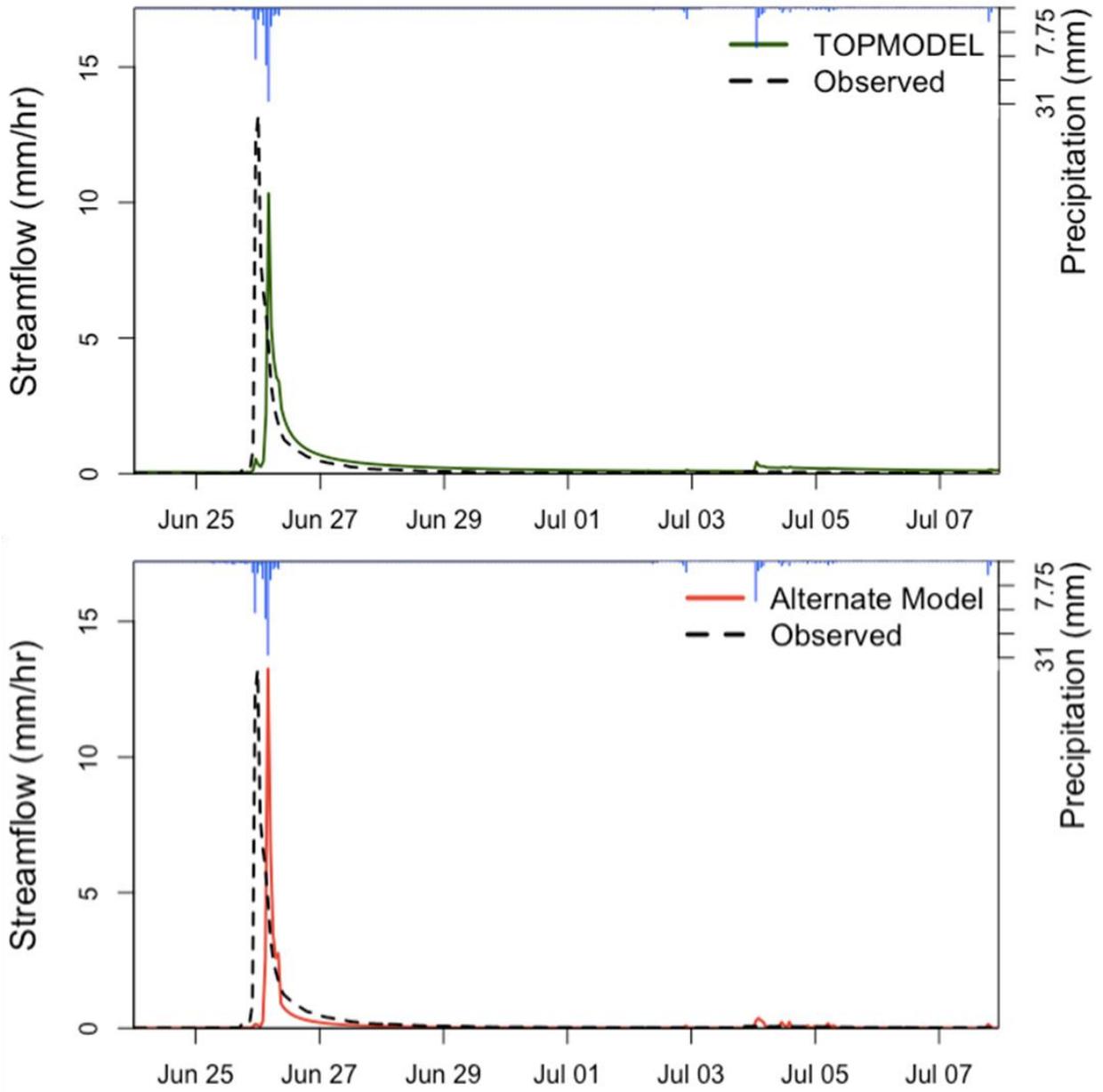


Figure 4. TOPMODEL and alternate model streamflow simulation in Hubbard Brook subcatchment 7 for the period from June 24th to July 7th, 2014. Decay parameter for transmissivity $m = 0.015$, as estimated from sensitivity analysis.

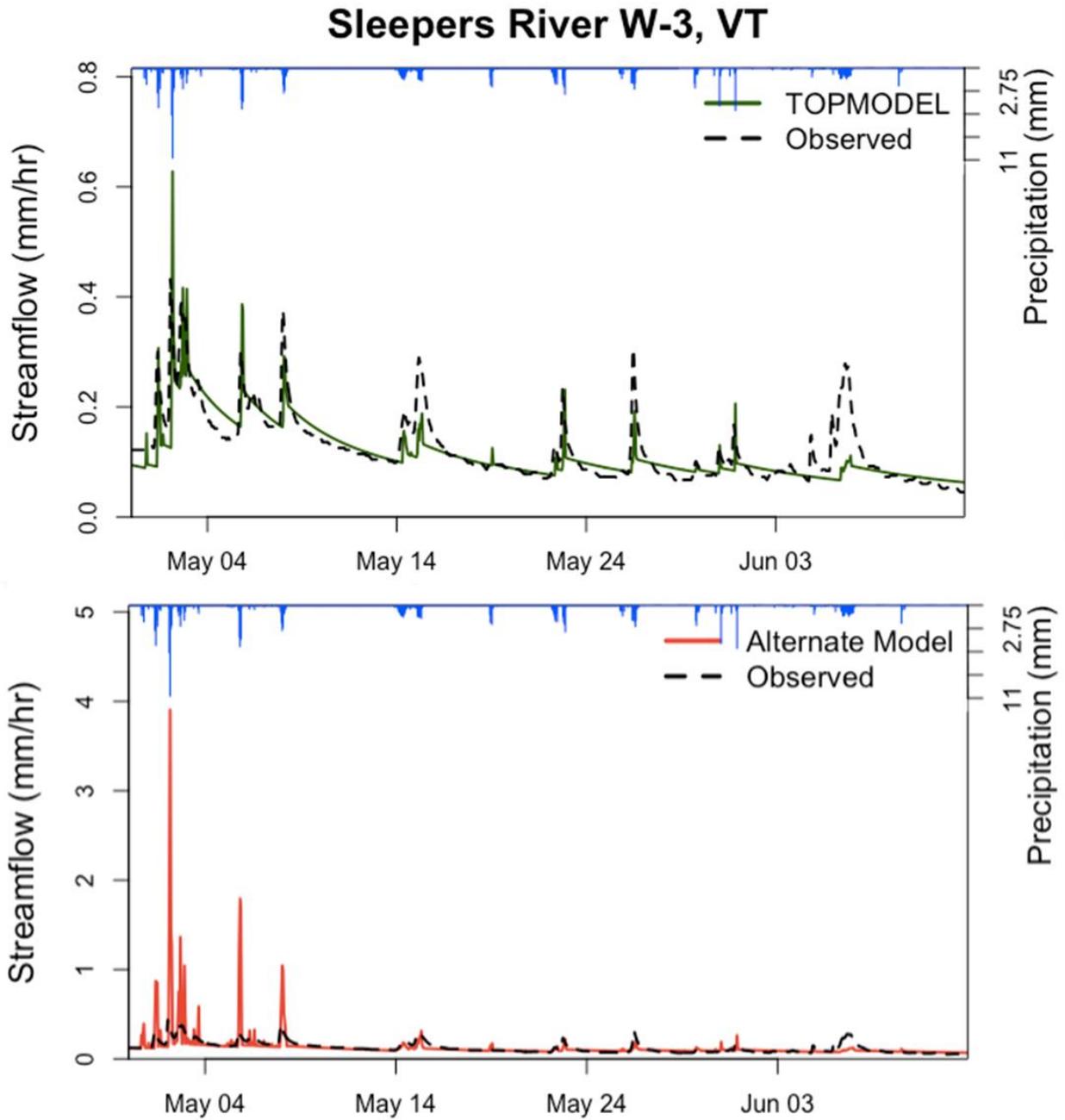


Figure 5. TOPMODEL and alternate model streamflow simulation in Sleepers River W-3 watershed for the period April 30th to June 12th, 2017. Decay parameter for transmissivity $m = 0.0279$, as estimated from sensitivity analysis.

To compare TOPMODEL’s performance with the alternate model, we computed the outflow per unit area for the top four peak events that occurred between May 1st and June 12th, 2017. We defined an event as three hours before to three hours after a peak occurrence. Thus, each event considered is a seven-hour duration. Volumes per unit area were computed by integrating flow per unit area over the seven-hour duration of each peak. Table 1 shows the alternate model performed better than TOPMODEL in all events considered, although in at least two cases the errors were very close. One explanation for the performance difference was that TOPMODEL does not explicitly model saturation excess, the main form of runoff in the study area. Another important difference was the way parameters were used in the two models. TOPMODEL characterized watershed hydrologic behavior mainly through TWI, determined exclusively by topographic information from the DEM. Other important hydrologic properties, such as soil

hydraulic conductivity and transmissivity decay, were introduced later as additional parameters. In contrast, the alternate model used topographic and soil information together to characterize watershed hydrologic behavior. Most relevant parameters were introduced at the beginning, and concurred to determine the relationship for saturation excess between soil water content and fraction of saturated area. The disadvantage of this approach was that it was more difficult to calibrate the model since a new relationship had to be built at each attempt. However, fewer parameters were used in the alternate model, and future development might be to add a specific calibration function, possibly incorporating sensitivity analysis rationale.

Table 1. *Volume per unit area of the top five peak events for Sleepers River during April 30 - June 12, 2017*

Date Range	Alternate Model volume per unit area (mm)	TOPMODEL Volume per unit area (mm)	Observed volume per unit area (mm)	Alternate % Error	TOPMODEL % Error
May 13 - May 17	10.8766	10.569	13.702	-21%	-23%
May 22 - May 23	2.850	2.591	3.401	-16%	-24%
May 25 - May 27	4.198	3.891	5.075	-17%	-23%
June 1- June 2	3.835	3.491	3.687	4%	-5%

6. Conclusion

Motivation for this study arose from the need for a hydrologic model that is accurate and simultaneously computationally efficient. Although conceptual models condense hydrologic concepts in simplified equations to represent water storage in catchments, they often overlook the physics beneath the process. Our study attempted to conjugate these two general trends and present a model formulation that follows TOPMODEL steps as a conceptual and parsimonious model, but also took advantage of high-resolution data and explicitly introduced physics.

We focused our efforts on saturation excess runoff generation, which is dominant in some U.S. regions. Groundwater tables rise and decline quickly, and terrain gets frequently saturated in those regions. The alternate model we created first established a relationship between soil moisture and saturated areas, then applied a simple rainfall-runoff routine using that relationship. The results were compared to TOPMODEL simulation and observed data, to assess model performance, and to highlight differences between them including weaknesses and potential application in regions needed.

Given the simplicity of parsimonious models, discrepancies between model results and observations are not surprising. The results indeed represent a satisfactory first step toward a successful application of TOPMODEL-based models. Some improvements and more testing can be made in the future, such as: refining model calibration; introducing a delay function based on watershed size and shape; analyzing larger watersheds in the context of scaling and continental climates, and possibly implementing other hydrologic routines, such as snowmelt and infiltration excess. TOPMODEL is known to work best in high precipitation areas with moderate topography and shallow, permeable soils [17], so it would be worthwhile to investigate whether the alternate model can be adapted to areas with different climates, landscapes and geology than those tested in this report.

The alternate model has considerable potential for incorporation into the NWM. Effective and fast simulations on headwater watersheds may allow results to be shared along the stream network and the fully distributed model currently in use to be applied with coarse grids, to just the main stem of major rivers. The results of this work have been constantly improving trend, especially during the final week of the Summer Institute, driving our optimism about the models' capability. The study is worth further investigation, and substantial improvements are within reach. Currently, the direction for hydrologic modeling is to integrate multiple approaches that

can be applied in different landscapes. The contribution of our study, which emphasizes model simplicity, flexibility and efficiency, becomes absolutely relevant in this context.

Supplementary Materials:

Our GitHub repository is found at <https://github.com/brittbarreto/Aquaholics> Anonymous, where data sets, scripts used from R's TOPMODEL and our alternate model's Python scripts are tracked. The material is also available in the Summer Institute 2019 Hydroshare group (<https://www.hydroshare.org/resource/1db946e29baa433b9ef5263a835f64e9/>).

References

1. Gochis, D. J.; Barlage, M.; Dugger, A.; FitzGerald, K.; Karsten, L.; McAllister, M.; McCreight, J.; Mills, J.; RafieeiNasab, A.; Read, L.; Sampson, K.; Yates, D.; Yu, W., The WRF-Hydro modeling system technical description, (Version 5.0). NCAR Technical Note. 2018. Source Code DOI: 10.5065/D6J38RBJ. Available online at <https://ral.ucar.edu/sites/default/files/public/WRFHydroV5TechnicalDescription.pdf>.
2. Buytaert, W., topmodel: Implementation of the hydrological model TOPMODEL in R, R package version 0.7.3, <https://CRAN.R-project.org/package=topmodel>. 2018.
3. Beven, K.; Lamb, R.; Quinn, P.; Romanowicz, R.; Freer, J., TOPMODEL. In *Computer Models of Watershed Hydrology*, 1 ed.; Singh, V., Ed.; Water Resources Publications: Highlands Ranch, CO, USA, 1995.
4. Nourani, V.; Roushani, A.; Gebremichael, M., TOPMODEL capability for rainfall-runoff modeling of the Ammameh watershed at different time scales using different terrain algorithms. *J. Urban Environ. Eng.* **2011**, 5 (1), 1-14, <https://doi.org/10.4090/juce.2011.v5n1.001014>.
5. Niu, G. Y.; Yang, Z. L.; Dickinson, R. E.; Gulden, L. E., A simple TOPMODEL-based runoff parameterization (SIMTOP) for use in global climate models. *J. Geophys. Res. Atmos.* **2005**, 110 (D21), <https://doi.org/10.1029/2005JD006111>.
6. Gedney, N.; Cox, P. M., The sensitivity of global climate model simulations to the representation of soil moisture heterogeneity. *J. Hydrometeorol.* **2003**, 4 (6), 1265-1275, [https://doi.org/10.1175/1525-7541\(2003\)004<1265:TSGCM>2.0.CO;2](https://doi.org/10.1175/1525-7541(2003)004<1265:TSGCM>2.0.CO;2).
7. USDA US Forest Service, Hubbard Brook Experimental Forest.
8. Cawley, K. M.; Campbell, J.; Zwilling, M.; Jaffé, R., Evaluation of forest disturbance legacy effects on dissolved organic matter characteristics in streams at the Hubbard Brook Experimental Forest, New Hampshire. *Aquat. Sci.* **2014**, 76 (4), 611-622, <https://doi.org/10.1007/s00027-014-0358-3>.
9. Shanley, J. B. *Sleepers River, Vermont: A water, energy, and biogeochemical budgets program site*; 2327-6932; US Geological Survey Fact Sheet-166-99: 2000, <https://doi.org/10.3133/fs16699>.
10. Kitzmiller, D. H.; Wu, W.; Zhang, Z.; Patrick, N.; Tan, X. The Analysis of Record for Calibration: A High-Resolution Precipitation and Surface Weather Dataset for the United States. AGU Fall Meeting Abstracts, 2018.
11. Kucherenko, S.; Albrecht, D.; Saltelli, A. Exploring multi-dimensional spaces: a Comparison of Latin Hypercube and Quasi Monte Carlo Sampling Techniques *arXiv e-prints* [Online], 2015. <https://ui.adsabs.harvard.edu/abs/2015arXiv150502350K> (accessed Aug 19, 2019).
12. Saltelli, A.; Ratto, M.; Andres, T.; Campolongo, F.; Cariboni, J.; Gatelli, D.; Saisana, M.; Tarantola, S., *Global sensitivity analysis: the primer*. John Wiley & Sons: 2008.
13. Sobol, I. M., Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. **2001**, 55 (1-3), 271-280, 10.1016/S0378-4754(00)00270-6.
14. Jenson, S. K.; Domingue, J. O., Extracting topographic structure from digital elevation data for geographic information system analysis. *Photogramm. Eng. Remote Sensing* **1988**, 54 (11), 1593-1600.
15. Beven, K.; Freer, J., A dynamic topmodel. *Hydrol. Process.* **2001**, 15 (10), 1993-2011, <https://doi.org/10.1002/hyp.252>.
16. McCuen, R. H.; Knight, Z.; Cutter, A. G., Evaluation of the Nash-Sutcliffe efficiency index. *J. Hydrol. Eng.* **2006**, 11 (6), 597-602, [https://doi.org/10.1061/\(ASCE\)1084-0699\(2006\)11:6\(597\)](https://doi.org/10.1061/(ASCE)1084-0699(2006)11:6(597)).
17. Sigdel, A.; Jha, R.; Bhatta, D.; Abou-Shanab, R. A.; Sapireddy, V. R.; Jeon, B.-H., Applicability of TOPMODEL in the catchments of Nepal: Bagmati River Basin. *Geosyst. Eng.* **2011**, 14 (4), 181-190, <https://doi.org/10.1080/12269328.2011.10541349>.

Chapter 5

National Water Model Dockerized Job Scheduler: A Reproducible Framework to Generate Parameter-based NWM Ensemble

Austin Raney¹, Iman Maghami², and Yenchia Feng³

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

²University of Virginia; im3vp@virginia.edu

³Stanford University; yenchia@stanford.edu

Summer Institute Theme Advisor: David L. Blodgett, United States Geological Survey, dblodgett@usgs.gov

Academic Advisors: Sagy Cohen, *University of Alabama*; Jonathan L. Goodall, *University of Virginia*; Oliver Fringer, *Stanford University*

Abstract: The process of hydrological model calibration is computationally burdensome and time consuming. Brute force and Monte Carlo style model evaluations are often used to perform this task, unequivocally requiring a colossal number of model simulations over a large domain of time. Frameworks capable of performing these approaches exist for the National Water Model (NWM), however, they 1) are not publicly available and 2) require the use of cluster computing (CC). In this report, we deliver an engine for perturbing NWM channel parameters, an ensemble generator for collating model outputs, and a dockerized job scheduler (DJS) capable of automating, managing, and executing numerous NWM simulations simultaneously. All these deliverables were developed using open source tools and programming languages capable of running agnostic to a given operating system. This framework was designed with the intent of removing the overhead of model setup and compilation, thereby lowering the barrier for entry without limiting performance, and also facilitating an environment for scientists to more easily test hypotheses. To test our tools, we ran our workflow on an uncalibrated catchment using NWM version 2.0. It is worth mentioning that performance of ensemble simulations reported was unfit, which was expected due to use of an uncalibrated catchment. However, this is of little importance as the robust, scalable computational performance of our framework is the greatest result of our work.

1. Motivation

Motivation for producing such a framework evolved from the desire to test our hypothesis: that an ensemble average produced by perturbing channel parameters may yield a more accurate result in the case of extreme events. This hypothesis was not sufficiently tested in this report due to time constraints and guidance provided by our advisors on the impact factor of a reproducible framework. However, a framework capable of answering this question was constructed, motivated originally by the requirements of sufficiently testing our hypothesis. Specifically, it automated repeat NWM runs with differing parameter value sets. Performing this task manually within a cluster computing (CC) environment would require much repetition, and thus be impractical as perturbation numbers increased. Hence, development of a reproducible automated framework transcending CC arose from necessity. However, while this framework met our project needs, it was not customized for our particular use case. Instead, we designed our framework with a high level of abstraction, hoping to deliver a product capable of meeting the needs of the greater NWM/Weather Research and Forecasting-Hydro (WRF-Hydro)

modeling communities. Our team strived to create a tool that could be used by both novice and veteran users.

2. Objectives and Scope

After developing a hypothesis, we set out to test it by creating a scalable automated framework capable of simultaneously running numerous NWM simulations of different channel parameters. To accomplish this we developed a perturbation engine, DJS, and an ensemble generator, to form the framework. We tested the framework's functionality and computational performance on a watershed in Pocono, Pennsylvania, an uncalibrated domain obtained from the National Center for Atmospheric Research (NCAR). This watershed includes both urbanization areas and reservoirs. Initially, we had planned to evaluate our hypothesis on a more idealized domain, without urbanized features and lakes, but this was not possible as the additional features pushed us beyond our initial scope. However, this report's impact was not diminished, as a higher impact was indicated both by National Water Center (NWC) staff and our theme leaders, in developing a reproducible, open-source framework that can run concurrent NWM simulations. Thus, we ranked development and testing our framework on NWM v2.0 a higher priority. Specifically, the framework's module, DJS, is capable of running all publicly available NWM versions and thus can be used to run NWM v1.2 on domains subset, for example, using the CUAHSI subsetter [1], and all other compatible NWM v2.0 domains.

3. Methodology

In this section we first elaborate the updated channel geometry of NWM 2.0. Then we outline the forcings and case study area of our simulations. Continuing on, we describe the perturbation engine used to vary channel parameters. Following that, we explain the different elements of the DJS, and how these elements make the tasks of automating, managing, and executing numerous NWM simulations occur simultaneously. Finally, we outline our ensemble generators algorithmic process to statistically amalgamate numerous model simulations.

3.1. NWM v2.0 Channel Formulation

The newest operational National Water Model (NWM v2.0) uses regression-based equations [2] to set up channel geometry parameters of an idealized trapezoidal channel cross-section:

$$T_w = 2.44 * (DA)^{0.34} \quad (1)$$

$$A = 0.75 * (DA)^{0.53} \quad (2)$$

where T_w is the bankfull elevation channel width (hereafter known as channel top width, in meters) and A is the channel cross-sectional area (in meters squared). The independent variable used to calculate T_w and A is DA , the upstream drainage area of the reach (in kilometers squared). coefficients and exponents in both equations (1) and (2) are empirically derived values based on regression fits over the contiguous U.S. (CONUS). Average bankfull channel depth, $d_{\text{regression}}$ (in meters), is calculated from T_w and A :

$$d_{\text{regression}} = A/T_w \quad (3)$$

To approximate trapezoidal channel depth, the NWM v2.0 models the bankfull channel depth by scaling $d_{\text{regression}}$ by a factor of 1.25:

$$d_{\text{model}} = 1.25 * d_{\text{regression}} \quad (4)$$

thus uniquely defining an isosceles trapezoidal channel cross-section with T_w , A and d_{model} . The rest of the channel geometry parameters can therefore be calculated from the three aforementioned parameters:

$$B_w = 2 * A / (d_{model} - T_w) \quad (5)$$

$$z = (T_w - B_w) / (2 * d_{model}) \quad (6)$$

where B is the channel bottom width (in meters) and z is the channel side-slope, defined as the ratio of run over rise.

It is worth noting that although NWM v2.0's method of establishing channel geometry parameters does not explicitly set up channel bottom width using channel top width, it does, in fact, effectively restrict the ratio of the channel bottom width and channel top width to be a fixed value of 0.6:

$$B_w = 0.6 * T_w \quad (7)$$

A detailed derivation of equation (7) is available on the GitHub repository for this project, provided at the end of this document in the Supplementary Materials section. As of the writing of this report, all channel bottom widths were 60% of channel top widths, in the CONUS route link file.

3.2. Forcing Data

Hourly quarter-degree NLDAS 2 primary forcings were obtained from GES Disc for the temporal period of January 1st, 2017 to June 30th, 2018. These forcings were subset to our Pocono, PA domain using the simple subset wizard on the GES Disc website [3]. The variables, air temperature, incoming shortwave and longwave radiation, surface pressure, u and v wind speed, and liquid precipitation rate were used to force the model.

3.3. Case Study Area

Our test case, provided by NCAR, resides in northeastern Pennsylvania, specifically encapsulating the town and greater region of Pocono. This watershed consists of head water streams that have a total drainage area of 291 km², maximum stream order of 5, and somewhat gently rolling topography (Figure 1). The watershed is dominated by forest and urbanization areas, and also contains lakes. There are two United States Geological Survey (USGS) streamflow gaging stations in the watershed: 01447720, Tobyhanna Creek near Blakeslee (located at the watershed outlet), and 01447680, Tunkhannock Creek near Long Pond, both in Pennsylvania, USA. In this study we compared NWM simulated streamflow, using our framework, versus the observed USGS gaging stations.

3.4. Perturbation Engine

Our perturbation engine produced new model parameter files (hereafter known as route link files) with perturbed channel parameters. This was accomplished by applying scalar values generated by Latin Hypercube Sampling (LHS) [4] to route link file parameters, for example channel bottom width. The user specifies the channel parameter(s) to perturb, and a scalar value range and number of samples for each parameter, as required by the LHS method. Then, input probability distributions for each parameter are stratified by dividing each cumulative curve into equal intervals (the number of intervals equal to the user-defined number of sample points). Finally, a sample point is randomly taken from each interval for each parameter while ensuring no two sample points share the same interval for any parameter. As opposed to a fully random

sampling methodology such as Monte Carlo, this methodology resulted in generating probability distribution for each parameter, using fewer number of sample points.

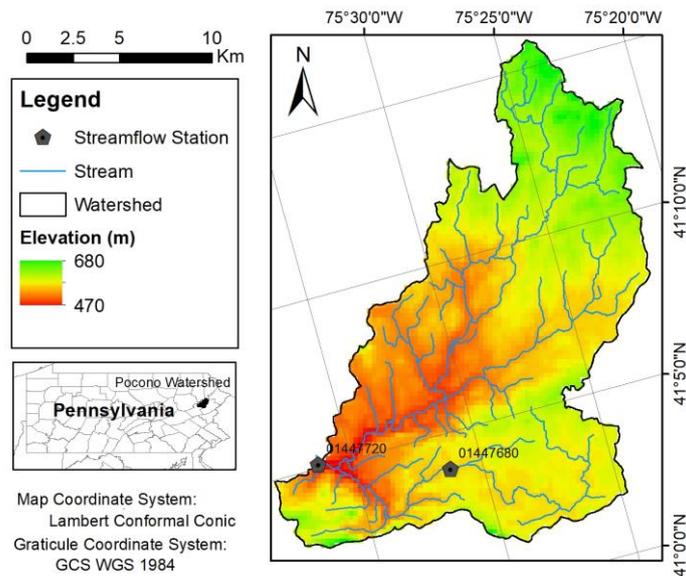


Figure 1. Study Area, The Pocono Watershed in Pennsylvania, USA.

Currently, the perturbation engine is capable of scaling the channel-related parameters relative to the default parameter values. When one parameter value depends on another, the dependent parameter is also updated once the independent parameter value is perturbed, (e.g., when keeping the top width and cross section area of a trapezoidal channel constant, if the bottom width is changed, the side slope also is changed accordingly, to satisfy related geometry equations).

Finally, corresponding input files containing target parameters are updated with new parameter values. These edited files are tagged with metadata (within their global variables list) to keep track of changes. The metadata contains changes made (i.e., scale values) and corresponding parameters, distinguishing between independent and dependent perturbed parameters.

3.5. Dockerized Job Scheduler (DJS)

Each DJS piece was designed to work independently. The DJS pieces are: docker images; jobs; the domain linker, and the job scheduler. Put simply, the *job scheduler* generates a queue of *jobs* where each *job* contains the properties needed for the *domain linker* to properly link files across the filesystem which then are used by a deployed *docker* container to run an NWM instance.

3.5.1. Docker Images

One main hindrance of using models that distribute tasks across Central Processing Units (CPUs) and cluster compute nodes using interfaces like the Message Passing Interface (MPI), is software and operating system dependencies for compiling a given model. Docker is an open-source, platform-independent, software-as-a-service alternative to classical Virtual Machine (VM) environments. Unlike traditional VMs, such as VirtualBox, Docker runs containers that use the user operating system's native kernel, thus significantly reducing the size of each container, compared to VMs. This underlying architecture reduction drastically improves both container spin up time and process runtime within containers, allowing numerous containers to run in tandem. Docker containers are spun up from Docker images with each image constructed using a Docker file containing instructions for forming the image (i.e. install dependencies). Importantly, Docker images require existing images to be stacked on one another, thus making

it possible to create a base environment image containing, for example, precompiled dependencies that can be used by other images to easily create images of differing software versions. For these reasons, Docker was chosen as the model running environment within DJS. All images and Docker files created during the Summer Institute are housed on GitHub and Docker Hub (Refer to the Supplementary Materials section).

With collaboration and reproducibility in mind, a set of Docker images were created for use with DJS. A base image containing all requirements for compiling the NWM was created based on version 3.10 of Alpine Linux [5], aptly named ‘aaraney/wrf-hydro_base’. This distribution was chosen to minimize overall container size. Images for v1.2, v2.0, and v2.1 beta of the NWM were created from the base image of each, with an overall image size of 274 megabytes (mb). Suitably named, ‘aaraney/nwm-djs’, each had its own model versioning tag.

3.5.2. Jobs

The class Job tracks location of alternative domain files, with primary and replica domain directory locations. Each job object also contains container-related information upon execution. The class is used by the Job Scheduler within a queue data structure of job objects, described later in this chapter.

3.5.3. Domain Linker

The domain linker creates replica directories of a user input primary domain directory that contains all necessary model inputs prior to runtime. The linker also handles placing the perturbed domain file in its proper replica directory, once created. To save disk space, each file linked to a replica domain is created using a hard link. It should be noted that many file explorer systems (e.g., Finder, Windows Explorer) double count hard link sizes, providing erroneous information on disk space taken up.

To avoid adjusting namelist files and to follow common WRF-Hydro file naming conventions, each perturbed input domain file is checked to identify its type, and link it with the appropriate name. Files provided by a user within the primary domain directory are also checked, and linked with conventional filenames to each replica domain, to follow standard file and directory naming conventions.

3.5.4. Job Scheduler

The task of spawning and tracking numerous jobs is handled by the Job Scheduler. This class object binds smaller pieces of DJS into a working unit, to form the main class a user would import and call to interact with the operational framework. The primary role of this module is to create a queue of jobs to be spawned once computational resources are available. That said, the maximum number of running jobs, the docker image to be used, the number of threads to allocate to each container, and the number of MPI processes each container launches are all set by the user as attributes in this module. For usage and further explanation, refer to the documentation on our GitHub repository link in the Supplementary Materials section.

3.6. Ensemble Generator

To test our hypothesis, that an ensemble average produced by perturbing channel parameters may yield a more accurate result in the case of extreme events, we developed the ensemble generator. First, it calculated any performance metrics sets chosen by the user (e.g, Nash–Sutcliffe Efficiency (NSE), Percent Bias (PB), etc.) for given sets of simulated streamflow time series versus observed values (e.g., USGS streamflow data). Then, for every performance metric, each simulated time series was ranked, and scored against the others based on the difference

between their calculated performance metric and the theoretical perfect value for the respective metric (e.g., 1 for NSE, and 0 for PB). Next, using either default equal metric weights (all metrics provided are of equal influence) or user defined metric weights (e.g., NSE is of greater importance than PB), a weighted average final score was given to each simulation. Finally, this score was used to generate an ensemble weighted average of all simulations.

4. Results

To implement the perturbation engine, DJS, and ensemble generator, we set out to test and benchmark our tools on a HUC 8 watershed in northeastern Pennsylvania. Greater detail about this watershed can be found in the study area section. In this section, we present our test case ensemble model simulation results, the effects of extreme channel geometries (near-triangular and near-rectangular channel formulations), and DJS computational benchmarking results. Prior to simulating any results found below, we spun-up NWM for a period of one year (2017-2018) using NLDAS forcing data. Each subsequent model simulation in this report, with the exception of benchmarking simulations, was performed over a six-month period, from January 1st, 2018 to June 30th, 2018.

4.1. Case Study: Parameter-Based Ensemble Outputs and Extreme Channel Geometries

4.1.1. Introduction

For the sake of simplicity, in this case study we only perturbed the channel bottom width parameter. Given that channel top width and channel cross section area values are both obtained from regression-based equations, as outlined in the NWM v2.0 Channel Formulation section, we decided to keep the channel top width and channel cross section area constant and adjust the channel side slope as a dependent parameter, to conserve the geometry for the trapezoidal channel. The NWM simulations were created by running 12 different perturbed runs: (1) Default bottom width, (2) A near triangular channel ($B_w * 0.0001$), and (3) A near rectangular channel ($B_w * 1.6$), (4-12) bottom width multiplied by near-randomly chosen scales, using the LHS method (scale range: 0.01-1.6). Added to these 12 runs we generated an ensemble weighted average of all simulations (excluding extreme channel cases: near-triangular and near-rectangular).

4.1.2. Findings

Figure 2 shows the results for (a) the extreme channel cases simulations, and (b) the ensemble weighted average for a 3-day subset (February 25-28, 2018) of our 6-month simulation period, to better show differences between simulations. Figure 2.a shows that, generally, the extreme cases of near-triangular and near-rectangular channels comprised the lower and upper bounds of streamflow simulations. During peak periods (in the rising limbs of hydrograph) the near-triangular channel predicted the highest streamflow values among all simulations, although it was the lowest in the falling limb of hydrograph. The near-rectangular channel, however, performed exactly the reverse: it was generally the lowest during the rising limb and highest during the falling limb. The time to peak was always smallest for the near-triangular channel while largest for the near-rectangular channel. These results were expected, as a narrow channel (e.g., triangular channel) yields a faster flow while a wide, shallow channel (i.e., rectangular channel) produces a slower flow, given the additional friction between the water and the stream-wetted perimeter. Figure 2.b shows the ensemble weighted average hydrograph having similar characteristics to the default simulation. This was because all perturbed simulations performed similarly, as channel bottom width was the only parameter perturbed, and the cross-section area of the channel and top width were kept constant. Thus, the results did not vary greatly from one another and consequently, the weighted average ensemble of these members did not reveal a

more accurate result. However, this can be attributed to use of an uncalibrated domain, therefore limiting the significance of our simulated results when compared to observed.

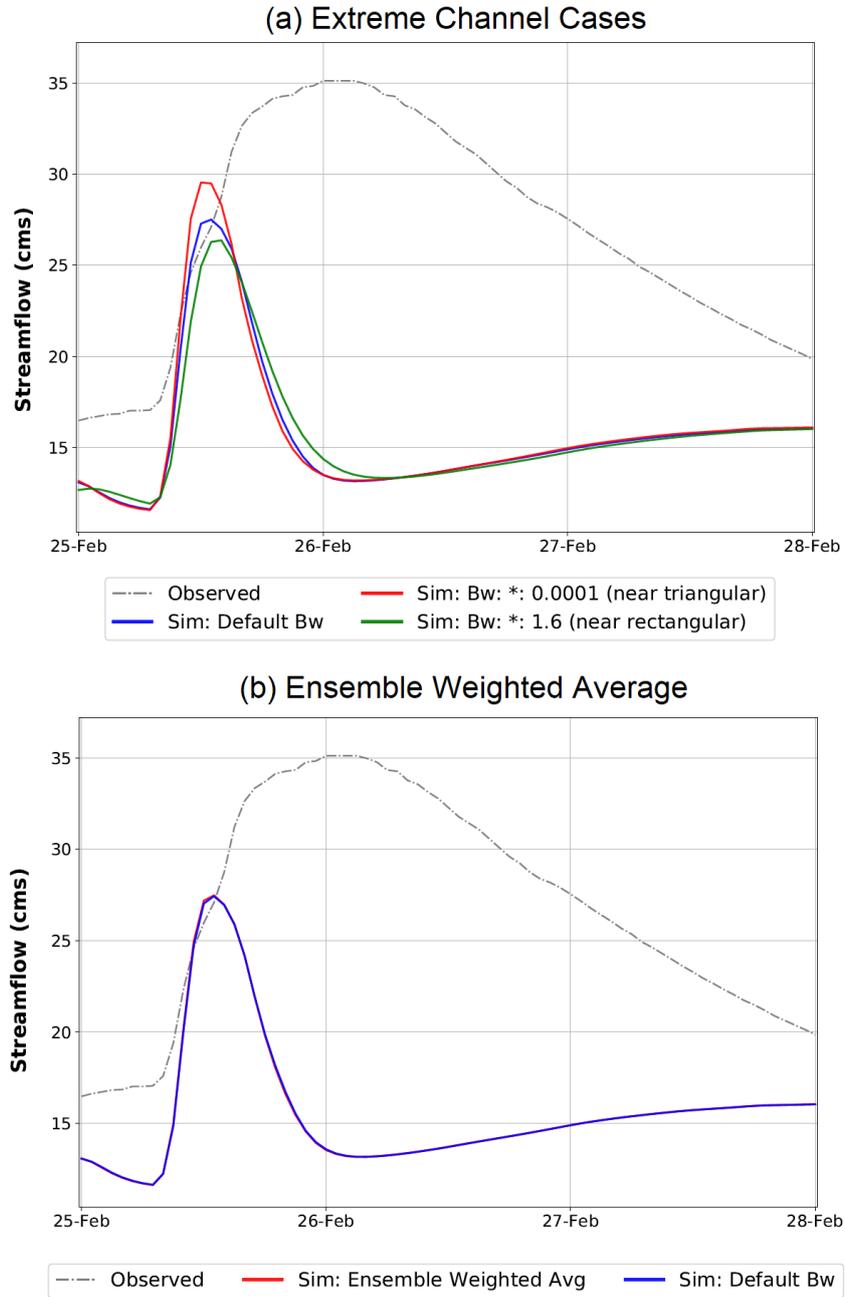


Figure 2. Streamflow observations and NWM simulations at the USGS station 01447720, (a) Extreme Channel Cases (b) Ensemble Weighted Average. (Focused on a 3-day subset of simulation period to better show differences between simulations)

4.2 DJS Benchmarking

A Dell Precision T7610 workstation housing two ten core Intel® Xeon® E5-2670 v2 processors clocked at 2.50 GHz and 64 GB of DDR3 RAM operating on version 18.04 LTS Bionic of Ubuntu was used to perform benchmarking. Several Docker containers and MPI process combinations were formulated to determine the scalability and robustness of our framework, based on runtime. To limit influence of potential bottlenecks introduced by Docker’s resource management system, each container was given access to all theoretical CPU threads, in our case

40. Table 1 outlines the aforementioned combinations. Although data read and write time should not contribute greatly to run time, it should be noted a Seagate 1TB external hard drive using the ext4 filesystem and USB 3.0 was used for all simulations. Each simulation consisted of a 31-day run period, using NLDAS forcing and restarts provided by a previous model spin-up simulation. The default Route_link.nc domain file, which contains all channel-based parameters, was used in place of alternative perturbed files.

Table 1. *DJS benchmark cases simulated for a one-month period with NLDAS forcings*

Active Containers Count	1	1	2	2	4	4	8	8	12	12	12 total limited to 6 containers
MPI Process Count	4	8	4	8	4	8	4	5	3	4	4

Figure 3 displays runtime results for each simulation, where the container number was equal to the number of NWM models running concurrently. Results of these simulations revealed three things. First, a hardware plateau, once the number of active containers multiplied by the number of MPI processes became greater than the theoretical thread count. This runtime hindrance can be observed in cases [8, 4] and [8, 5] in figure 3. The three right-most cases in figure 3, [12, 3], [6, 4] and [12, 4] reveal this point most clearly. Case [6, 4] differs from all cases in that 12 NWM simulations were completed, however the number of containers running concurrently was limited to six containers running four MPI processes. The importance of benchmarking a given system is shown in case [12, 3] which performed on par with all other cases running a lower number of containers, and adhered to the active container times MPI process rule. Second, the input and output runtime effects caused by writing each hourly land surface model output to disk were most clear in the [1, 4] and [12, 3] cases, where runtimes of one simulation were marginally different despite vastly different numbers of concurrent simulations. Third, and most important, the robust scalability of DJS was displayed in these tests. Specifically, when the number of active containers times the number of MPI processes remained at or below the maximum number of CPU threads, the runtime between a single model run and, in our case, twelve model runs remained similar.

5. Conclusion

This research set out to create a scalable, automated framework capable of running numerous NWM simulations with different channel parameters simultaneously building and utilizing open source programming languages. Our hope was to benefit the greater WRF-Hydro/NWM community. Our tools; the perturbation engine, DJS, and ensemble generator, all exist as modular units that when used together, have the potential of asking countless scientific questions. Despite our limited findings regarding our test case, Pocono, PA, we believe perturbing channel parameters on a calibrated domain will produce findings that will be much more compelling. We hope the deliverables offered in this report lay the foundation for both veteran and novice modelers to contribute to, and improve the nation's hydrological model.

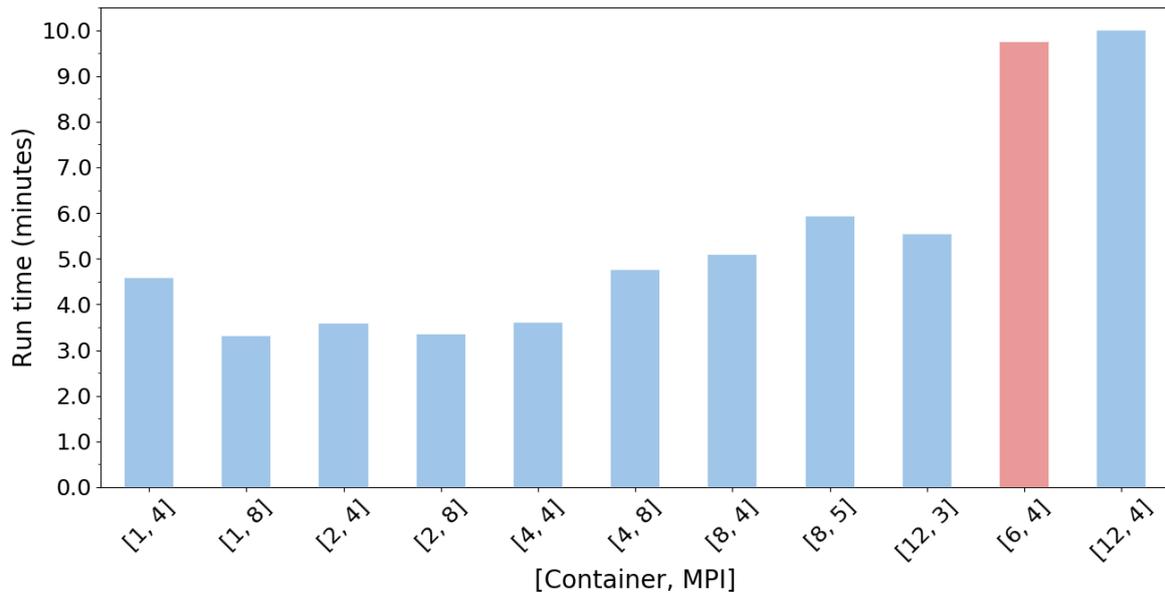


Figure 3. Runtimes for benchmarking Container-MPI combinations

Supplementary Materials: GitHub, Dockerhub, and Hydroshare repository of models and codes developed for this project are available online at: <https://github.com/aaraney/NWM-Docker-Ensemble-Framework>, <http://Hub.docker.com/u/aaraney>, and <https://www.hydroshare.org/user/4606/>, respectively.

References

1. Castronova, A.; Tijerina, D. (2019, July 19). CUAHSI/nwm_subsetting: CUAHSI Subsetter (Version 0.2). Zenodo. <http://doi.org/10.5281/zenodo.3343763>
2. Blackburn-Lynch, W.; Agouridis, C.T.; Barton, C.D. Development of Regional Curves for Hydrologic Landscape Regions (HLR) in the Contiguous United States. *JAWRA* **2017**, *53*(4), 903-928. 10.1111/1752-1688.12540, Available online: <https://doi.org/10.1111/1752-1688.12540>
3. Xia, Y., et al., NCEP/EMC (2009), NLDAS Primary Forcing Data L4 Hourly 0.125 x 0.125 degree V002, Edited by David Mocko, NASA/GSFC/HSL, Greenbelt, Maryland, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed: July, 2019, [10.5067/6J5LHHOHZHN4](https://doi.org/10.5067/6J5LHHOHZHN4)
4. McKay, M. D., Beckman, R. J.; Conover, W. J. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, **2000**, *42*(1), 55-61. 10.1080/00401706.2000.10485979, Available online: <https://doi.org/10.1080/00401706.2000.10485979>
5. Alpine Linux, <https://alpinelinux.org/>

Chapter 6

A Visualization Workflow for Quantifying Parameter Sensitivities to Uncertainties for Hydrologic Models

Kyla Semmendinger¹ and Catherine Finkenbiner²

¹ Department of Biological and Environmental Engineering, Cornell University; kts48@cornell.edu

² Department of Biological and Ecological Engineering, Oregon State University;
finkenbc@oregonstate.edu

Summer Institute Theme Advisors: David Blodgett, United States Geological Survey, dblodgett@usgs.gov; Kyle Mandli, Columbia University, kyle.mandli@columbia.edu

Academic Advisors: Scott Steinschneider, *Cornell University*; Stephen P. Good, *Oregon State University*

Abstract: Due to factors such as data availability and physical attributes, models require modular components for representing individual hydrologic processes, from regional to continental scales. Therefore, the need for a common procedure within the hydrologic field to evaluate model output based on parameter sensitivities and uncertainties compared to performance metrics (i.e. objective functions) is evident. We developed a reproducible workflow for evaluating hydrologic models, to objectively evaluate model output as a function of parameter choice, using both numerical and visualization techniques. The workflow was implemented using case studies provided by Fellows of the National Water Center Innovators Program: Summer Institute 2019. The resulting output can be reproduced and visualized using Jupyter notebooks within a community GitHub code repository (https://github.com/ksemmendinger/Hydro_Parameter_Sensitivity_Visuals). Model parameter sensitivity was evaluated using various global sensitivity indices (i.e. Sobol, Delta, R^2) and Bayesian theory. Uncertainty in parameter spaces was quantified, to highlight the impact of unreliable input data on model output. Model parameter sensitivities and uncertainties were evaluated numerically and visually to provide a comprehensive outlook on their model output impact. For each case study, we provide a summary and an interpretation of workflow results. Our workflow can be integrated into hydrologic modeling frameworks for objective modular model and parameter set evaluations, based on a data-driven approach for model selection.

1. Motivation

Conceptual and mathematical models are used to represent complex and unconstrained hydrologic environments [1]. Difficulties arise in model performance evaluation due to multiple interpretations, differing mathematical descriptions and uncertainty present in the input and parameter datasets [2, 3]. Parameter values often are not a direct physical interpretation and rarely can be measured in the field. Therefore, models are calibrated using objective functions (e.g. Nash-Sutcliffe Efficiency [NSE], root-mean-square error [RMSE], peak discharge) and the 'best' model structure and parameter sets are often selected [4]. To effectively model hydrology it is necessary to understand how model output is impacted by each parameter, and to focus model calibration on decreasing uncertainties of more influential parameters. The hydrologic modeling community needs common frameworks organizing development of an open source multi-criteria approach that is reproducible, adaptable and transferable, for evaluating model parameter

sensitivities and uncertainties to advance computational hydrology across disciplines and platforms [2,4,5].

2. Objectives and Scope

The study objective was to develop an adaptable and reproducible workflow for evaluating hydrologic model parameter sensitivities and uncertainties. To achieve the objective the workflow needed to impartially evaluate model output as a function of parameter choice, using both numerical and visualization techniques. This paper details workflow implementation and sensitivity analyses on eleven input parameters for Storm Water Management Model (SWMM) simulated streamflow. Visualization techniques are included in this paper for one objective function; however, visuals for five additional objective functions are available via our GitHub repository (refer to Supplementary Materials). Additional case studies, provided by Fellows of the National Water Center Innovators Program: Summer Institute 2019, also are included within our GitHub repository. All Python/R scripts are included in the GitHub repository along with an executable Jupyter Notebook for each case study. The range in model outputs and objective functions for each case study directly highlights the reproducibility of our workflow, as each case study has different parameter sets, model configurations and outputs. The workflow was developed for easy integration into hydrologic modeling frameworks (e.g. evaluation of the NWM) for objective modular model and parameter scheme evaluations, based on a data-driven model selection approach.

3. Previous Studies

Previous studies have called for adaptation of reproducibility and transparency within hydrology to add credibility to scientific claims. Research should be rooted in evidence supporting the methodology applied, data acquired, data analysis and interpretation of results and outcomes [6]. Advancements in how we disseminate information should be incorporated into hydrologic modeling frameworks, by including tools that will document and publish computational workflow and all associated digital objects [5, 7]. It is assumed reproducibility and scientific provenance, a digital record of data and methods used to achieve results, will be key review criterion for future geoscience publications [5]. Gil et al. [5] introduces the Geoscience Paper of the Future concept, with a best practice requiring authors to make data, software and methods openly accessible, citable, and well documented. Challenges to reproducibility include technical barriers, limited documentation of the research to be replicated, and potentially complex requirements for software packaging, installation and execution [7].

While sensitivity analyses have been implemented across various hydrologic models, a comprehensive and adaptable methodology has yet to be adopted in common practice. Many packages are available for calculating objective functions (e.g. the R package hydroGOF, <https://www.rdocumentation.org/packages/hydroGOF>) and implementing sensitivity analyses (e.g. the python library SALib, <https://salib.readthedocs.io/en/latest/>), but rarely do these packages communicate with each other. Future research should focus on integrating existing coding packages within universal workflows, for robust hydrologic model analyses.

4. Methodology

The workflow was developed to be adaptable and reproducible for applications across various hydrologic models, and their respective outputs or objective functions (i.e. performance metrics) of interest. The workflow was organized to be easily integrated with the National Water Center (NWC) Water Resources Evaluation Service (WRES, v1.8) tool that calculates metrics using hydrologic model outputs and United States Geological Survey (USGS) stream gauge data.

WRES is under development, backed by .xml files, and calculates metrics from single-value forecasts, ensemble forecasts and observations/simulations. The WRES-calculated metrics can be used as objective functions for multiple parameter sensitivity and uncertainty calculations performed in our workflow. All analyses incorporated into the resulting workflow should encompass any user-defined objective function and allow for testing new hydrologic evaluation metrics. The remainder of this section explains different sensitivity and uncertainty methods incorporated into the workflow, based on user-defined or calculated objective functions.

4.1 Parameter Sensitivity and Uncertainty Analyses

One method integrated into the workflow was Approximate Bayesian Computation (ABC), which represents the combination of model parameter values that maximize the probability of representing the observed data. ABC applies Bayesian theory to parameter spaces to estimate posterior distributions of model parameters. ABC is advantageous because it bypasses calculating the likelihood function, by using model simulations compared to observed data [3,8-10]. The steps to compute ABC are as follows.

- 1) Calculate observed data's statistics (e.g. mean, standard deviation) and choose model-specific objective functions (e.g. NSE);
- 2) Assume a uniform sampling interval for the parameter space. Draw a total of n parameters from prior data, and simulate the model for each parameter point, producing n sequences of simulated data.
- 3) Calculate objective functions for each simulated data sequence
- 4) Determine the distance between observed and simulated transition frequencies for all parameter points. Remove parameter points beyond a user-specified tolerance interval (e.g. $NSE \geq 0.0$), to approximate samples from the posterior distribution.
- 5) Estimate posterior distribution with the parameter points within the tolerance interval [9, 10].

Three sensitivity analyses are incorporated into the workflow: a variance-based sensitivity analysis; a moment-independent sensitivity analysis, and an ordinary least squares regression. The Sobol method [11] is a variance-based global sensitivity analysis that yields first-order, second-order, and total-order sensitivity indices. Sobol's method can effectively handle nonlinear responses, and it measures effects of interactions within non-additive systems. It decomposes the variance of model output into fractions that can be attributed to inputs or sets of inputs. The first-order sensitivity index (i.e. main effect index) quantifies parameter impact on model output variance by averaging variations in other input parameters. The second-order sensitivity index decomposes model variance by parameter interactions with one another. The total-order sensitivity index (i.e. total effect index) measures the contribution each parameter had on model output across the first-order index and all higher-order indices. Our workflow employs the Saltelli scheme [12, 13], which allows for the calculation of the first-order, second-order, and total-order sensitivity indices with fewer model runs than a traditional approach. However, since the calculation of Sobol second-order indices requires $N * (2K + 2)$ model runs (where N is preset model runs and K is number of parameters) [13], some models may be too computationally expensive to run with multiple input parameters. Consequently, we included two additional sensitivity analyses that have no minimum number of model runs.

The Delta index [14-16] is a moment-independent global sensitivity analysis. While less robust than indices returned by variance-based sensitivity analysis, moment-independent sensitivity analysis is a popular technique, due to its computational efficiency and insensitivity to dependent parameters [17]. The Delta sensitivity analysis searches for parameters with the greatest impact on the probability density function of model output. Delta indices capture non-linear and non-monotonic parameter-output dynamics. Finally, the ordinary least squares (OLS) regression

yields an R^2 coefficient that quantifies the linear effects of model input parameters on model output variance. OLS regressions have long been employed for model sensitivity analyses, and assume explicit interaction between model output and any given parameter [18-20].

4.2 Visualization Workflow

The reproducible and adaptable visualization workflow (Figure 1) handles input files of parameter sets, model simulations, observed data, and objective functions (e.g. NSE). The necessary input formats are a: 1) $[n \times K]$ matrix where n rows are the models/indexes and m columns are the K parameters for each model configuration; 2) $[n \times t]$ matrix where n rows are the models/indexes and t columns are the model simulations at each temporal interval; 3) $[1 \times t]$ matrix where t is the observed data at each temporal interval, and 4) $[n \times S]$ matrix where n rows are the model/index and S columns are the S number of objective functions of interest. Coding packages are available for calculating common objective functions (see Section 3) and we used R package hydroGOF. Additionally, the WRES tool can be used to calculate objective functions for the various model parameter sensitivity and uncertainty visualization approaches.

The workflow is modular, allowing users to run only scripts applicable to their parameter sets, model outputs, and objective functions. Moving from left to right across the different modular components in Figure 1, the ABC analysis provides a priori and posteriori parameter distribution plots. The workflow saves text files of all models that have parameter sets within a user specified tolerance interval (e.g. $NSE \geq 0.0$) and high-resolution figures of the resulting histograms, cumulative distribution functions (CDFs) and probability density functions (pdfs). ABC output is intended to highlight parameter distributions as one function/role of the objective functions of interest. The user can alter the code to consider any objective function and tolerance interval. To evaluate model performance, the workflow plots observation datasets (i.e. historical time series) and magnitude percentile plots, comparing frequency distributions of the modeled versus observed data magnitude. The next workflow module is the delta and OLS (R^2) sensitivity analyses, which produce scatterplots of the various user-specified objective functions versus the different model parameters. This allows the user to visualize single parameter impacts on the objective functions. Next, portrait and spider plots summarize OLS and Delta sensitivity analysis of each parameter, for single-to-multiple objective functions. A Sobol second-order, and total-order sensitivity analysis is specific for Saltelli parameter sampling schemes, and may not be appropriate for all models due to computational limitations or large parameter spaces. Sobol provides a global sensitivity analysis and is advantageous for observing first-order, second-order and total-order parameter interactions and their impacts on objective functions.

5. Results

5.1 Case Study 0: SWMM Streamflow Simulation of Fall Creek Watershed, NY

The Fall Creek watershed is a 325 km² basin, located near Ithaca, New York. The Fall Creek watershed was selected for Case Study 0 domain area because all forcing data and parameter ranges were readily available. SWMM was selected as the hydrological model because it is computationally inexpensive and has the ability to complete a model run in approximately four seconds. Additional information regarding SWMM models and their configurations is available in our GitHub repository. To complete Sobol, Delta, and OLS sensitivity analyses on SWMM simulated streamflow to input parameters, we employed the Saltelli sampling scheme to create 24,000 unique parameter sets. Eleven parameters, shown in Table 1, were sampled for the sensitivity analysis.

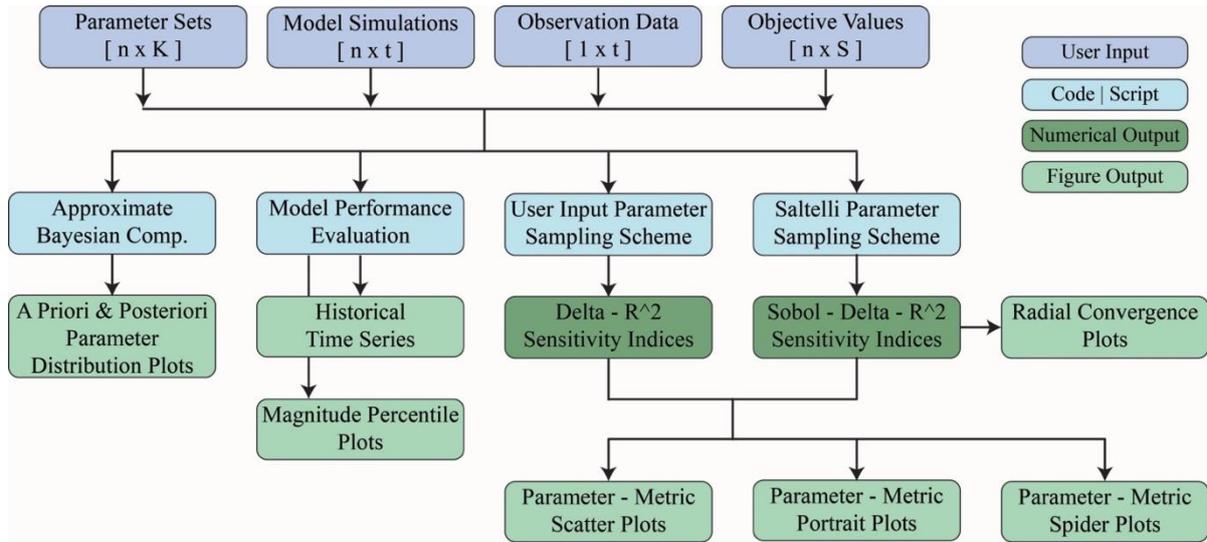


Figure 1. Workflow illustrating the various datasets and analysis for visualization of parameter sensitivities and uncertainties

Table 1. The eleven parameters sampled in the SWMM sensitivity analysis

Parameter	Description
w	Basin width
n_imperv	Impervious cover Manning's roughness coefficient
n_perv	Pervious cover Manning's roughness coefficient
s_imperv	Impervious cover storage volume
s_perv	Pervious cover storage volume
k_sat	Hydraulic conductivity
per_routed	Percent routed to pervious cover
cmelt	Snow melt coefficient
Tb	Base temperature
A1	Groundwater flow coefficient
B1	Groundwater flow exponent

Each parameter set was run through the SWMM from January 1, 2013 through June 30, 2013, and the simulated streamflow time series was extracted. Simulated streamflow was compared to USGS stream gauge data (<https://waterdata.usgs.gov/nwis/uv?04234000>). Six objective functions were calculated and used as model performance metrics using the R hydroGOF package [21]: mean absolute error (MAE); mean error (ME); mean squared error (MSE); Nash-Sutcliff efficiency (NSE); percent bias (pbias), and root mean squared error (RMSE). Section 5.2 visualized the complete workflow with NSE as the objective function. Refer to the GitHub repository (link at end of chapter) for a complete analysis across all objective functions and a replication of all visualizations.

5.2 Workflow Visualizations and Outcomes

Workflow numeric output includes Sobol first-order, second-order, and total-order sensitivity indices, as well as Delta and R^2 sensitivity indices. These files are saved in the output directory,

together with the figures shown below. The numeric output files can be found in our GitHub repository. In a numerical sensitivity analysis, a larger sensitivity index implies greater influence of input parameter over model output. For the NSE objective function of Case Study 0, the Sobol, Delta, and R^2 sensitivity indices, together with their corresponding confidence intervals, are shown in Table 2.

Table 2. The Sobol first-order (S1), S1 Confidence Interval (CI), Sobol total-order (ST), ST CI, Delta, Delta CI, and R^2 sensitivity indices for the NSE objective function

Parameter	S1	S1 CI	ST	ST CI	Delta	Delta CI	R^2
w	0.00932451	0.01024303	0.01116858	0.00425999	0.05726384	0.003195	0.38168023
n_imperv	0	0.01407207	0.02102163	0.00297597	0.06736131	0.00219677	0.29685502
n_perv	0.00066735	0.0118365	0.02635968	0.0154946	0.06141614	0.00468541	0.30461684
s_imperv	0.00088842	0.00138211	0.00019754	4.54E-05	0.05762623	0.00354774	0.30196962
s_perv	0.00441332	0.00388486	0.0022433	0.00165579	0.05347048	0.00343368	0.30710408
k_sat	0.05411055	0.03422148	0.08045932	0.04190596	0.0706289	0.00597163	0.27664867
per_routed	0.02830262	0.02034452	0.04283429	0.00740148	0.23760169	0.00545487	0.25279263
cmelt	0.02640971	0.03276811	0.07950454	0.03096019	0.06506388	0.00326249	0.34164728
Tb	0.02902564	0.02758362	0.07529993	0.02679876	0.05676743	0.00457267	0.01074236
A1	0.48937156	0.06469653	0.57903878	0.07151498	0.29388964	0.00737714	0.66406945
B1	0.2218755	0.05058228	0.26088142	0.03282531	0.16087131	0.00373687	0.27357454

Since the objective of this study was to visualize model parameter sensitivity, in this section we describe the figures created to visualize model output frequency, distribution, and sensitivity indices for each SWMM input parameter and the NSE objective function. We supply some inferences and figures interpretations, but model performance evaluation was not the focus here.

Comprehensive sensitivity analyses require large numbers of model runs. This results in overlapping plots, often with thousands of time series. The overlapping historical and simulation data time series provides little insight into the frequency of various model output values. Figure 2 shows a historical data time series together with a frequency magnitude percentile curve. Magnitude percentile curves allow interpretation of more- and less-frequent observation values by grouping individual time series into percentile ranges. As shown in Figure 2 for Case Study 0, the majority of simulated streamflow values fall below the historical data, implying SWMM tends to underestimate streamflow, especially in higher flow (80th-100th percentile) events.

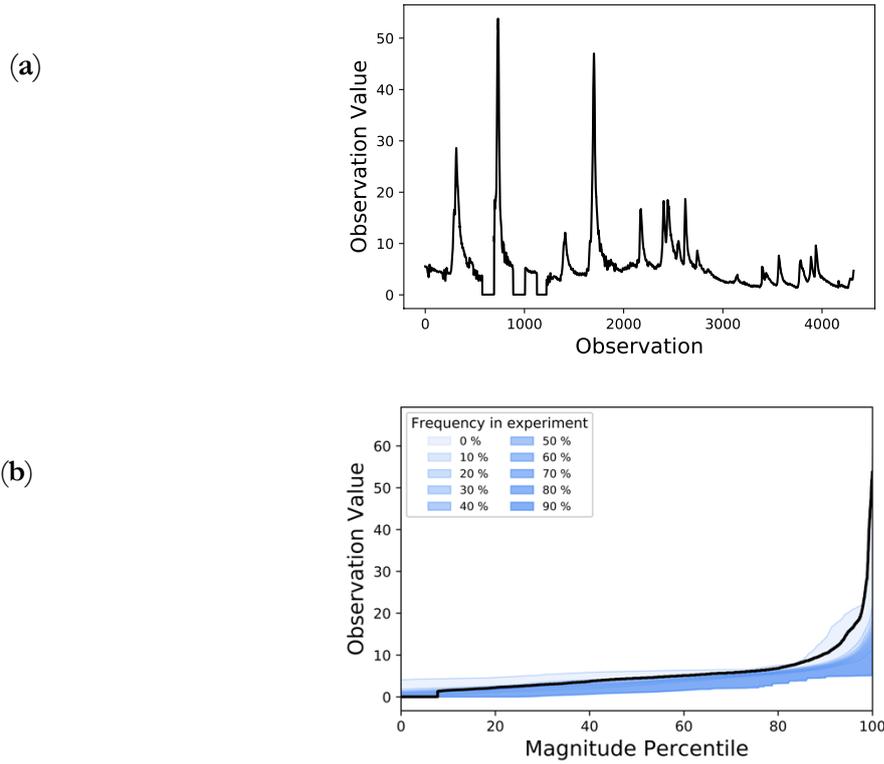
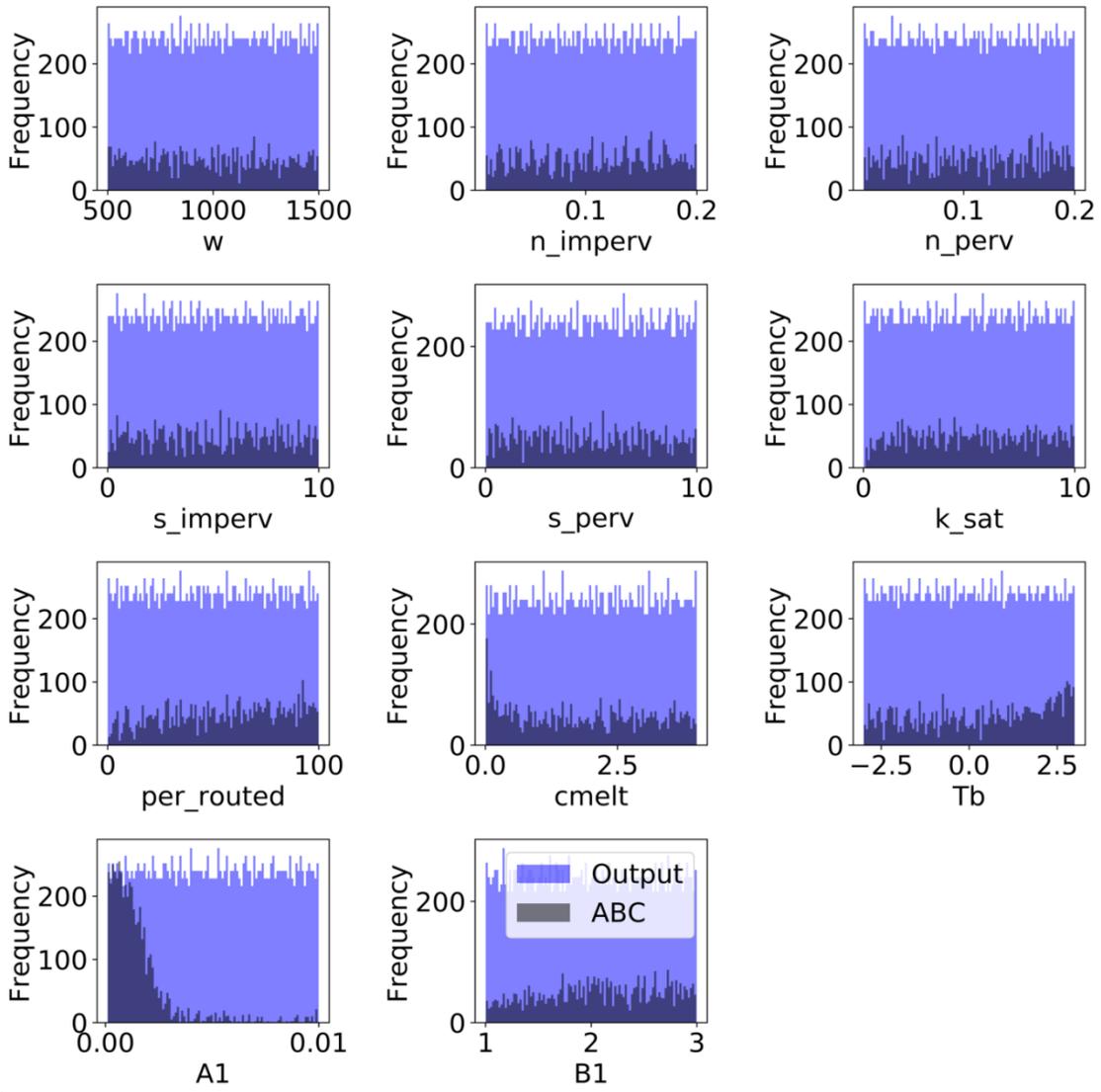


Figure 2. (a) A time series of historical streamflow data (cms/hour) (b) along with a magnitude percentile plot showing the spread and frequency of SWMM simulated streamflow compared to the historical observations (cms/hour).

ABC was implemented with a tolerance specified at $NSE \geq 0.0$ and workflow results were visualized as histograms, CDFs and pdfs. The histograms illustrate the frequency distributions of parameter sets constrained by ABC specified tolerances (Figure 3[a]). Changing the tolerance will adjust the frequency distributions of parameters as a function of user-specified objective functions. CDFs (see GitHub repository) and pdfs (Figure 3[b]) provide the user with additional visualizations of constrained parameter sets based on ABC. Based on Figure 3, the SWMM model configurations with the most positive NSE values have A1 parameters equaling values near zero. One might conclude the A1 parameter should be sampled from a constrained range of values, skewed toward zero. Constraining A1 may adjust the histograms, CDFs and pdfs of other parameters depending on the user-specified objective function and tolerance interval. However, it is impossible to conclude this without understanding the relationships and dependencies between the various model parameters. Additional sensitivity analyses are required to adequately describe the parameter interactions within hydrologic models.



(a)

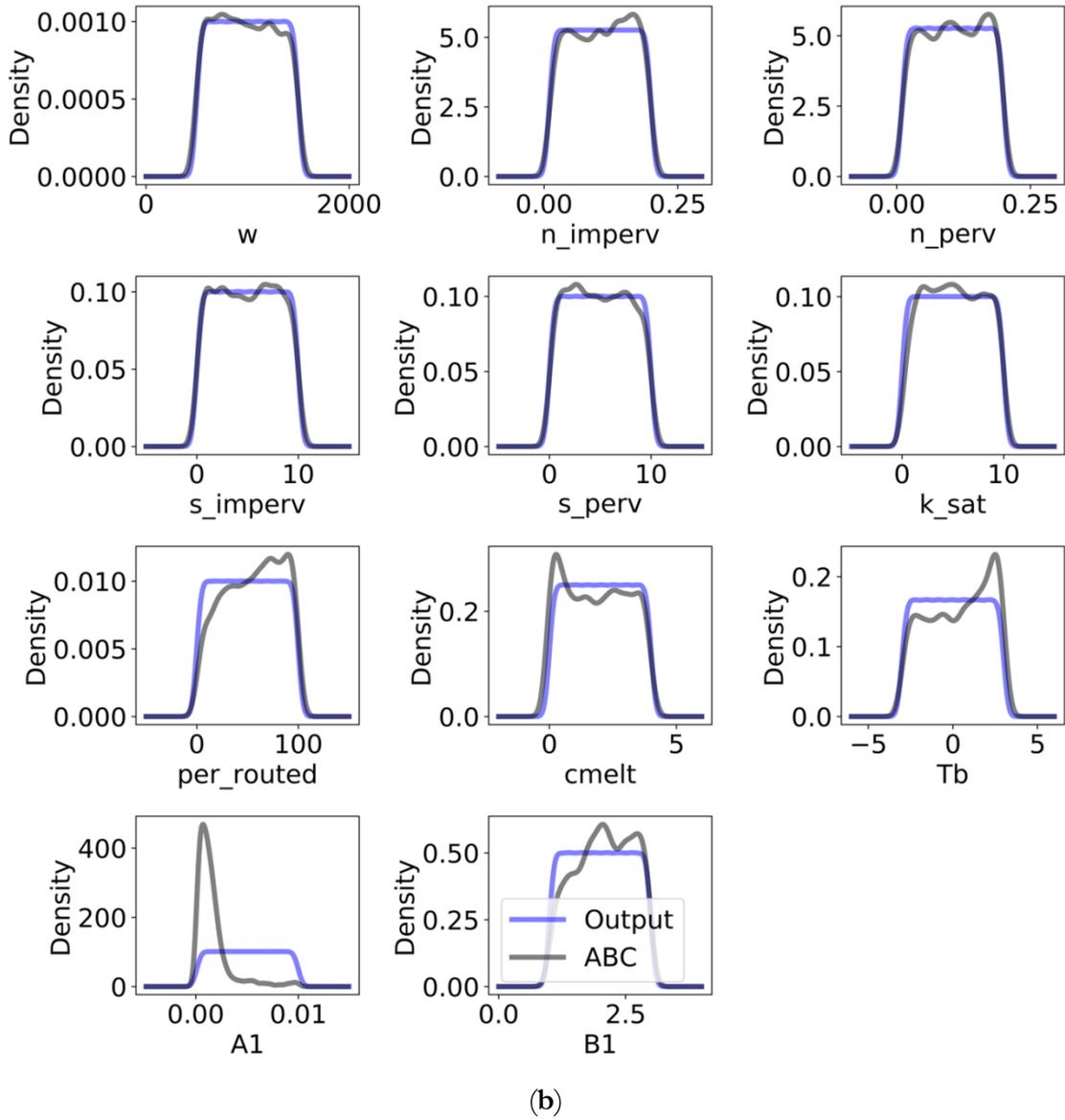


Figure 3. (a) Histograms (bins = 100) and (b) pdfs illustrating frequency of the modeled discharge (Output) versus the different SWMM model parameters. Approximate Bayesian computation (ABC) applied for $NSE \geq 0.0$ is shown.

Figure 4 illustrates scatterplots of objective function values versus parameter values. A correlation between the two directly corresponded to the linear relationship indicated by the R^2 value from the OLS sensitivity analysis. For Cast Study 0, it can be inferred that A1 and B1 groundwater coefficients had the strongest linear impacts on the objective function (i.e. NSE). A spider plot, shown in Figure 5, illustrates the difference in sensitivity index magnitudes across model parameters for each objective function and sensitivity analysis. The results reinforce OLS regression results of A1 and B1 groundwater coefficients as the most impactful on model output.

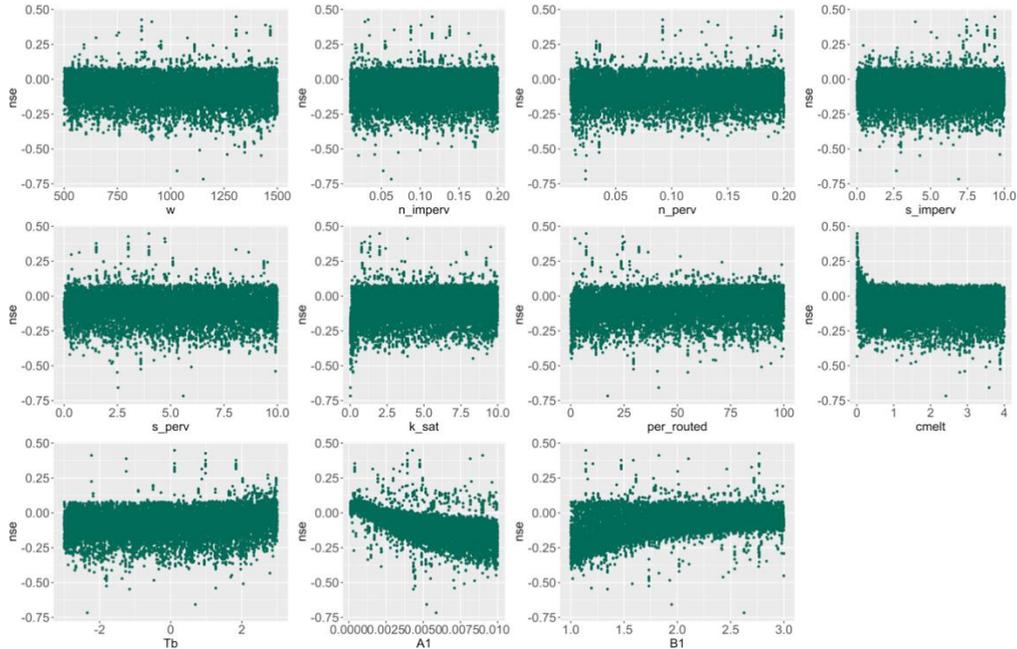


Figure 4. Scatterplots highlighting linear influences of input parameters on NSE objective function

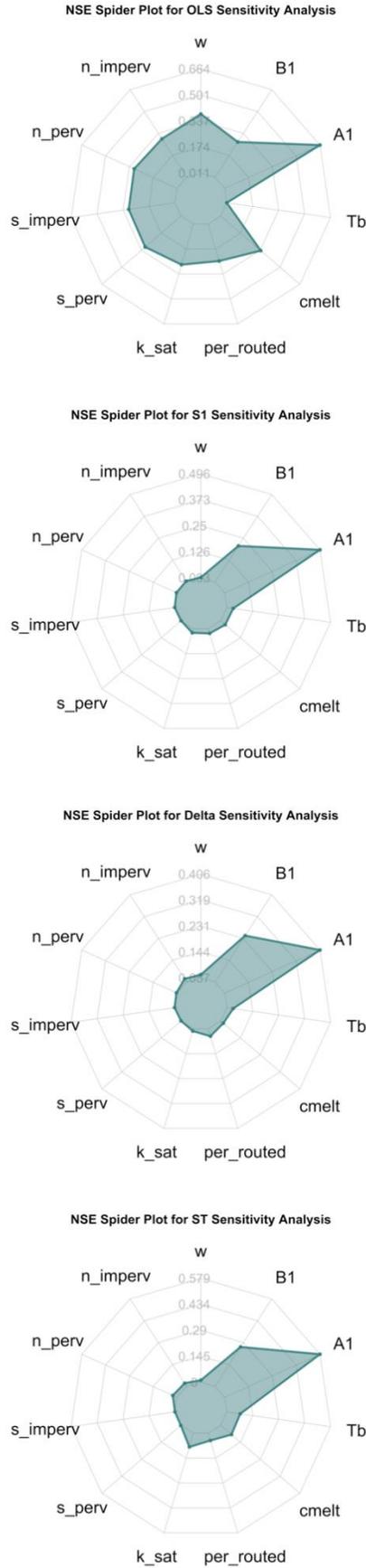


Figure 5. Spider plot illustrating magnitudes of the OLS, Sobol (S1, ST) and delta sensitivity indices for NSE and each SWMM input parameter

A radial convergence plot (RCP), shown in Figure 6, was produced from the Sobol sensitivity analysis. RCPs visualize not only first-order and total-order sensitivity for each parameter, but also second-order interactions between model parameters. Large first-order indices indicate the parameter is influential over model output or objective function. Small first-order indices indicate the parameter is unimportant in model output, thus allowing model simplification. Second-order indices describe how much interactions between two parameters influence model output, and total-order indices describe how much the sum of all parameter interactions influences model output.

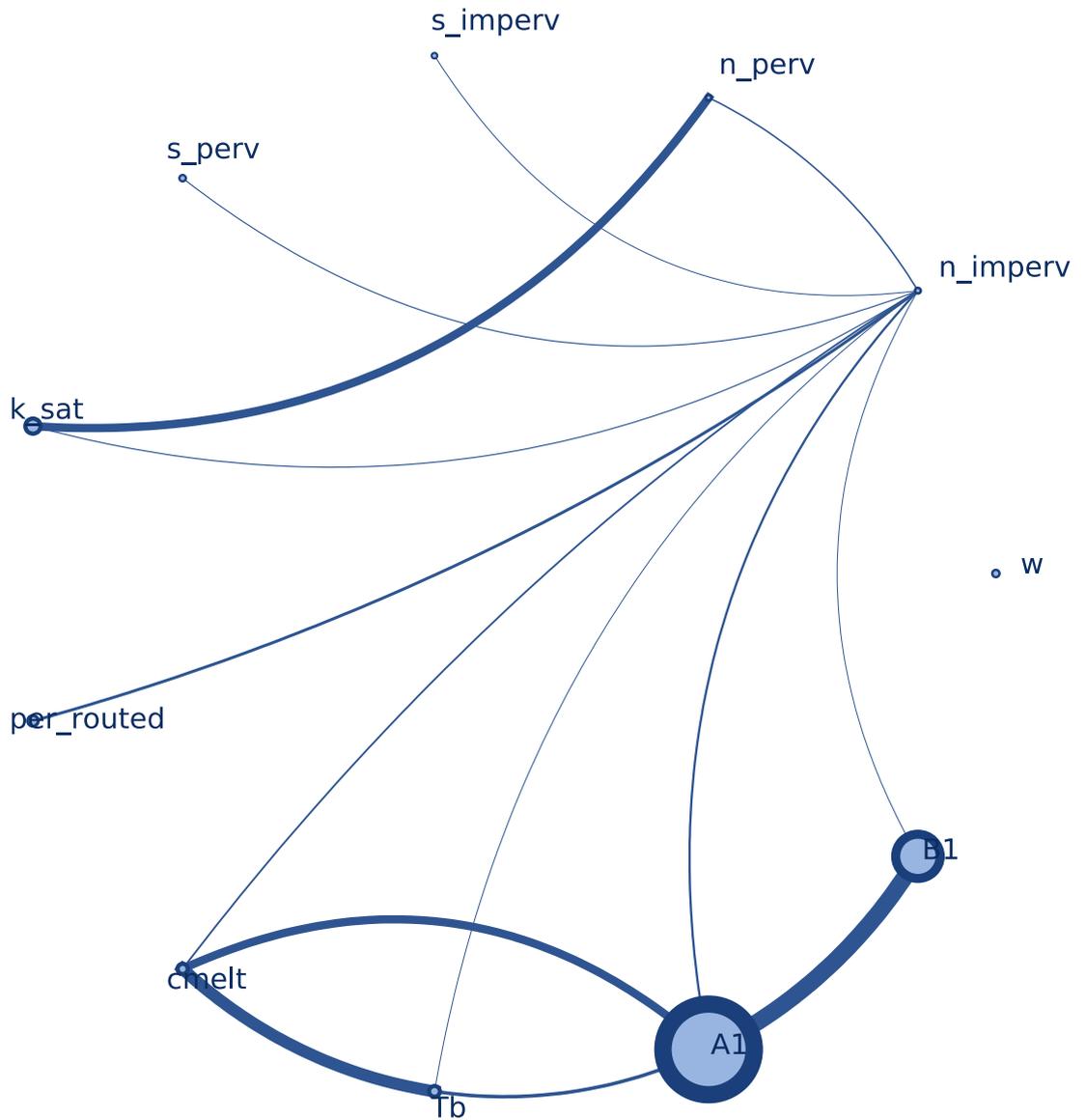
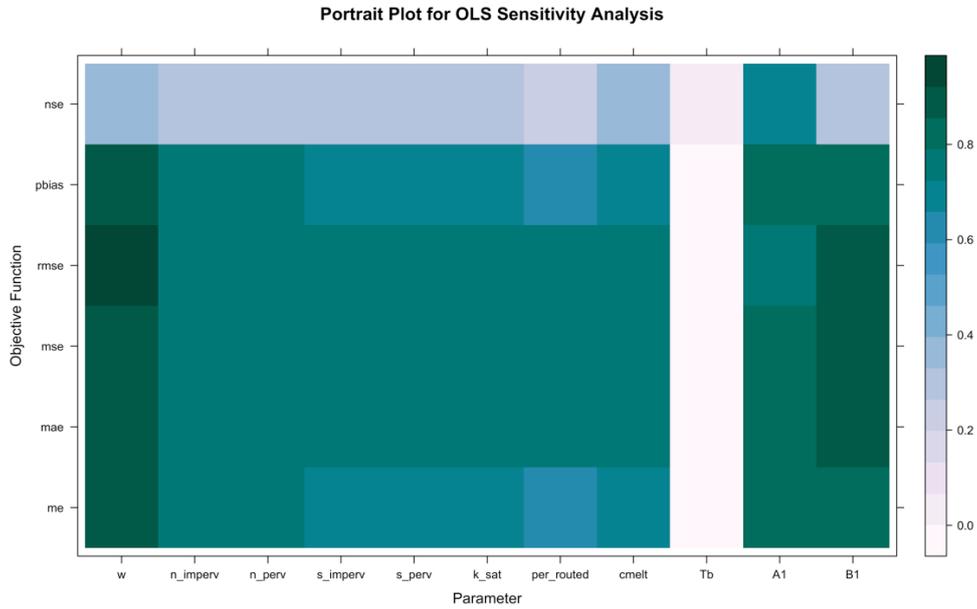
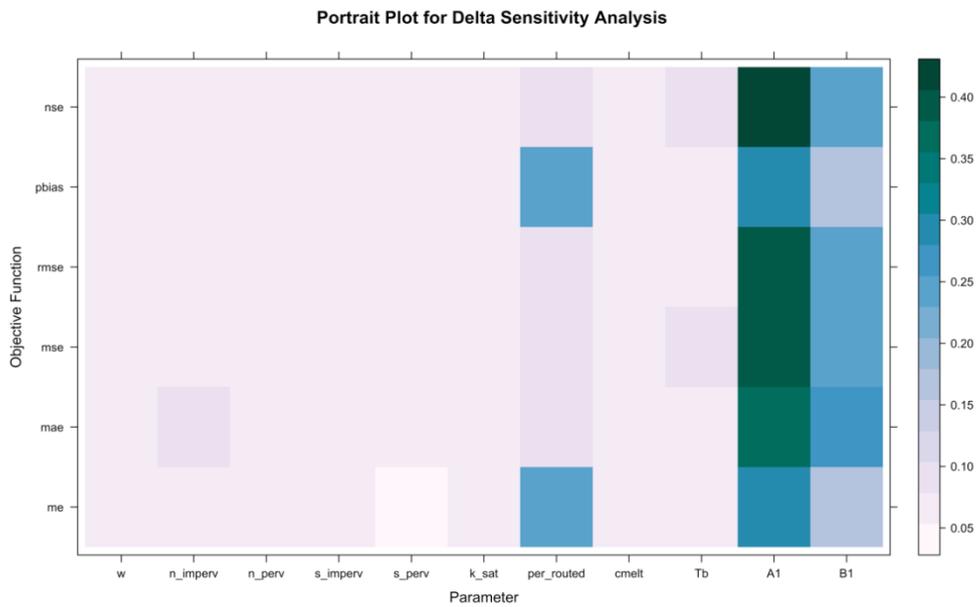


Figure 6. A radial convergence plot demonstrating the first-order (light blue circle), second-order (dark blue lines), and total-order (dark blue circle border) Sobol indices for NSE objective function

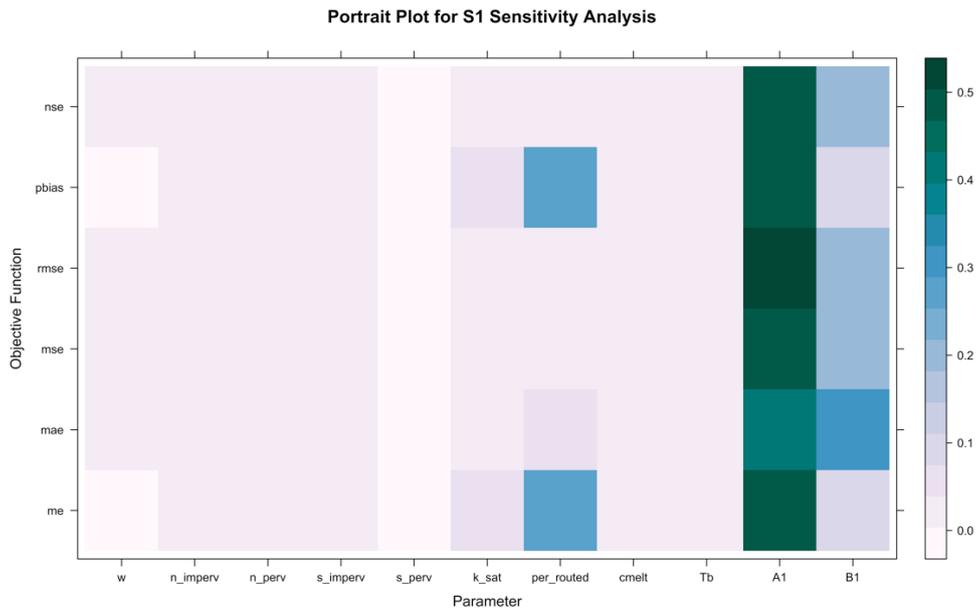
The following visuals, produced by the workflow, are portrait plots (Figure 7). Portrait plots allow direct comparison of all objective functions and parameters for each sensitivity analysis. In the plots below, darker green indicates higher objective function sensitivity to model parameter value. While several parameters seem to have linear impact on objective functions in the OLS regression, for Sobol and Delta sensitivity analyses, A1 and B1 groundwater parameters show the most significant influence over the various objective function values.



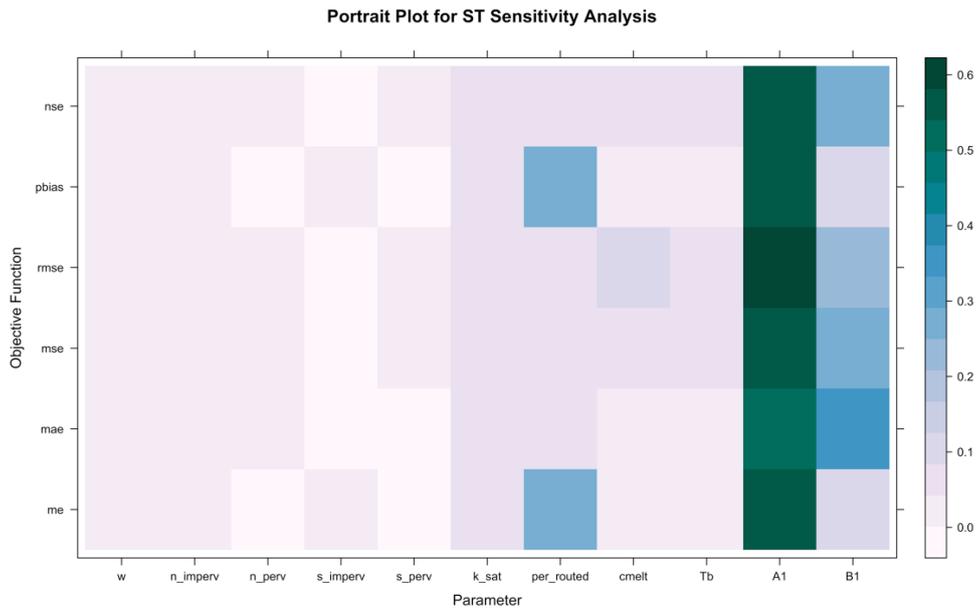
(a)



(b)



(c)



(d)

Figure 7. Portrait plots showing the (a) OLS, (b) Delta, (c) Sobol first-order, and (d) Sobol total-order sensitivity indices with respect to each combination of objective function and model parameter.

The workflow is organized modularly, to incorporate additional sensitivity analyses and visualization techniques. Future work could further develop the process to allow parameter sampling iteration based on the various sensitivity analyses results.

6. Conclusion

In this study we developed an easily adaptable and transferrable workflow, to visualize parameter sensitivities and uncertainties for hydrologic models. This workflow was implemented for multiple case studies, validating its applicability across several hydrologic research fields. Numeric and visual workflow output can aid hydrologic modelers with model selection, simplification, and parameter definition. Additionally, the workflow can be integrated with the existing model performance metric tool, WRES, used by the NWC. Due to the modular workflow setup there is potential to expand on sensitivity analyses and visualization techniques, based on user needs. We hope to see an adaptation of our workflow integrated into operational use for evaluation of the NWM, to meet NWC needs as they move toward a modular set-up of streamflow analysis across various temporal and spatial scales.

Supplementary Materials: Code repository and reproducible case study visualizations available at: https://github.com/ksemmendinger/Hydro_Parameter_Sensitivity_Visuals

References

1. Neuman, S. Maximum likelihood Bayesian averaging of uncertain model predictions, *Stoch. Environ. Res. Risk Assess.* **2003**, *17*, 291. doi:10.1007/s00477-003-0151-7
2. Hutton, C., T. Wagener, J. Freer, D. Han, C. Duffy, and B. Arheimer. Most computational hydrology is not reproducible, so is it really science? *Water Resour. Res.* **2016**, *52*, 7548–7555, doi:10.1002/2016WR019285.
3. Kavetski, D.; Kuczera, G.; Steward W.F. Bayesian analysis of input uncertainty in hydrological modeling: 1. Theory, *Water Resour. Res.* **2006**, *42*, W03407. doi:10.1029/2005WR004368
4. Wagner, T.; Boyle, D.P.; Lees, M.J.; Wheater, H.S.; Gupta, H.V.; Sorooshian, S. A framework for development and application of hydrological models, *Hydrol. Earth Sys. Sci.* **2001**, *5*(1), 13-26. doi: 10.5194/hess-5-13-2001
5. Gil, Y., David, C.H., Demir, I., Essawy, B.T., Fulweiler, R.W., Goodall, J.L., Karlstrom, L., Lee, H., Mills, H.J., Oh, J.H., Pierce, S.A., Pope, A., Tzeng, M.W., Villamizar, S.R., Yu, X. Toward the Geoscience Paper of the Future: Best practices for documenting and sharing research from data to software to provenance. *Earth Sp. Sci.* **2016**, *3*, 388-415. doi.org/10.1002/2015EA000136
6. Munafò, M.R.; Nosek, B.A.; Bishop, D.V.M.; Button, K.S.; Chambers, C.D.; Percie du Sert, N.; Simonsohn, U.; Wagenmakers, E.J.; Ware, J.J.; A. Ioannidis, J.P. A manifesto for reproducible science, *Nature Human Behav.* **2017**, *1*(21). doi:10.1038/s41562-016-0021
7. Essawy, B.T.; Goodall, J.L.; Zell, W.; Voce, D.; Morsy, M.M.; Sadler, J.; Yuan, Z.; Malik, T. Integrating scientific cyberinfrastructures to improve reproducibility in computational hydrology: Example for HydroShare and GeoTrust, *J. Environ. Modeling & Software* **2018**, *105*, 217-229. doi: 10.1016/j.envsoft.2018.03.025.
8. Engeland, K.; Gottschalk L. Bayesian estimation of parameters in regional hydrological model, *Hydrol. Earth Sys. Sci.* **2002**, *6*(5), 883-898. doi:10.5194/hess-6-883-2002
9. Sunnåker, M.; Busetto A.G.; Numminen E.; Corander J.; Foll M.; Dessimoz C. Approximate Bayesian Computation, *PLoS Comput. Biol.* **2013**, *9*(1). doi:10.1371/journal.pcbi.1002803
10. Vrugt, J.A.; Mojtaba, S. Toward diagnostic model calibration and evaluation: Approximate Bayesian computation, *Water Resour. Res.* **2013**, *49*, 4335-4345. doi:10.1002/wrcr.20354

11. Sobol, I.M. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Math. Comput. Simul.* **2001**, *55*, 271-280. doi.org/10.1016/S0378-4754(00)00270-6
12. Saltelli, A. Making best use of model evaluations to compute sensitivity indices. *Comput. Phys. Commun.* **2002**, *145*, 280-297. doi.org/10.1016/S0010-4655(02)00280-1
13. Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., Tarantola, S. Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Comput. Phys. Commun.* **2002**, *181*, 259-270. doi.org/10.1016/j.cpc.2009.09.018
14. Borgonovo, E. A new uncertainty importance measure. *Reliab. Eng. Syst. Saf.* **2007**, *92*, 771-784. doi.org/10.1016/j.res.2006.04.015
15. Borgonovo, E., Castaings, W., Tarantola, S. Model emulation and moment-independent sensitivity analysis: An application to environmental modelling. *Environ. Model. Softw.* **2012**, *34*, 105-115. doi.org/10.1016/j.envsoft.2011.06.006
16. Plischke, E., Borgonovo, E., Smith, C.L. Global sensitivity measures from given data. *Eur. J. Oper. Res.* **2013**, *226*, 536-550. doi.org/10.1016/j.ejor.2012.11.047
17. Wei, P., Lu, Z., Yuan, X. Monte Carlo simulation for moment-independent sensitivity analysis. *Reliab. Eng. Syst. Saf.* **2013**, *110*, 60-67. doi.org/10.1016/j.res.2012.09.005
18. Kleijnen, J.P.C. Sensitivity analysis and optimization of system dynamics models: Regression analysis and statistical design of experiments. *Syst. Dyn. Rev.* **1995**, *11*, 275-288. doi.org/10.1002/sdr.4260110403
19. Pannell, D.J. Sensitivity analysis of normative economic models: Theoretical framework and practical strategies, *Agr. Econ.* **1997**, *16*, 139-152. doi.org/10.1016/S0169-5150(96)01217-0
20. Zobitz, J.M., Keener, J.P., Schnyder, H., Bowling, D.R. Sensitivity analysis and quantification of uncertainty for isotopic mixing relationships in carbon cycle research. *Agric. For. Meteorol.* **2006**, *136*, 56-75. doi.org/10.1016/j.agrformet.2006.01.003
21. Mauricio Zambrano-Bigiarini. hydroGOF: Goodness-of-fit functions for comparison of and observed hydrological time series R package version 0.3-10. **2017**. <http://hzambran.github.io/hydroGOF/>. DOI:10.5281/zenodo.840087.