TECHNICAL REPORT 18

July 2022

OWP | OFFICE OF WATER PREDICTION

CUAHSI
universities allied for water research

National Water Center Innovators Program Summer Institute Report 2022

National Water Center Innovators Program
Summer Institute Report 2022

Editors:

Emily Deardorff

Arash Modaresi Rad

Jerad Bales

Trey Flowers

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

CUAHSI Technical Report No. 18

DOI: ____

Version 1.2

July 2022

Suggested Citation:

Deardorff, E., A. Modaresi Rad, et al. (2022). National Water Center Innovators Program - Summer Institute, CUAHSI Technical Report, HydroShare, http://www.hydroshare.org/resource/096e7badabb44c9f8c29751098f83afa

Contents

Preface

The first National Water Center Innovators Program: Summer Institute (NWC-SI) was held in the summer of 2015 by the National Water Center and Consortium of Universities for the Advancement of Hydrologic Science Inc. (CUAHSI). CUAHSI is a nonprofit organization of more than 130 member institutions that facilitates interdisciplinary water science and cooperation between researchers and students from different home institutions. The NWC-SI is held each year at the National Water Center (NWC), which is a National Oceanic and Atmospheric Administration (NOAA) facility that was established to enable cooperation across federal agencies to deliver high quality water forecasts and improve the nation's understanding of flood, drought, and water quantity issues at every level of government. The partnership between the NWC and CUAHSI has allowed the NWC-SI to facilitate collaboration and ideation across disciplines and backgrounds, helping developing scientists collaborate to create innovative research and sustainable solutions.

Throughout the program, graduate student fellows and academic faculty work with NWC staff and senior scientists to tackle challenging research projects related to modeling streamflow, predicting river geomorphology, and demystifying the relationship between oceans and rivers in coastal systems. This year, each project interfaced in some way with the Next Generation National Water Model (Nextgen) framework, which is a new, model-agnostic framework that is being developed for the U.S. National Water Model.

Each project revolves around a specific "theme" that has been established as important by the NWC and the theme leads. This year, there were three themes: "Next Generation hydrologic model framework," "Hydroinformatics," and "Coupling inland and coastal hydraulics." Each theme had two theme leads to guide the projects and two groups of students working on a project. In the first two weeks of the program, fellows were introduced to the themes, the NWC mission, and the ongoing goals of the Nextgen framework.

In the first week, fellows participated in training on how to use the Nextgen framework, using CUAHSI computational resources, collaborating through GitHub, and project management. Together, the theme leads and fellows got to know each other's skill sets and brainstormed project ideas. By the end of the first week, the fellows had formed their project groups and decided on a preliminary research plan. The fellows and theme leads spent the second week doing specialized training relating to their project. Fellows learned how to calibrate the GeoClaw coastal model and the National Water Model conceptual functional equivalent (CFE) model, how to use T-Route, advanced machine learning concepts, and how to run long short-term memory (LSTM) models in the NeuralHydrology framework. Over a seven-week program, the fellows worked hard in their groups to gather datasets, run models, and evaluate the results. The program culminated with capstone presentations, where students shared their findings to an audience of their advisors, NWC personnel, and other members of the CUAHSI community.

Fellows

The 7th NWC-SI Fellows class consisted of 22 graduate students, 6 M.S. students, and 16

PhD candidates from 22 universities across the United States. This cohort represented various academic backgrounds including civil engineering, geosensing systems engineering and sciences, geography, hydrology, and climate science. The diversity of fellows is a celebrated aspect of the NWC-SI that facilitates strong research projects and the inclusion of unique perspectives throughout the summer.

Themes and Theme Leads

The NWC-SI 2022 themes and theme leads were:

- The "Next Generation hydrologic model framework" theme, led by Fred Ogden (Chief Scientist for Water Prediction, NOAA NWC) and Jonathan Frame (Senior Hydrologist, Cloud to Street). Additional technical support was provided by Matthew Williamson (NOAA NWC) , Nels Frazier (NOAA NWC), Xia Feng (NOAA NWC), Luciana Kindl da Cunha (NOAA), and Austin Raney (CUAHSI).

- The "Hydroinformatics" theme was led by Chaopeng Shen (Pennsylvania State University) and Sagy Cohen (University of Alabama). Additional technical support was provided by Austin Raney (CUAHSI).

- The "Coupling inland and coastal hydraulics" theme was led by Celso Ferreira (George Mason University) and Kyle Mandli (Columbia University). Additional technical support was provided by Matthew Williamson (NOAA NWC), Nels Frazier(NOAA NWC), Xia Feng (NOAA NWC), Luciana Kindl da Cunha (NOAA), Austin Raney (CUAHSI).

Project Summaries

The following provides a brief summary of the 2022 NWC-SI projects, which are presented in more detail in Chapters 1-6 of this report.

1. Projects Related to the 'Next Generation hydrologic model framework' Theme

"Automated decision support for model selection in the Nextgen National Water Model" (Chapter 1) develops a novel method for selecting an optimal model (or models) for catchments in CONUS. Models selected for this study were Long short-term memory (LSTM) and National Water Model Conceptual Functional Equivalent (CFE). A machine learning approach (Random Forest Regressor) was used to predict the optimal choice of model for a given catchment. This offers additional capability to the Nextgen Framework as new models are incorporated into the Nextgen.

"Data assimilation of USGS streamflow data in the Nextgen Framework" (Chapter 2) explores other avenues to reduce model uncertainty and improve streamflow prediction by introducing a data assimilation (DA) method into the Nextgen framework. The team focused on producing simulations of the CFE model using the Ensemble Kalman Filter (EnKF) to assimilate the United States Geological Survey (USGS) streamflow observations.

2. Projects Related to the 'Hydroinformatics' Theme

"Data-driven approaches for estimating river channel geometry" (Chapter 3) explores the potential of using machine learning approaches and a large dataset of observed channel geometry to model channel geometry. Acoustic Doppler Current Profiler (ADCP) and National Hydrography Dataset Plus (NHDPlusV2) were used as point data alongside satellite imagery to establish a relationship between

the easily-available input data and the target variables (channel width and depth). A variety of model configurations showed promising results, including tree-based regressors and artificial neural networks..

"Large scale prediction of channel roughness coefficient using machine learning methods" (Chapter 4) introduces a machine learning method to estimate the roughness coefficient of Manning's equation using Acoustic Doppler Current Profiler (ADCP) and National Hydrography Dataset Plus (NHDPlusV2) datasets. The machine learning model showed the linkage between the influencing factors and channel roughness.

3. Projects Related to the 'Coupling inland and coastal hydraulics' Theme

"QuiCFIM, a quick GIS-based combined flood inundation mapping framework" (Chapter 5) introduces the Quick Combined Flood Inundation Mapping (QuiCFIM) approach, a simple but accurate method for flood depth and extent estimation. Their method combines inland and coastal flood inundation mapping (FIM). This approach uses HAND-SRC based inland FIM and c-HAND based coastal FIM, but the developed QuiCFIM is model agnostic and different inundation models can be tested in its framework. The framework was tested with the Hurricane Florence event in 2018 and the results were validated using ground observations and satellite imagery. QuiCFIM enhanced the results in transition zones and it was observed that the accuracy of this framework heavily relies on the quality of input FIM.

"Coupling coastal and hydrologic models through BMI and Next Generation National Water Model Framework in low gradient coastal regions of Galveston Bay, Texas, USA" (Chapter 6) tackles the daunting task of coupling coastal and hydrological models, specifically the GeoClaw costal model and CFE hydrological model. The team developed the first coastal BMI for the Nextgen framework and tested the coupled model using historical records of Hurricanes Harvey and Ike in a watershed that terminates in Galveston Bay. They demonstrated and documented the coupling process outside the Nextgen framework. They note the need for a plugin to have discharge as an input to GeoClaw.

Acknowledgements

A program of this magnitude and prominence requires the work of many hands. Initial organization, travel arrangements, pre-program webinars, and program leadership was provided by Jerad Bales, Veronica Sosa Gonzalez, Julia Masterman from CUAHSI. Trey Flowers, Fred Ogden, and Lauren Stewart of the NWC provided crucial assistance in coordinating the program from the NOAA side and providing both technical and administrative support to the program. Housing and on-campus support was provided by the University of Alabama office of Professional Development & Conference Services. Additional assistance with on-campus and in-town events was provided by Dr. Sagy Cohen, who made sure the fellows had proper on-campus work environments and worked with the university to get through bureaucratic hoops. Technical training provided by Matt Williamson, Xia Feng, Nels Frazier, and Mike Johnson from the NWC provided critical expertise for fellows that used NOAA modeling and data products. Austin Raney (CUAHSI) and Jonathan Frame (Cloud to Street) also contributed exceptional instruction and technical support when students were learning how to use GitHub, Python, and the CUAHSI computational environment. Course Coordinators Emily Deardorff and Arash Modaresi Rad were instrumental in planning the first two weeks, helping to structure the research timeline and deliverables, and providing on-site participant support throughout the seven-week program.

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 185 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 National Water Center - Office of Water Prediction for hosting and supporting this innovative activity, and for the opportunity to contribute to the enhancement of water prediction for our nation.

Emily Deardorff
Student Course Coordinator, NWC Innovators Program - Summer Institute 2022
Department of Geography, San Diego State University

Arash Modaresi Rad
Student Course Coordinator, NWC Innovators Program - Summer Institute 2022
Computing Ph.D. Program, Boise State University

Chapter 1

Automated decision support for model selection in the Nextgen National Water Model

Lauren A. Bolotin[1], Francisco Haces-Garcia[2], Mochi Liao[3], Qiyue Liu[4]

[1] San Diego State University; *lbolotin3468@sdsu.edu*
[2] University of Houston; *fhacesgarcia@uh.edu*
[3] Duke University; *mochi.liao@duke.edu*
[4] University of Illinois at Urbana-Champaign; *qiyuel3@illinois.edu*

Academic Advisors: Hilary McMillan, *San Diego State University*[1]; Craig L. Glennie, *University of Houston*[2]; Hanadi S. Rifai, *University of Houston*[2]; Ana P. Barros, *University of Illinois at Urbana-Champaign* [3, 4]

Summer Institute Theme Advisors: Dr. Jonathan M. Frame, University of Alabama, *jmframe@crimson.ua.edu*; Dr. Fred L. Ogden, Office of Water Prediction - National Water Center, *fred.ogden@noaa.gov*

Abstract: The Next Generation Water Resources Modeling Framework (Nextgen), which is currently under development, is planned to deploy version 4.0 of the National Water Model (NWM). The model agnostic Nextgen Framework allows users to mosaic multiple hydrologic models for one modeling task and evaluate the performance of different models using a unifying structure and standard. However, there is no existing methodology for choosing an optimal model (or models) for a given catchment. We develop a model selection framework by training a random forest regressor (RF) to predict the performance of two hydrologic models in the Nextgen Framework, namely a Long short-term memory (LSTM) and Conceptual Functional Equivalent (CFE) models. The RF predicted hydrologic model performance (Normalized Nash-Sutcliffe Equivalent) with an RMSE of 0.133 for LSTM and 0.106 for CFE, and suggested that LSTM performs better in 458 of 495 basins whereas CFE performs better in the remaining 37 basins, correctly identifying the model with the highest NNSE in 81% of cases. This methodology was designed to be adaptable as new hydrologic models are implemented into the Nextgen Framework, therefore providing a flexible and robust workflow for taking full advantage of the modeling capabilities offered by the Nextgen Framework.

---

1. Motivation

1.1 Why predict streamflow?

Accurate streamflow prediction is crucial for flood control, irrigation planning, and food management [1]. However, due to the spatial heterogeneity of nature, the cost of hydrologic instrumentation, and difficulty accessing remote areas, it is challenging to obtain streamflow data in most remote areas. As a result, models are frequently used to simulate streamflow in data-sparse regions. In 2003, the International Association of Hydrological Sciences launched the Predictions in Ungauged Basins (PUB) initiative for the following decade to improve scientific understanding of hydrological processes and associated uncertainties [2]. Dominant hydrological processes vary spatiotemporally [3]. Therefore, models with different process representations may more accurately simulate streamflow across different basins and times. For this reason, multiple models may be considered in one region. The research gap in a methodology for selecting such models is addressed by this study.

1.2 The Next Generation Water Resources Modeling Framework (Nextgen)

Recently, the National Water Center (NWC) established the Next Generation Water Resources Modeling Framework (hereafter referred to as Nextgen) to allow users to evaluate the performance of different models, or model combinations, using a unified structure and standard [4]. At the time of this study, there are three rainfall-runoff models implemented in Nextgen (CFE, LSTM, and TOPMODEL). The appeal of the modular Nextgen Framework is the ability to use the optimal model for each catchment for multi-catchment hydrologic modeling tasks; however, there is no established process for making automated model selection decisions.

2. Objectives and Scope

We propose a data-driven methodology to select appropriate model formulations in the Nextgen Framework that will be suitable for use in ungauged basins in the contiguous U.S. (CONUS). This method is designed to be adaptable as new models are added and tested within Nextgen. Using two models already embedded in Nextgen (CFE and LSTM), we determine which model to use in which catchment by calibrating/training and running these models on a large sample of catchments across CONUS. The resulting performance metrics and catchment attributes are used to train a machine learning model (random forest regressor) to predict the comparative model performance and thus advise model selection.

After the models are configured within the decision support system, an a-priori estimate of model performance for any basin can be obtained using only catchment attributes, without requiring model calibration and parameter estimation. This is particularly useful in ungauged basins, where such procedures carry large uncertainties and are computationally prohibitive at the CONUS scale. We expect to provide users of Nextgen with a reproducible workflow for selecting which model to use in which catchment using only widely available catchment attributes, ensuring that this framework is usable anywhere in CONUS. While our results will be limited to an evaluation of the method on our study catchments using two specific models, we expect that this method will be adaptable as additional models become compatible with Nextgen.

3. Previous Studies

3.1 Challenges of Prediction in Ungauged Basins

The majority of small basins (<100 km^2) in CONUS are ungauged, posing a significant threat to flood prediction and water management. Since the PUB initiative, different approaches have been developed to improve hydrological model performances in ungauged basins such as using transfer learning [5] and remote sensing [6]. Nevertheless, the large number of parameters in some hydrologic models may lead to high degrees of freedom, allowing good results by calibrating a large number of parameters against unreasonable physics-based constraints (i.e. getting the right answer for the wrong reason). Therefore, a demand has emerged for either 1) improving the relationship between parameters and physics, or 2) eliminating unreasonable physic-constraints [7,8].

3.2 Regionalization

Several studies have used regionalization approaches to overcome the challenges of PUB, relying on the idea that physiographically and climatically similar catchments will produce similar hydrologic responses. While this is not a perfect method due to the uniqueness of each catchment [9], it is a current best practice given the lack of direct hydrologic observation in ungauged basins [10]. Regionalization studies often cluster similar catchments by their catchment attributes [11] or by their hydrologic signature values, using machine learning to derive meaning from these groups in terms of physiographic and climatic drivers a posteriori

[3]. A regression relationship between catchment attributes and model performances can also be developed to predict how models might behave in ungauged basins [12]. Addor et al. [13] determined what aspects of catchment hydrology (via hydrologic signatures) can be predicted well by catchment attributes, as well as which subset of attributes are most important for the accuracy of those predictions.

Several studies suggest that climatic attributes are the dominant control for hydrological responses [3,13]. We expect hydrologic model parameters to be related to the catchment attributes that control specific hydrologic processes. Based on this assumption, regionalization is a helpful starting point for model parameterization in ungauged basins [10]. The use of regionalization to inform model selection, however, has not been explored as extensively. In this study, we explore the predictability of hydrological model performances using catchment attributes.

4. Methodology

4.1 The CAMELS dataset

Recent advances in the effort of prediction in ungauged basins [3,13] have been possible due to the increased availability of large sample hydrometeorological datasets. In the US in particular, the 671-catchment CAMELS dataset, with near-complete daily USGS streamflow and forcing data from three different sources (1980-2014), as well as catchment-averaged physiographic and climatic attribute data, has promoted the frequency and impact of continental scale hydrologic studies [14,15]. An hourly version of the streamflow and National Land Data Assimilation System (NLDAS) forcing has been developed to promote high-frequency rainfall-runoff predictions for 516 of the CAMELS catchments [16]. We use the hourly dataset and catchment attributes to design a model selection process for use within the Nextgen Framework, as described in the subsequent sections.

4.2 LSTM Training

We trained 10 LSTMs, splitting up the 516 CAMELS basins into 10 training and testing sets (90% of basins for each training set, 10% of basins for each testing set, where each basin is present in only one test set and nine training sets). This setup was used to demonstrate performance in ungauged basins for our LSTMs (predicting out-of-sample in terms of space and time). This means, however, that our LSTM models were trained on only 516*0.9 basins. We adapted the model configurations from the NeuralHydrology package documentation [17] to be compatible with the hourly CAMELS data and the training/testing sets of basins we created. We trained the model with 6 years of data (water years 2008-2013), a training period with comparable performance (area under the distribution curve (AUC) of Nash-Sutcliffe Efficiency (NSE) [18] = 0.31) to a longer, more computationally intensive training period of 9 years (AUC of NSE = 0.30), but a substantially better performance than a period of just 3 years (AUC of NSE = 0.36) [19]. We skipped validation for the LSTMs to maximize the number of catchments and the variability in catchment attributes used for training. We did not perform hyperparameter tuning, further eliminating the need for validation. We used a testing period of 5 years (water years 2003-2007). We applied the optimized hyperparameters for an hourly LSTM determined in Gauch et al. [16], aside from the number of epochs (n = 3), the sequence length (1440 hours), and the included static catchment attributes. We use a reduced set of catchment attributes to reflect the information expected to be publicly available through an R package in the future as part of Nextgen.

4.3 CFE Calibration

CFE is a conceptualized hydrological model [7] that simplifies the runoff schemes but provides functionality similar to the current WRF-Hydro based version of the NWM. As with most conceptual models, CFE requires a range of parameters to predict streamflow. Calibrating CFE on the gauged CAMELS basins results in finding the optimal set of parameter values that minimizes the difference between model estimations and observations.

We chose nine parameters and two runoff subroutines (Schaake and Xinanjiang) to calibrate in CFE and used the Dynamically Dimensioned Search optimization algorithm and the Kling-Gupta Efficiency objective function to identify a performant set of parameter values for each CAMELS basin. A python package, SPOTPY, was used to automatically perform 300 iterations of calibration for each basin. The calibration and validation periods used for CFE are the same years as the training and testing periods used for the LSTM, respectively. A one-year spin-up period was implemented prior to each calibration and validation period. Default parameter values were provided by the NWC [20]. Due to insufficient observation data for the entire calibration and validation periods, 21 basins were left out in the calibration process. In total, 495 CAMELS basins were calibrated and validated. CFE is typically paired with a snowmelt module for cold regions, however we skipped this step for simplicity, which may impact the final parameter values used for validation.

4.4 Proposed Decision Support

For each the LSTM and CFE simulation results, we developed a 5-fold, cross-validated Random Forest Regressor model (RFR) to predict hydrological model performance (i.e. Normalized NSE (NNSE), Eq. 1) using the same static attributes used to train the LSTM.

$$NNSE = \frac{1}{2-NSE} \tag{1}$$

For a given basin, the predicted NNSE from a RFR is then used to select the model with the higher predicted NNSE, or the Best Predicted Model (BPM). The NNSE obtained from the LSTM training and CFE calibration were compared to determine the Best Actual Model (BAM). BAM is included as a baseline comparison to the BPM, but this would not be available in ungauged basins since it would require model calibration and training.

Following the calculation of the BPM and the BAM, we tested several model ensembling methods. Literature on ensembling Deep Learning methods with process-based methods suggests a weighting scheme based on performance metrics shown in Eq. 2 [21].

$$\hat{Q} = \sum_{i=1}^{n} \left(\frac{\hat{NNSE}_i}{\sum_{j=1}^{n} (\hat{NNSE}_j)} * \hat{Q}_i \right) \tag{2}$$

In summary, this decision support framework can provide three NNSE metrics for each ungauged basin to be compared to the BAM: 1) the predicted NNSE value if CFE is used 2) the predicted NNSE value if LSTM is used, 3) the predicted NNSE value when the BPM is used.

5. Results

5.1 LSTM and CFE

Across the 10 LSTMs trained, the median NNSE from the test sets is 0.70. Calibrated CFE has a median NNSE of 0.56 for the validation period. **Figure 1C** demonstrates the summary

statistics for each model's performance. As shown in **Figures 1A** and **1B**, the lowest LSTM and CFE NNSE values occur in the mountain west and northern plains. Higher NNSE values concentrate along the coasts. The cumulative distribution function of NNSE for LSTM and CFE (**Figure 1D**) and a map of the difference between the NNSE for each model (**Figure 1E**) demonstrate that LSTM produces higher NNSE values overall across CAMELS, even when LSTM is formulated for PUB and CFE is calibrated and validated on each individual basin. Some of the CFE parameters may overcompensate during calibration to generate streamflow predictions similar to observation in regions with a significant snowmelt component, whereas LSTM can learn these processes from the training data.

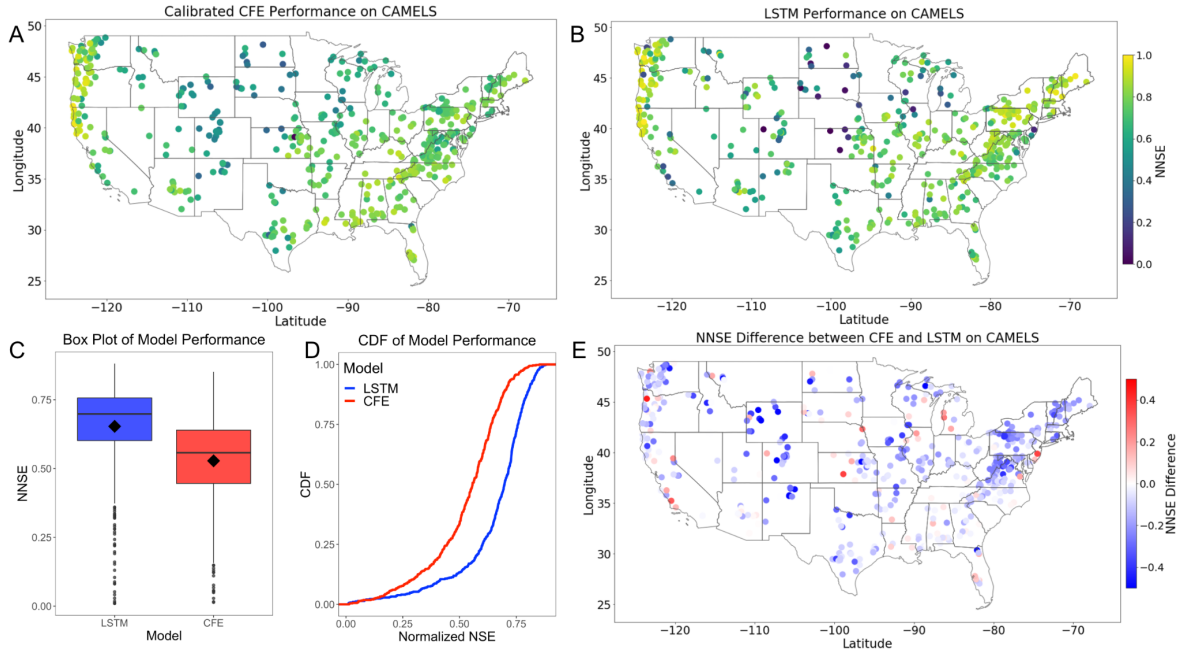


Figure 1. *A) Normalized NSE (NNSE) for CFE across CAMELS. B) NNSE for LSTM across CAMELS. C) Boxplot of NNSE values for LSTM and CFE across CAMELS. D) Cumulative Distribution Function (CDF) of NNSE values for LSTM and CFE. E) Difference in NNSE for LSTM and CFE, where blue values suggest LSTM had a higher NNSE and red values suggest CFE had higher NNSE.*

To add further context to the performance metrics from each model, we calculated 8 hydrologic signatures on the observed, LSTM simulated, and CFE simulated streamflow using the NeuralHydrology package [19]. As shown in **Figure 2**, the LSTM simulated signatures are generally closer to the observed signature values than CFE simulated signatures. CFE simulates the baseflow index and runoff ratio particularly poorly, which is likely an artifact of the lack of snowmelt information being given to the model, requiring it to compensate by optimizing parameters during calibration. For this reason, one may consider eliminating CFE as a candidate model in regions dominated by baseflow and snowmelt, even in some instances where CFE has a high NNSE.

5.2 Random Forest Regressions

Prior to any result evaluation, overfitting tests were performed on each of the folds of the RFRs to establish whether the decision support system could be generalized. Since performance on a test set was largely similar to the out of bag errors for each of the folds, the models were found to be generalizable. The RMSEs and R^2 on the predicted NNSEs of each model are reported in **Figure 3.A**. 5-fold, Cross-Validated−Random RFR is developed to predict hydrological model performance (NNSE) using basin attributes. A comparison graph is shown in **Figure 3**.

The correlation coefficient for the training phase is significantly higher than the testing phase for both CFE-RFR and LSTM-RFR. The RFRs correctly predicted the most performant model in 81% (n=400/495) of basins. To improve overall predictions, Weighted Ensembling (WE) as shown in Eq. 2 was performed. Candidate basins were identified as those with RFR-NNSEs within 0.2 for both models (i.e., neither model is predicted to be clearly better). Of the 381 candidate basins, ensembling improved upon the RFR prediction (BPM) in 56% of cases (n=216/381) and upon the BAM in 46% of cases (n=177/381), as shown in **Figure 4**. This suggests that, beyond serving as a selection framework, the RFRs add value to ensembling schemes by providing metrics to ensemble upon, even outperforming single-model results.

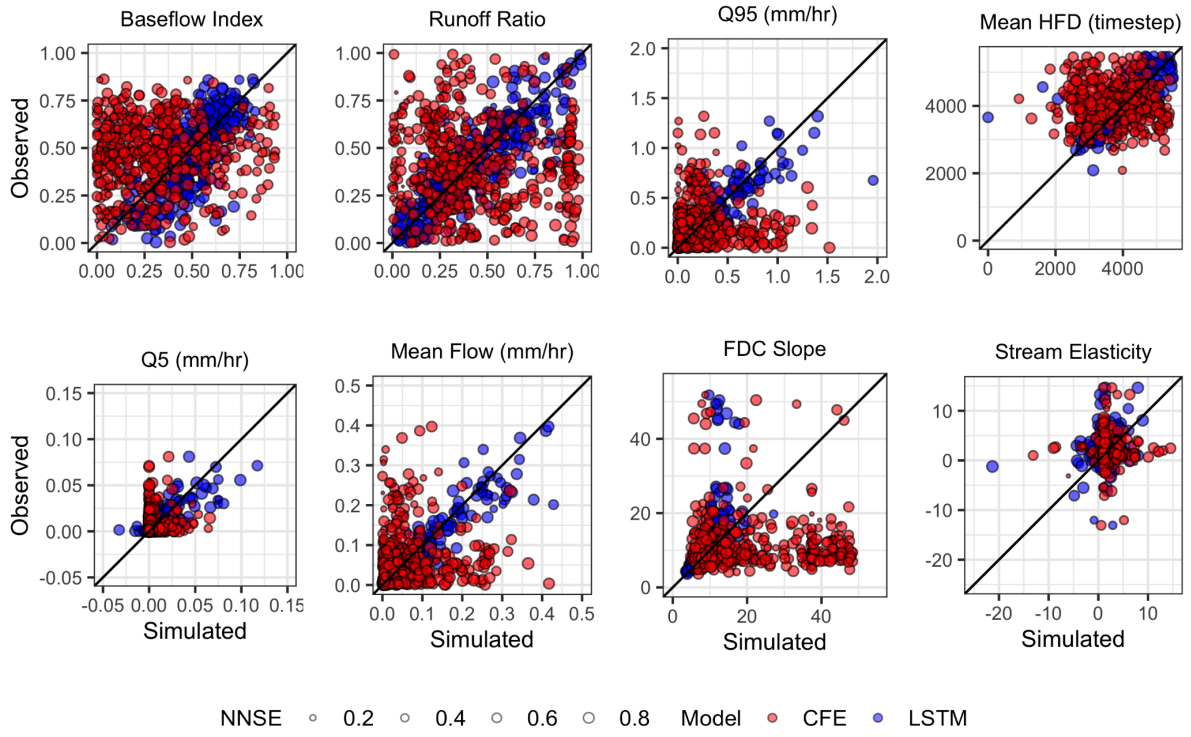


Figure 2. *Comparison of hydrologic signatures calculated on observed USGS streamflow data and streamflow simulated by LSTM and CFE. Points on the 1:1 line represent the greatest similarity between the observed and simulated signature values, whereas points above the line imply an overprediction and points below imply an underprediction. Q95 is the 95th percentile flow, Q5 is the 5th percentile flow, HFD is half flow date, and FDC is flow duration curve.*

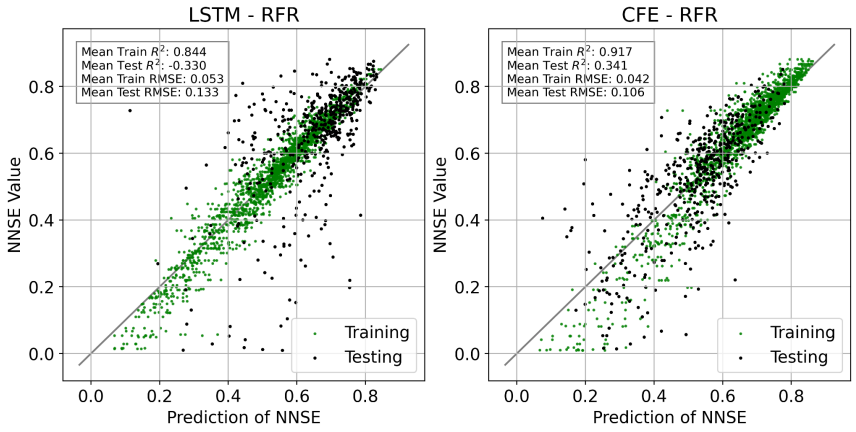


Figure 3. *Scatter plot of RFRs for training phase (green scatters) and testing phase (testing scatters). CFE-RFR and LSTM-RFR have similar predictability during the training phase. However, the transferability of CFE-RFR is better than LSTM-RFR given that the testing R^2 are 0.341 and -0.330 for the testing phase respectively.*

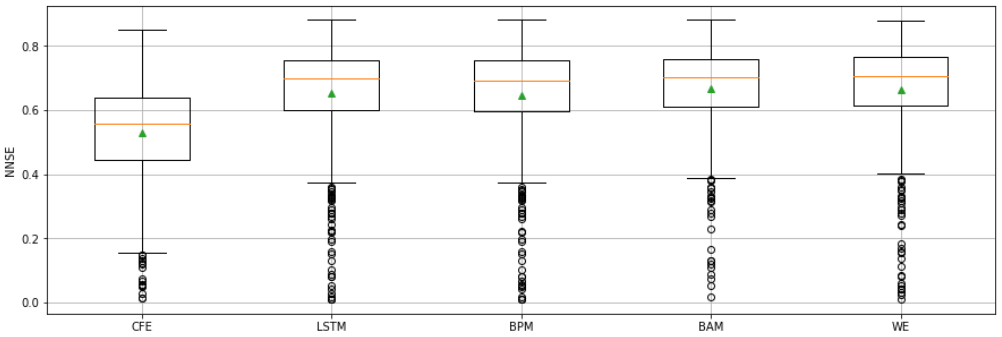


Figure 4. *Box plot of model performances for validated CFE, calibrated LSTM, BPM (Best Predicted Model), BAM(Best Actual Model) and WE(Weighted Ensembling). The median is denoted as the red line and the mean is plotted as a green triangle. The box main body covers the 25th to 75th percentile of the data. Calibrated LSTM, BPM and BAM share similar predicted NNSE data structures. Validated CFE has slightly worse performance than the rest of the model schemes.*

6. Conclusion

Addor and Melsen [22] investigated the motivation behind the selection of models in 1,500+ hydrologic studies, concluding that user familiarity or institution affiliation with a certain model, rather than model performance, was the key determinant for model selection. A user that is highly familiar with a model may be able to obtain better performance from that model than a new user due to experience with the parameters, however only an intercomparison of several models can elucidate the best performance that could potentially be achieved. These comparisons can require significant time and effort to set up all models individually, however modular modeling frameworks, like the model agnostic Next Generation Water Resources Modeling Framework, which is planned to be implemented in version 4.0 of the National Water Model, can expedite this process [24]. Even with the development of such a framework, informed guidance for selecting the most performant model in each basin remains invaluable, which is what we aim to provide here. LSTM and CFE were retrained and calibrated, respectively, using hourly forcing and streamflow observations in 495 CAMELS catchments. Two RFRs were developed to predict the NNSE of the CFE and LSTM streamflow simulations based on catchment attributes. This research found that RFRs can accurately predict model performance a-priori, which presents an important selection pathway for model formulations within the Nextgen National Water Model. Moreover, we show that constructing ensembles of models based on the predicted model performance can improve overall fit in certain basins, which could signify that selecting combinations of models to ensemble is worthwhile for future study.

Supplementary Material: For supplemental information and figures, please visit here. Additionally, all code is available on GitHub: https://github.com/bolotinl/nextgen-form-eval

References

1. Yaseen, Z. M.; Awadh, S. M.; Sharafati, A.; Shahid, S. Complementary Data-Intelligence Model for River Flow Simulation. *Journal of Hydrology* **2018**, *567*, 180–190. https://doi.org/10.1016/j.jhydrol.2018.10.020.
2. Sivapalan, M.; Takeuchi, K.; Franks, S. W.; Gupta, V. K.; Karambiri, H.; Lakshmi, V.; Liang, X.; McDONNELL, J. J.; Mendiondo, E. M.; O'Connell, P. E.; Oki, T.; Pomeroy, J. W.; Schertzer, D.; Uhlenbrook, S.; Zehe, E. IAHS Decade on Predictions in Ungauged Basins (PUB), 2003–2012: Shaping an Exciting Future for the Hydrological Sciences. *Hydrological Sciences Journal* **2003**, *48* (6), 857–880. https://doi.org/10.1623/hysj.48.6.857.51421.

3. Wu, S.; Zhao, J.; Wang, H.; Sivapalan, M. Regional Patterns and Physical Controls of Streamflow Generation Across the Conterminous United States. *Water Resources Research* **2021**, *57* (6). https://doi.org/10.1029/2020WR028086.
4. Ogden, F.; Avant, B.; Bartel, R.; Blodgett, D.; Clark, E.; Coon, E.; Cosgrove, B.; Cui, S.; Kindl da Cunha, L.; Farthing, M.; Flowers, T.; Frame, J.; Frazier, N.; Graziano, T.; Gutenson, J.; Johnson, D.; McDaniel, R.; Moulton, J.; Loney, D.; Peckham, S.; Mattern, D.; Jennings, K.; Williamson, M.; Savant, G.; Tubbs, C.; Garrett, J.; Wood, A.; Johnson, J.. The Next Generation Water Resources Modeling Framework: Open Source, Standards Based, Community Accessible, Model Interoperability for Large Scale Water Prediction; New Orleans, LA, 2021.
5. Qi, J.; Li, S.; Bourque, C. P.-A.; Xing, Z.; Meng, F.-R. Developing a Decision Support Tool for Assessing Land Use Change and BMPs in Ungauged Watersheds Based on Decision Rules Provided by SWAT Simulation. *Hydrol. Earth Syst. Sci.* **2018**, *22* (7), 3789–3806. https://doi.org/10.5194/hess-22-3789-2018.
6. Fortin, J.-P.; Turcotte, R.; Massicotte, S.; Moussa, R.; Fitzback, J.; Villeneuve, J.-P. Distributed Watershed Model Compatible with Remote Sensing and GIS Data. I: Description of Model. *J. Hydrol. Eng.* **2001**, *6* (2), 91–99. https://doi.org/10.1061/(ASCE)1084-0699(2001)6:2(91).
7. Ogden, F.L. Justification of and parameter estimation for a conceptual functional equivalent (CFE) formulation of the NOAA-NWS National Water Model, *NOAA-NWS Office of Water Prediction*, 205 Hackberry Ln, Tuscaloosa, AL 35405, Revised, June **2022**, 19 pp.
8. Frame, J. M., Kratzert, F., Gupta, H. V., Ullrich, P., and Nearing, G. S. On strictly enforced mass conservation constraints for modeling the rainfall-runoff process. *Hydrological Processes*, in review. **2022.**
9. Beven, K. J. Uniqueness of Place and Process Representations in Hydrological Modelling. *Hydrol. Earth Syst. Sci.* **2000**, *4* (2), 203–213. https://doi.org/10.5194/hess-4-203-2000.
10. Mosavi, A.; Golshan, M.; Choubin, B.; Ziegler, A. D.; Sigaroodi, S. K.; Zhang, F.; Dineva, A. A. Fuzzy Clustering and Distributed Model for Streamflow Estimation in Ungauged Watersheds. *Sci Rep* **2021**, *11* (1), 8243. https://doi.org/10.1038/s41598-021-87691-0.
11. Jehn, F. U.; Bestian, K.; Breuer, L.; Kraft, P.; Houska, T. Using Hydrological and Climatic Catchment Clusters to Explore Drivers of Catchment Behavior. *Hydrol. Earth Syst. Sci.* **2020**, *24* (3), 1081–1100. https://doi.org/10.5194/hess-24-1081-2020.
12. Guo, Y.; Zhang, Y.; Zhang, L.; Wang, Z. Regionalization of Hydrological Modeling for Predicting Streamflow in Ungauged Catchments: A Comprehensive Review. *WIREs Water* **2021**, *8* (1). https://doi.org/10.1002/wat2.1487.
13. Addor, N.; Nearing, G.; Prieto, C.; Newman, A. J.; Le Vine, N.; Clark, M. P. A Ranking of Hydrological Signatures Based on Their Predictability in Space. *Water Resour. Res.* **2018**, *54* (11), 8792–8812. https://doi.org/10.1029/2018WR022606.
14. Addor, N.; Newman, A. J.; Mizukami, N.; Clark, M. P. The CAMELS Data Set: Catchment Attributes and Meteorology for Large-Sample Studies. *Hydrol. Earth Syst. Sci.* **2017**, 21.
15. Newman, A. J.; Clark, M. P.; Sampson, K.; Wood, A.; Hay, L. E.; Bock, A.; Viger, R. J.; Blodgett, D.; Brekke, L.; Arnold, J. R.; Hopson, T.; Duan, Q. Development of a Large-Sample Watershed-Scale Hydrometeorological Data Set for the Contiguous USA: Data Set Characteristics and Assessment of Regional Variability in Hydrologic Model Performance. *Hydrol. Earth Syst. Sci.* **2015**, *19* (1), 209–223. https://doi.org/10.5194/hess-19-209-2015.

16. Gauch, M.; Kratzert, F.; Klotz, D.; Nearing, G.; Lin, J.; Hochreiter, S. Rainfall–Runoff Prediction at Multiple Timescales with a Single Long Short-Term Memory Network. *Hydrol. Earth Syst. Sci.* **2021**, *25* (4), 2045–2062. https://doi.org/10.5194/hess-25-2045-2021.
17. Kratzert, F.; Gauch, M.; Nearing, G.; Klotz, D. NeuralHydrology — A Python Library for Deep Learning Research in Hydrology. *JOSS* **2022**, *7* (71), 4050. https://doi.org/10.21105/joss.04050.
18. Nash, J. E.; Sutcliffe, J. V. River Flow Forecasting through Conceptual Models Part I — A Discussion of Principles. *Journal of Hydrology* **1970**, *10* (3), 282–290. https://doi.org/10.1016/0022-1694(70)90255-6.
19. Gauch, M.; Mai, J.; Lin, J. The Proper Care and Feeding of CAMELS: How Limited Training Data Affects Streamflow Prediction. *Environmental Modelling & Software* **2021**, *135*, 104926. https://doi.org/10.1016/j.envsoft.2020.104926.
20. Luciana Kindl da Cunha, personal communication, July 5, 2022
21. Nourani, V.; Gökçekuş, H.; Gichamo, T. Ensemble Data-Driven Rainfall-Runoff Modeling Using Multi-Source Satellite and Gauge Rainfall Data Input Fusion. *Earth Sci Inform* **2021**, *14* (4), 1787–1808. https://doi.org/10.1007/s12145-021-00615-4.
22. Addor, N.; Melsen, L. A. Legacy, Rather Than Adequacy, Drives the Selection of Hydrological Models. *Water Resour. Res.* **2019**, *55* (1), 378–390. https://doi.org/10.1029/2018WR022958.

Chapter 2

Data assimilation of USGS streamflow in the Nextgen Framework

Fitsume T. Wolkeba[1], Kenneth O. Ekpetere[2], Motasem S. Abualqumboz[3], and Zachariah J. Butler[4]

[1] University of Alabama; *fwolkeba@crimson.ua.edu*
[2] University of Kansas; *kennethekpetere@ku.edu, kenneth.ekpetere@gmail.com*
[3] Utah State University; *motasem.abualqumboz@usu.edu*
[4] Oregon State University; *butterz@oregonstate.edu*

Academic Advisors: Mesfin Mergia Mekonnen, *University of Alabama*[1]; Xingong Li, *University of Kansas*[2]; David G Tarboton, *Utah State University*[3]; Stephen Good, *Oregon State University*[4]

Summer Institute Theme Advisors: Fred L. Ogden, NOAA Federal, *fred.ogden@noaa.gov*, Jonathan M. Frame, University of Alabama, *jmframe@crimson.ua.edu*

Abstract: Streamflow predictions derived from a hydrologic model are hampered by many sources of errors, including uncertainties in meteorological inputs and inadequate representation of natural systems' physical processes. One way to reduce the impact of these uncertainties on models' accuracy and improve streamflow predictions is to implement Data Assimilation (DA) methods. DA methods use streamflow observations to update models' state variables to improve streamflow predictions. This study aimed to build and demonstrate DA as a module ready for use in the Next Generation Water Resources Modeling Framework (Nextgen). All components of this module conform to the Basic Model Interface (BMI) standards to be compatible with the Nextgen framework. This includes a BMI-enabled python-module that uses the Ensemble Kalman Filter (EnKF) to derive state transition functions for other Nextgen modules. An additional BMI-enabled python-module was also created to collect USGS streamflow observations to assimilate streamflow predictions. The developed EnKF and USGS modules were tested using streamflow simulations from the BMI-enabled Conceptual Functional Equivalent (CFE) version 2.1 of the National Water Model (NWM). The existing BMI-enabled CFE was used to create a new version of the BMI-enabled CFE with added perturbations to produce an ensemble of streamflow simulations required to run the EnKF method. The DA module was tested on synthetic data and a real-world example at a USGS gauged basin. The outcomes showed that implementing the EnKF DA method improved the CFE streamflow predictions. The difference between the assimilated streamflow predictions and USGS streamflow observations was smaller than the difference between original CFE simulations and USGS observations. When compared to the USGS streamflow observations, the assimilated CFE streamflow predictions improved correlation coefficient (r), Kling-Gupta Efficiency (KGE), Nash-Sutcliffe Efficiency (NSE), and percent bias (PBIAS) statistics.

1. Motivation

Nextgen is an open-source large-scale hydrological modeling framework currently being developed by the National Oceanic and Atmospheric Administration to forecast streamflow over the continental United States (CONUS). The motivation of Nextgen is to facilitate a research community collaboration, promote model interoperability and allow evidence-driven model selection. The Nextgen modeling framework consists of (1) hydrological forecast models and modules, (2) CONUS hydrofabric, and (3) forcing input data. The framework is

model agnostic and can be used to operate single or multiple multi-language hydrological models within the study area as long as they are written in Fortran, C, C++, or Python. The Nextgen framework interacts with forecast models using the open-access BMI standard methods, which make hydrological models easier to couple with other software elements and models supported by different programming languages [1]. This allows the framework to perform heterogeneous modeling simulations over multiple catchments for research-based investigations and real-time operations. This would be extremely helpful in the case of having to operate specific forecast models over specific catchments as they represent the catchment's dominant hydrological processes more than other models. However, a science-based method for a proper model selection process is yet to be defined.

The current U.S. NWM incorporates a simple nudging functionality that allows direct insertion of USGS streamflow observations. This forces the model to agree with stream gauges up until the beginning of a forecast cycle. Although the streamflow nudging process improves streamflow simulation, it does not account for the uncertainty associated with models' streamflow simulations and USGS streamflow observations [2]. The nudging process may also add noise signals to streamflow simulations. Furthermore, because the simple nudging method does not adjust model states to agree with the assimilated streamflow values, the simulation immediately diverges from the assimilated value the first time step of the forecast period, and continues to do so. Therefore, this study aims to fill that gap by creating an optimal BMI-enabled DA module to be used in the Nextgen framework to properly assimilate USGS observed streamflow data. This should enhance streamflow simulations and improve the short-term (1-18 hours) streamflow forecast accuracy. Unlike the current nudging functionality, the EnKF DA method is used to update the CFE state variables based on measuring the error characteristics of hydrological models and USGS observations. This would significantly improve the accuracy of models' streamflow predictions.

2. Objectives and Scope

This study aims to build a complete BMI-enabled DA module to assimilate the USGS streamflow observations in the Nextgen framework. The built-in DA module would improve models' streamflow predictions and the short-time (1-18 hours) forecasts. Among many DA methods, the EnKF method has been chosen because of its efficiency in improving models' streamflow predictions [3,4]. The EnKF module was created according to the BMI standard methods to be compatible with the Nextgen framework. The collection process of the USGS observed streamflow was also carried out using a BMI-enabled module created for this purpose. The EnKF method was used to assimilate USGS measured streamflow observations to their respective streamflow CFE model components. Similar to other BMI-enabled hydrological models, Nextgen can use this complete BMI-enabled DA module with the USGS and EnKF BMI-enabled sub-modules. The EnKF module was tested using streamflow simulations from the CFE version 2.1 of the NWM [5]. The performance of the EnKF method was assessed using several metrics including the r, KGE, NSE, and PBIAS statistics.

2.1 Study Area

The study area shown in **Figure 1** is a CAMELS (Catchment Attributes and Meteorology for Large-sample Studies) basin located in the southeastern part of Connecticut (USA). The basin has an area of 84.2 km^2. CAMELS basins are known for having an unregulated natural inflow [6]. The basin is made up of five comparatively small sub-basins with areas ranging from 8 km^2 to 31 km^2. In addition to the five sub-basins, the basin's hydrofabric also includes flow paths and three nexuses. The streamflow is measured using a USGS gauge (ID: 01123000) located in the downstream sub-basin. The USGS gauge and the downstream nexus are approximately 3.5 km apart. True streamflow observations were collected from the USGS gauge using the USGS

BMI-supported module. USGS observations were then used to assimilate the CFE-simulated streamflow data collected at the downstream nexus. It should be noted that the USGS gauge is located less than 100 km away from a Next Generation Weather Radar (NEXRAD) site (not shown on the map), which could cause a minimal error in forcing data measurements. Nextgen uses High-Resolution Rapid Refresh atmospheric data that incorporates NEXRAD data. The hydrofabric of the study area was created using an R script that used the USGS reference hydrofabric to produce the aggregate hydrofabric needed to run the Nextgen.

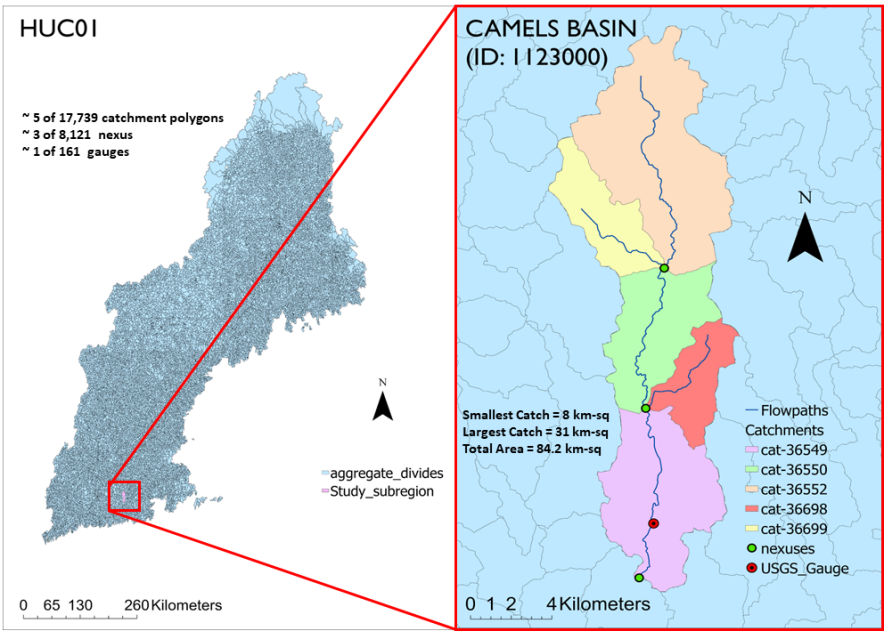


Figure 1. *The study area location and hydrofabric including five sub-basins, flow paths and three nexuses.*

3. Previous Studies

DA methods are used to merge information from measurements and models of the Earth's systems into an optimal estimate of the geophysical fields of interest. The principal goal of DA is to find an accurate prediction in spite of the many uncertainties that come from observations and model estimates [7]. Improving model run-up or forecast initialization based on DA of streamflow observations through time and space is critical for ensuring accurate predictions, specifically for hydrologic models [2,7]. DA can be implemented using different methods such as the Bootstrap Filter [3], Extended Kalman Filter [4], and the EnKF [7,8], which is widely used in atmospheric and hydrologic sciences. The EnKF method normally uses a comparably small ensemble of model trajectories and simulations to capture the relevant parts of the error structure [7].

3.1 Ensemble Kalman Filter (EnKF)

EnKF is an approximate version of the Kalman filter where the covariance matrix in the Kalman filter is replaced by sample covariance; it was originally developed for large problems [15]. The EnKF uses the Monte Carlo method for state estimation by generating a large number of sigma points [9]. These Sigma points are sample points that optimally map the probability distribution by using the input's initial mean and standard deviation values. These points are then passed through a non-linear transformation and used to build the histogram of the results. EnKF starts by randomly generating many sigma points distributed about the filter's initial state. This distribution is proportional to the filter's covariance [9]. When the filter is initialized, a large number of points are drawn from the initial state and covariance. The algorithm then proceeds to the prediction step where the sigma points are passed through the

state transition function and then perturbed. The perturbation is done by adding a bit of noise to account for the process noise [10]. At the update step, the sigma points are translated into measurement space by passing them through the measurement function; they are perturbed by a small amount to account for the measurement noise [11]. The general equations for implementing the EnKF are expressed mathematically as:

$$x_h = h(x, u) \tag{1}$$

$$Z_{mean} = \frac{1}{N} \sum_{1}^{N} xh \tag{2}$$

$$P_{zz} = \frac{1}{N-1} \sum_{1}^{N} \left[xh - Z_{mean}\right][xh - Z_{mean}]^T + R \tag{3}$$

$$P_{xz} = \frac{1}{N-1} \sum_{1}^{N} \left[x - x^-\right][xh - Z_{mean}]^T \tag{4}$$

$$K = P_{xz} P_{zz}^{-1} \tag{5}$$

$$x = x + K[z - xh + V_R] \tag{6}$$

$$X = \frac{1}{N} \sum_{1}^{N} x \tag{7}$$

$$P = P - KP_{zz}K^T \tag{8}$$

where χ is the sigma points, h is the measurement function, χ_h is the resulting points, N is the number of sigma points, Z_{mean} is the mean of the measurement sigmas, P_{zz} is the covariance for every measurement sigma point, P_{xz} is the cross variance for the measurement sigma points vs the sigma points, x is the state mean, P is the covariance, K is the Kalman gain of the update step, and R is the measurement noise matrix [12].

4. Methodology

Three BMI-enabled modules were built to achieve the goal of this study of improving the short-term flow forecast accuracy in the Nextgen framework. These modules are referred to as 1) CFE Perturbed, 2) USGS Data Retrieval and 3) EnKF DA modules. The EnKF DA process was implemented at the nexus downstream of the USGS gauge for the period from May 15th - June 15th, 2017. The DA was done directly to CFE as a stand-alone module. The T-ROUTE Muskingum channel routing module of the Nextgen model was not considered. The CFE model was run with precipitation forcing data inputs only and no evaporation data were considered.

4.1 CFE Perturbed

The CFE Perturbed model was created to produce 500 ensembles of streamflow simulations to account for the uncertainty associated with the predictions of the CFE model. Streamflow ensembles were created using a random distribution of 75% of the model's streamflow data. This added a random systematic noise to the CFE model's streamflow. Running this systematic error on the model's streamflow 500times created ensembles needed to implement the EnKF DA method. From the 500 ensembles, the mean and covariance of streamflow data could then be used in the EnKF DA method.

4.2 USGS Data Retrieval

The USGS Data Retrieval module was built using the data retrieval Python package [13] to obtain the instantaneous 15-minute average streamflow data from the USGS's National Water Information Service (NWIS) database. The module converts the 15 minute stream flow data to hourly so it can be directly used in analysis. It also provides a check for availability of data. The information needed to operate the module includes the simulation start and end times, USGS site ID, and the parameter code, which refers to the type of streamflow data needed (e.g. 00060 refers to instantaneous 15-minute average streamflow). The USGS streamflow gauge located in the study area is referred to as the Little River Near Hanover, CT and has an ID of 01123000 (**Figure 1**).

4.3 EnKF Data Assimilation and CFE state transitions

The EnKF DA module uses the filterpy python package [12]. It was prepared as a BMI-enabled module for the Nextgen framework. This module is designed to be used to assimilate USGS observations into other Nextgen modules. The EnKF DA module is used to update the state variables of the CFE model to improve its streamflow simulations and short-term predictions. The updated CFE state variables were the "soil_reservoir_storage_deficit" and "runoff_queue_m_per_timestep". The "soil_reservoir_storage_deficit" variable refers to the difference between current soil moisture content and the field capacity, whereas the "runoff_queue_m_per_timestep" controls the process of runoff generation. The "soil_reservoir_storage_deficit" and "runoff_queue_m_per_timestep" state variables were chosen because of their impact on streamflow simulations. The "soil_reservoir_storage_deficit" is updated first because it defines how much water can infiltrate, setting the leftover amount for runoff generation. The "runoff_queue_m_per_timestep" has a significant impact on streamflow simulations. For instance, more streamflow would be expected at the downstream nexus in case of increasing the "runoff_queue_m_per_timestep" variable. The "runoff_queue_m_per_timestep" state transition is applied as a multiplier to a convolution integral that delays the surface runoff according to the geomorphic instantaneous unit hydrograph. This forces the precipitation partitioning function to cause more overland runoff than groundwater recharge. Updating the state variables using the EnKF module was limited to their physical constraints. For instance, soil deficit would range from the soil's maximum field capacity (minimum deficient = saturated soil) to its effective porosity (maximum deficit = dry soil).

5. Results

The original CFE streamflow simulations (CFE Open), the assimilated CFE streamflow (CFE Assimilated), USGS streamflow observations (USGS Obs), and study area hourly averaged precipitation are presented in **Figure 2**. The plot shows that the simulated and observed streamflow data followed precipitation patterns closely. For instance, streamflow peaks were observed at the precipitation peak time. The plot presented in **Figure 2** also shows that the EnKF method improved the CFE streamflow simulation. The assimilated streamflow data lies between the original streamflow data and the USGS streamflow observations, suggesting that the EnKF DA method dragged the CFE streamflow simulations closer to the USGS observations. The CFE original simulations had an average streamflow of 26.03 cubic feet per second (cfs), whereas the CFE assimilated streamflow had an average of 40.1 cfs. The USGS mean streamflow was 67.01 cfs. The plot also shows that the EnKF DA method increased the CFE assimilated streamflow in response to a precipitation event even when USGS data is missing.

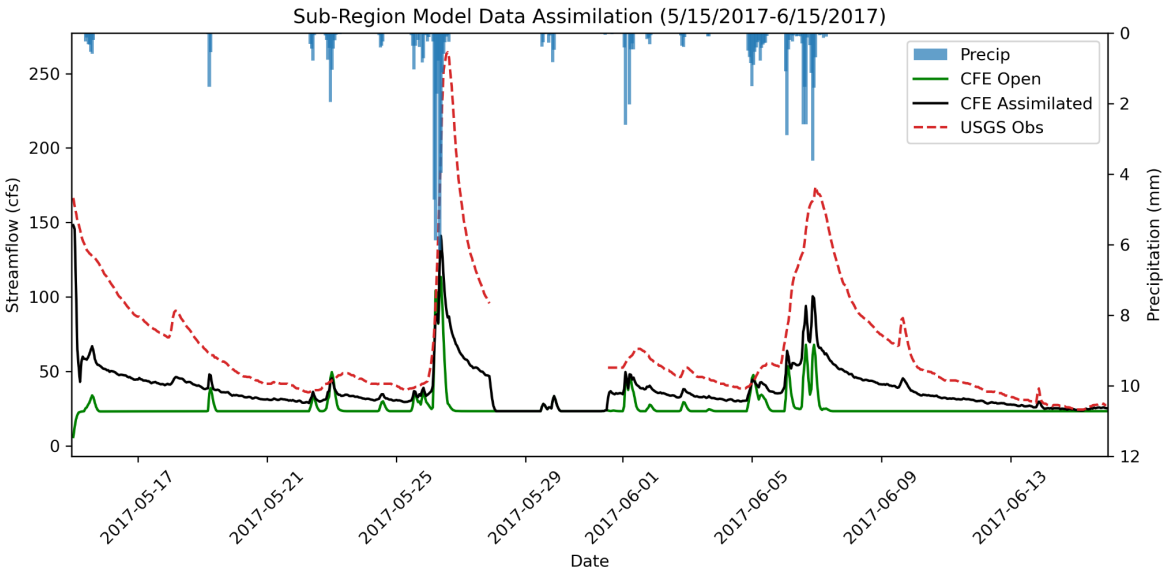


Figure 2. *CFE Open, CFE Analysis, USGS Observations, and observed precipitation from 15 May, 2017 to 15 June, 2017.*

The statistics shown in **Table 1** also highlight the fact that the EnKF improved streamflow predictions. For instance, the r increased from 0.31 to 0.88, suggesting stronger correlation between the USGS observations and the CFE Assimilated streamflow simulations. Similarly, the value of the KGE statistics improved from -2.84 to -0.82, whereas the NSE value enhanced from -35.92 to -5.46. The PBIAS statistic of the CFE simulation data was also improved from -157.39 to -67.11

Table 1. *Performance statistics for Open and assimilated CFE simulations*

Metric	CFE Open	CFE Assimilated	USGS
r	0.31	0.88	-
KGE	-2.84	-0.82	-
NSE	-35.92	-5.46	-
Percent Bias	-157.39	-67.11	-
Mean streamflow (cfs)	26.03	40.66	67.016.94

An additional example was used to test whether our DA method could improve 1-18hr hydrologic forecasts. We used synthetic forcing data to run CFE with synthetic USGS observations. In **Figure 3**, CFE Assimilated used EnKF DA with updated state variables up until hour 768. This can be observed by the dashed red line, which indicates the synthetic USGS observations that were not used within the forecasted CFE Assimilated from hour 767 to 785. The above example is another signal that the DA method used with updating state variables will improve hydrologic forecasts in the short term. We can see an improved 8 hour forecast, until CFE Open and CFE Assimilated eventually converge. This would need to be tested in other scenarios to confirm the improved forecasting capability of CFE Assimilated.

Other state variables, such as ground water reservoirs in CFE could be updated to improve forecasts past 8 hours.

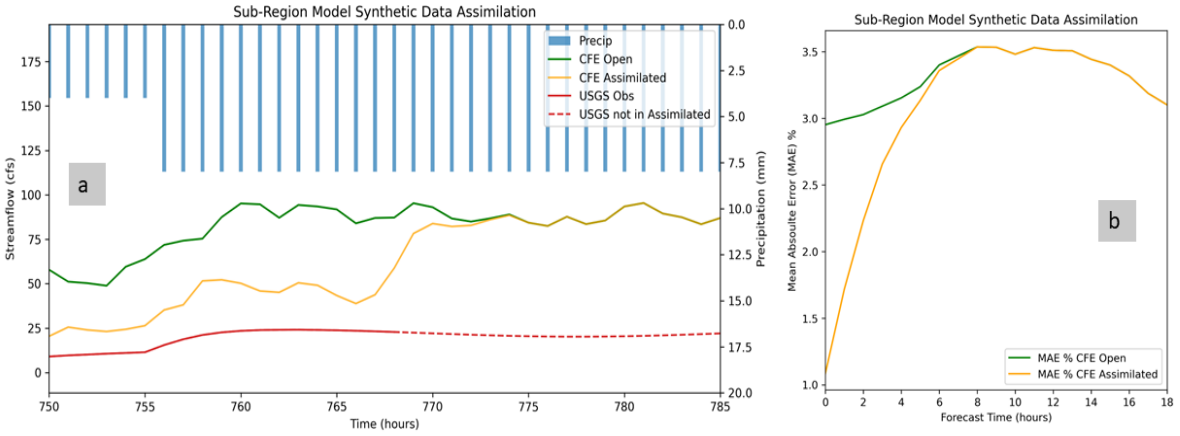


Figure 3. *CFE Open, CFE Analysis, USGS Observations, and observed precipitation using synthetic forcing data. The dashed red line indicates when USGS observations were not used within the "forecast" period from hour 767 to hour 785.*

It should be noted that these results were based on an uncalibrated CFE version and the statistics shown in **Table 1** are only used to illustrate how the CFE streamflow simulations improved due to the EnKF DA method. The BMI-enabled modules developed in this study will provide the groundwork for future studies to incorporate other modeling capabilities with CFE and the Nextgen framework to further improve on the results. As expected, updating the CFE model's state variables brought the CFE streamflow simulation closer to USGS streamflow observations.

6. Conclusion

TheEnKF DA method was implemented in this study to assimilate USGS streamflow observations in the Nextgen modeling framework to improve streamflow prediction and the short-term (1-18 hour) forecast. CFE Assimilated improved streamflow predictions up to 8 hours. Three Python BMI-enabled modules were created to (1) collect USGS streamflow observations, (2) produce an ensemble of CFE streamflow predictions and (3) implement the EnKF DA method using the USGS observations and the ensemble streamflow perturbations. The modules were created according to the BMI standards to be compatible with the Nextgen framework. The results demonstrated that the EnKF DA method improved the CFE Assimilated streamflow as illustrated by many statistics including the NSE, KGE, PBIAS and r. The study did not incorporate evapotranspiration and the T-ROUTE channel routing method of the Nextgen framework, which could improve streamflow assimilation. These components were not considered as the goal was to implement the EnKF DA method while updating the CFE state variables in the Nextgen framework. Additionally, an uncalibrated model was used, but these do not hinder the usefulness of the methodology built for others to use in future hydrologic modeling. Future research might consider addressing our limitations and using the developed DA modules with other hydrological models, other study areas, and updating additional state variables.

Supplementary Materials:

The main inputs and outcomes of this study are available online at https://github.com/zbutler33/NextGen---The-Simulators-.git.

References

1. Cunha, L., Jennings, K., Wood, A., Mizukami, N., Ogden, F., Feng, X., Liu, Y., Peckham, S., Garrett, J., Frame, J., Cui, S., Johnson, D., Mattern, D., Johnson, J., Frazier, N., Bartel, R., Avant, B., McDaniel, R., & Williamson, M. "Next Generation National Water Model: Strategy and Preliminary Performance of Initial Model Formulations." In *AGU Fall Meeting Abstracts*, vol. 2021, pp. H54G-04. 2021.
2. Seo, B. C., Krajewski, W. F., & Quintero, F. Multi-Scale Hydrologic Evaluation of the National Water Model Streamflow Data Assimilation. *JAWRA Journal of the American Water Resources Association*, **2021**, *57(6)*, 875–884. https://doi.org/10.1111/1752-1688.12955
3. Gordon, N. J., Salmond, D. J., & Smith, A. F. M. A novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEE Proceedings F Radar and Signal Processing*, (**1993**)*140(2)*, 107. https://doi.org/10.1049/ip-f-2.1993.0015
4. Sun, L., Seidou, O., & Nistor, I. Data Assimilation for Streamflow Forecasting: State–Parameter Assimilation versus Output Assimilation. *Journal of Hydrologic Engineering*, (**2017**) *22(3)*, 04016060. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001475
5. Ogden, F.L. (2021). Justification of and parameter estimation for a conceptual functional equivalent (CFE) formulation of the NOAA-NWS National Water Model (version 2.1). NOAA-NWS Office of Water Prediction, 205 Hackberry Ln, Tuscaloosa, AL, 35401, revised June 2022, 19 pp.
6. Addor, N., Newman, A. J., Mizukami, N., & Clark, M. P. The CAMELS data set: Catchment attributes and meteorology for large-sample studies. *Hydrology and Earth System Sciences*, **2017**, *21*, 5293–5313. https://doi.org/10.5194/hess-21-5293-2017
7. Reichle, R. H. Data assimilation methods in the Earth sciences. *Advances in Water Resources*, **2008**, *31(11)*, 1411–1418. https://doi.org/10.1016/j.advwatres.2008.01.001
8. Pan M., & Wood E. F. Data assimilation for estimating the terrestrial water budget using a constrained ensemble Kalman filter. *J Hydrometeorol*, **2006**, *7*, 534–547. https://doi.org/10.1175/JHM495.1
9. Shen, Z., & Tang, Y. A modified ensemble Kalman particle filter for non-Gaussian systems with nonlinear measurement functions. *Journal of Advances in Modeling Earth Systems*, **2015**, *7(1)*, 50–66. https://doi.org/10.1002/2014MS000373
10. Yamanaka, A., Maeda, Y., & Sasaki, K. Ensemble Kalman filter-based data assimilation for three-dimensional multi-phase-field model: Estimation of anisotropic grain boundary properties. *Materials & Design*, **2019**, *165*. https://doi.org/10.1016/j.matdes.2018.107577
11. Roth, M., Hendeby, G., Fritsche, C., & Gustafsson, F. The Ensemble Kalman filter: A signal processing perspective. *EURASIP Journal on Advances in Signal Processing*, **2017**, *56*. https://doi.org/10.1186/s13634-017-0492-x
12. Labbe, R. (2022). Filterpy, GitHub, GitHub repository, Link: https://github.com/rlabbe/filterpy, commit: c762273f7a64276f051215a31958df51c7a6009c
13. Jem, S. (2018). *dataretrieval: Download hydrologic data* [Python]. USGS-python. https://github.com/USGS-python/dataretrieval (Original work published 2018)
14. Matthias Katzfuss, J. R. (2016). Understanding the Ensemble Kalman Filter. THE AMERICAN STATISTICIAN, 350-357. http://dx.doi.org/10.1080/00031305.2016.114.1709

Chapter 3

Data-driven approaches for estimating river channel geometry

Shuyu Chang[1], Mohammad Erfani[2], Zahra Ghahremani [3]and Laura Manuel[4]

[1]Pennsylvania State University; *shuyu.chang@psu.edu*
[2]University of South Carolina; *serfari@email.sc.edu*
[3]Boise State University; *zahragahremani@u.boisestate.edu*
[4]Tulane University; *lmanuel2@tulane.edu*

Academic Advisors: Kimberly Van Meter, *Pennsylvania State University*[1]; Erfan Goharian, *University of South Carolina*[2]; Jen Pierce, *Boise State University*[3]; Ehab Meselhe, *Tulane University*[4]

Summer Institute Theme Advisors: Sagy Cohen, University of Alabama, *sagy.cohen@ua.edu;* Chaopeng Shen, Pennsylvania State University, *shen.chaopeng@gmail.com*

Abstract: The representation of river channel geometry is important for hydrologic and hydraulic modeling of natural systems. Often, in numerical modeling, the channel geometry is simplified and estimated by methods that can be applied generally across a variety of spatial and temporal scales. Despite their utility, commonly applied channel estimation methods, such as regional hydraulic geometry curves, are limited to power law relations that do not employ many potentially relevant catchment and reach attributes. Considering the recent availability of large-scale observed channel geometry datasets and advances in machine learning techniques, this study aims to develop data driven approaches to provide a better estimation of river geometries (depth and width) over the Continental United States (CONUS), compared to current regional hydraulic geometry curves. Multiple machine learning techniques were employed here, including multilinear regression, random forest, extreme gradient boost, multilayer perceptron, and convolutional neural network. The models showed similar results for predicting both stream width and depth, with the general highest performance for predicting stream median width. This study also suggests additional research to 1) refine variable transformation methods for machine learning models to optimize performance, 2) further supplement large scale datasets, and 3) advance model architecture to utilize imagery datasets in channel geometry estimation.

1. Motivation

River channel geometry attributes influence riverine processes. Geometric parameters, seen in **Figure 1**, respond dynamically to their local and large-scale systems and serve as indicators of riverine processes.

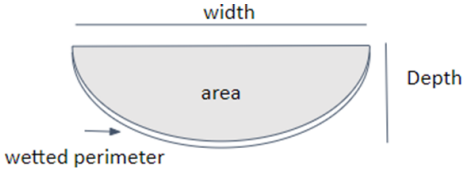


Figure 1. *Simplified river cross-section*

Specifically, the river channel geometric attributes of width and depth are key parameters for determining when bank full discharge is exceeded and flood inundation is triggered. The

definition of channel features is particularly important for flood modeling and mapping. Predicting river geometry under different flow conditions is a nonlinear problem with a high degree of spatial and temporal variabilities, which can be further complicated by uncertainties in parameter estimates [1].

Presently, extensive stream surveys do not exist to accurately model all streams within the Contiguous United States(CONUS). Therefore, to define channel geometry for large-scale model application, methods such as the empirical hydraulic geometry equations [2], are used to provide relationships between stream depth and width to drainage area or discharge. These techniques are presently utilized in the National Water Model (NWM) to provide general geometry attributes for water routing calculations within the model's stream network. The equations fall short in generating a reliable stage product and suffer from the assumption of "infinite" channel depth. This makes it impractical to represent overbank flow and limits the model's ability to realistically represent natural fluvial water dynamics. The simplification of the hydraulic geometry methods provides the opportunity for the development of enhanced models capable of providing improved channel attributes. A number of datasets have recently been made available which include river channel geometry and basin and fluvial network attributes for the US and globally. Notably, extensive Acoustic Doppler Current Profiler (ADCP) records from the Surface Water and Ocean Topography mission [3], the National Hydrography Dataset Plus (NHDPlusV2) [4], and satellite imagery, are primary datasets that provide the opportunity to pursue large scale data driven approaches to estimate channel geometry. Advances in ML allow for the more robust utilization of these datasets for applications in hydrological sciences in developing robust predictive models.

2. Objectives and Scope

The project aims to apply data driven approaches, including the recent advances in ML models, to introduce a framework for determining channel geometry (ie., width and depth), using publicly available catchment and channel characteristics datasets. It also aims to explore the potential use of satellite imagery for better estimation of channel width using Deep Learning-based approaches. This requires collecting and extracting data from different datasets and satellite geospatial data for training and testing models. This knowledge paves the way for developing end-to-end models customized for the NWM's Next Generation framework, which aims to integrate different models and promote model interoperability.

3. Previous Studies

A study by Brackins et al [5] tackled the issue of flood mapping by hypothesizing that a more realistic representative channel geometry, as opposed to the trapezoidal channel representation that is currently used for forecasting in the NWM can enhance the performance of hydrological models. They tested both the trapezoidal channel representation and improved representation on hydrological model predictions and showed that a more realistic channel geometry has a significant impact on the accuracy of simulated channel routing and stage-discharge predictions. Raney [6] used Canova et al. [3] (HydroSWOT) and Lin et al. [7] datasets for training a deep learning (DL) regression model to predict the mean river depth for the CONUS and compared the model performance with the commonly used power-law method. That study showed that the DL regression does not need to know the instantaneous discharge for estimating the mean river depth. Results demonstrated that the DL model improves estimations of river depth by decreasing RMSE by 22% compared to the power-law method. However, this model underestimated the mean depth and had limited predictive power for large mean river depth values.

4. Methodology

4.1 Data collection and preprocessing

The dataset utilized for this study was created from several nation-wide datasets. The HydroSWOT dataset [3] consists of 200,000+ (ADCP) river measurements conducted across the CONUS at discrete river locations. These surveys were conducted by the United States Geological Survey from the 1940s-2014. The NHDPlusV2 [4] dataset contains national flow network dataset with combined stream and catchment attributes for each stream reach (over 2.7M). The gSSURGO dataset [8] provides soil geographic data derived from properties stored in the National Soil Information System. Information regarding the aridity index ($AI = \frac{P}{PET}$) was retrieved from the Global Aridity Index and Potential Evapotranspiration (ET0) Climate Database v3 [9]. A subset of these datasets was extracted to provide predictor variables to channel geometry and 2 target variables: mean stream depth and width (**Table 1**).

Table 1. *Extracted variable name and description*

Variable	**Description**
mean_depth_va [3]	Measured stream mean depth (ft)
stream_wdth_va [3]	Measured stream cross-sectional width (ft)
q_va [3]	Discharge (cfs)
ACC_NDAMS2010 [4]	Accumulated number of dams built on or before 2010 based on total upstream accumulation
AI1[9]	Mean aridity index within 500m site buffer
CAT_POPDENS10 [4]	Population density per catchment
D50[mm] [4]	Median bed-material sediment particle size (mm)
MINELEVSMO [4]	Minimum elevation (smoothed) in centimeters
EVI_JAS_2012 [4]	Enhanced Vegetation Index 2012 -Summer
EVI_JFM_2012 [4]	Enhanced Vegetation Index2012 - Winter
NLCD_agriculture_16 [4]	Percent agricultural land use in 2016
NLCD_developed_16 [4]	Percent developed land use in 2016
NLCD_forest_16 [4]	Percent forested land use in 2016
SLOPE [4]	Slope of flowline based on smoothed elevations

StreamOrde [4]	Modified Strahler Stream Order
TotDASqKM [4]	Total upstream catchment area (km^2)
Percent_Clay1km_0_100cm[8]	Mean clay content (mass percentage, %) within 500m site buffer
percent_sand1km_0_100cm[8]	Mean sand content (mass percentage, %) within 500m site buffer
Percent_Silt_1km100cmfinal[8]	Mean silt content (mass percentage, %) within 500m site buffer

Using the point locations of the ADCP measurements as a base dataset, NHDPlusV2 reach attributes were joined with the nearest site within a 500m buffer. The 500m buffer was applied to extract mean values for soil characteristics and aridity index. To account for unrealistic values reported in the datasets, the combined dataset was cleaned from all outliers, suspicious and nan values. Additionally, for the sites having multiple observations, outliers of target variables (width and depth) beyond the 90% confidence level were excluded. The final number of observations and sites for each target variable dataset is recorded in **Table 2**.

Table 2. *Finalized dataset records*

Target Variable	Number of Sites	Number of Observations
Depth	5,776	51,040
Width	8,788	17,291

For sites with multiple observations, the median value of the target variable and its corresponding predictor variable attributes were assigned as the representative values per site. This provided a time stationary dataset for each target variable (single observation) in addition to the existing time varying dataset (multiple observations). Finally, the datasets were shuffled and randomly distributed to a 75:25% split of training and testing datasets. In addition to scalar data extraction for ADCP site measurements, geospatial data was collected for 8827 sites from Earth Engine Data Catalog in the form of DEM tiles (U.S. Geological Survey, 3D Elevation Program 10-Meter Resolution Digital Elevation Model) with 10 meter resolution and National Agriculture Imagery Program (NAIP) imagery with one meter resolution (USDA Farm Production and Conservation- Business Center, Geospatial Enterprise Operations). For our study, we filtered captured NAIP images between 2014-2016. The DEM and 4 bands of NAIP imagery (Red, Green, Blue, and Near-Infrared) were retrieved as 500x500m tiles for each unique site location in the finalized dataset (**Figure 2**).

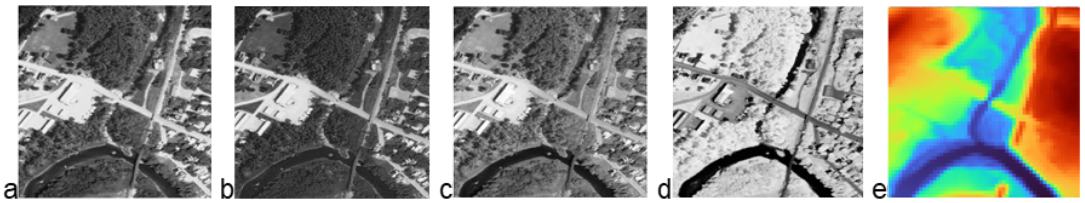


Figure 2. *NAIP Imagery: a.Red, b.Blue, c.Green, d.Near-Infrared and e.DEM*

4.2 Model development

4.2.1 Multi-Linear regressions

Multi-Linear regression models, conducted on JMP statistical software, were optimized for each target variable. To satisfy the assumptions of regression modeling, the median width and median depth datasets were utilized and all variables log transformed, given the high degree of skewness in most variables. Using forward stepwise regression to determine significant variables to model performance, an optimal model was refined for each target variable.

4.2.2 Tree-based model

Random forest regression (RFR), a flexible and robust nonparametric technique that can accommodate nonlinear relationships, high-order interactions, and missing values. This approach is increasingly utilized to model and understand environmental systems. Grid search was performed for hyperparameter tuning, such as the maximum depth of the tree, and the number of trees. To avoid problems like overfitting and selection bias as well as improving models' performances, cross-validation was conducted.

4.2.3 Multilayer Perceptron

Multilayer Perceptron (MLP), a supervised learning approach to approximate functions by connecting units called neurons across layers, is well suited to tackle complex problems, with relatively high levels of accuracy, few input variables and knowledge requirements [10]. The MLPs utilized in this project have two fully connected hidden layers following ReLU as activation functions. Training procedure was conducted using the Adam optimizer, MSE loss function. A random grid search was applied for tuning hyperparameters with respect to NSE on the test dataset.

4.2.4 CNN-based Regression Models

We developed DL-based models using both geospatial data and scalar covariates at the same time. First, five geospatial channels, including DEM, NAIP red, NAIP green, NAIP blue, and NAIP near infrared, were resized to create a tensor of 5x400x400 pixels, showing the region of interest for each site measurement. We hypothesized that the geospatial data, as they contain information about the terrain of the floodplain, shape signatures of the river, and the surrounding environments, could provide predictive power for the geometry of the river. The resulting tensors, along with scalar covariates, were then fed into the DL-based model. Considering available computational power, two different CNN architectures are introduced. ZaRA, the first model, uses ResNet-18 (18 layers) to extract features. ResNet uses "residual blocks" to avoid gradient dissipation during training the model. The extracted features from the ResNet block are concatenated with the features received from single feedforward neural network mapping scalar covariates to 64 features. The features coming from different units are concatenated and fed into a Multilayer Perceptron (MLP unit) for the regression task. LaRA uses VGG for the CNN unit to extract features from geospatial data. VGG architecture benefits from the simple structure of 3x3 convolution filters across all layers with periodic pooling. The last fully connected layer before going to the layer of classes is used to represent features of geospatial data. In order to avoid the problem of vanishing gradients in the VGG module, an intermediate regression module followed by an auxiliary loss function is added for the features extracted from the CNN path.

Models were implemented using PyTorch v1.12.0, Torchvision v0.13.0, CUDA v11.2 running on Python v3.10.4. The networks were trained using the Mean Square Error loss function and the Stochastic Gradient Descent (SGD) optimizer with a momentum of 0.9, weight decay of

0.009, and batch size of 64. The learning rate was first set to 1.75E-3, then was adjusted based on the decaying rate of the loss function during training. Figure 3 illustrates this CNN.

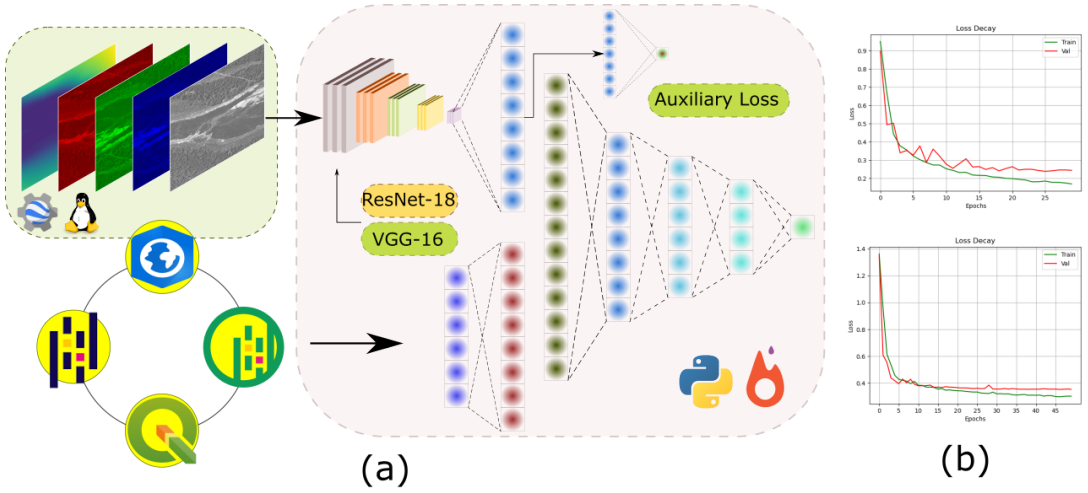


Figure 3. *(a) DL-based Models pipeline, 5 channel geospatial data is fed into CNN module, 64 features are extracted for each tensor (5x400x400). The single value covariates are fed into the warm-up layer to map to 64 features. The resulting 128 features are then fed into the following regression module. For LaRA, the auxiliary loss is considered immediately after CNN based to avoid gradient dissipation. (b) The loss function decay rate of ZaRA (b-Top) on the training set over 30 epochs, and loss function decay rate of LaRA (b-Bottom) on the training set over 50 epochs.*

4.3 Performance indicators

To quantify model performances and uncertainties, metrics were used as performance indicators by comparing observed river geometry parameters (E) against modeled (M) values, depending on transformation of target variables in model construction. The objective functions set included: the coefficient of determination (R^2), Nash Sutcliffe efficiency (NSE), and percent bias ($\%BIAS$).

5. Results

5.1 Width

The regression model produced for width estimation is described in equations (1).

$$Width(ft) = 10^{0.855} Q^{0.316} Ai^{0.165} MinElev^{-0.040} S^{-0.028} DrainA^{0.128} \quad (1)$$

The variable Q represents discharge (cfs), Ai is aridity index, $MinElev$ is the minimum elevation of the channel (cm), S is slope, $DrainA$ is drainage area of the catchment (sqkm). The equation suggests an inverse relationship between the channel width and the slope and minimum bed elevation, which may reflect steeper channel reaches (with higher bed elevations) existing, primarily, in the more headwater regions of catchments.

Model performances and corresponding statistics matrices for the target variable width are reported in **Tables 3&4**. In general, models have better results for median width than time dependent width, with the CNN models having the highest performance (test **R^2=0.82, NSE=0.82**), which may highlight the value of extra information extracted from images. The MLP performs better than the tree-based models for the time-dependent width datasets, which could potentially be explained by the variable transformations that were applied to resolve different distributions and ranges.

Table 3. *Model Performance for Target Variable: Median Width (ft)*

Model	R^2		**NSE**		**% Bias**	
	Train	Test	Train	Test	Train	Test
Multi-Linear Regression	0.72^1	0.73^1	0.73^1	0.73^1	0.10^1	-0.25^1
Random Forest	0.85	0.76	0.82	0.76	-0.19	-2.53
XGBoost	0.92	0.70	0.91	0.69	0.90	-2.73
MLP	0.80^1	0.79^1	0.79^1	0.78^1	2.2^1	1.9^1
CNN (ZaRA)	0.83^1	0.76^1	0.83^1	0.76^1	0.32^1	0.79^1
CNN (LaRA)	0.85^1	0.82^1	0.85^1	0.82^1	-0.16^1	0.29^1

Table 4. *Model Performance for Target Variable: Time dependant Width(ft)*

Model	R^2		**NSE**		**% Bias**	
	Train	Test	Train	Test	Train	Test
Random Forest	0.74	0.37	0.74	0.37	0.01	0.93
XGBoost	0.70	0.44	0.70	0.43	0.59	1.39
NN	0.86^1	0.77^1	0.86^1	0.77^1	-0.06^1	-0.03^1

[1] Log applied

5.2 Depth

The multi-linear regression model produced for depth estimation is described in equation (2).

$$Depth(ft) = 10^{0.342} Q^{0.248} MinElev^{-0.038} S^{-0.035} \%Sand^{-0.127} \%Silt^{-0.127} D50_{mm}^{-0.044} \quad (2)$$

The variable Q represents discharge (cfs), *MinElev* is the minimum elevation of the channel (cm), S is slope, *%Sand* is the local percent sand content, *%Silt* is the local percent silt content, and $D50_{mm}$ is the is the median diameter of the channel bed material (mm). The equation indicates channel depth is sensitive to different parameters than the channel width equations, namely, sediment characteristics. The inverse relationship between sand, silt, and median bed

particle size may suggest that channels are deeper in more depositional environments than in upstream environments.

Model performances and corresponding statistics matrices for the target variable depth are reported in Tables 5&6. Overall, models for time-dependent depth have better performances than median depth predictions, which could potentially be explained by the fact that river depth is more temporally dynamic in most rivers and, thus, more strongly related to discharge. Another potential reason is the scattered surveys do not really capture the true median flow conditions. That being said, the model learned more and was able to yield better estimations when trained on a larger dataset (dataset for time-dependent depth is about 9-times larger than median depth). In time-dependent depth predictions, random forest, XGBoost, and NN have equally solid performances (test R^2=0.79, NSE=0.77) with negligible biases, but NN has fewer issues of overfitting (the $R^2$and the NSE of training for NN is smaller than the test, but these models performed equally well in the test). The XGBooost and NN model in median depth predictions, remarkably outperform random forest and Multi-Linear Regression, but it should be noted that XGBoost shows significant moderate overfittings.

Table 5. *Model Performance for Target Variable: Median Depth (ft)*

Model	R^2		**NSE**		**% Bias**	
	Train	Test	Train	Test	Train	Test
Multi-Linear Regression	0.53^1	0.52^1	0.53^1	0.52^1	0.10^1	0.10^1
Random Forest	0.74	0.64	0.73	0.63	0.00	1.26
XGBoost	0.91	0.67	0.91	0.66	0.66	1.67
MLP	0.70	0.67	0.71	0.67	2.39	2.84

Table 6. *Model Performance for Target Variable: Time dependant Depth(ft)*

Model	R^2		**NSE**		**% Bias**	
	Train	Test	Train	Test	Train	Test
Random Forest	0.91	0.79	0.91	0.77	0.10	4.15
XGBoost	0.94	0.79	0.94	0.77	0.00	5.36
MLP	0.85	0.79	0.84	0.77	2.78	6.60

[1] Log applied

5.3 Regional Curve Comparison

We compared our results to the Blackburn-Lynch et al. (2017) [2] hydraulic geometry (physiographic) regional curves. These power law equations calculate bankfull depth and bankfull width as a function of drainage area, used by the NWC for estimating CONUS-wide channel geometry. Using the equations unique to 8 physiographic divisions (**Figure 4**), target variables were calculated and compared to the corresponding median depth and width in the HydroSWOT dataset. The Blackburn-Lynch et al. (2017) equations were developed and evaluated using a lower fidelity dataset, comprised of only 2,856 observational sites from 65 studies published between the late 1970s-early 2000s. **Table 7** outlines the results of this analysis.

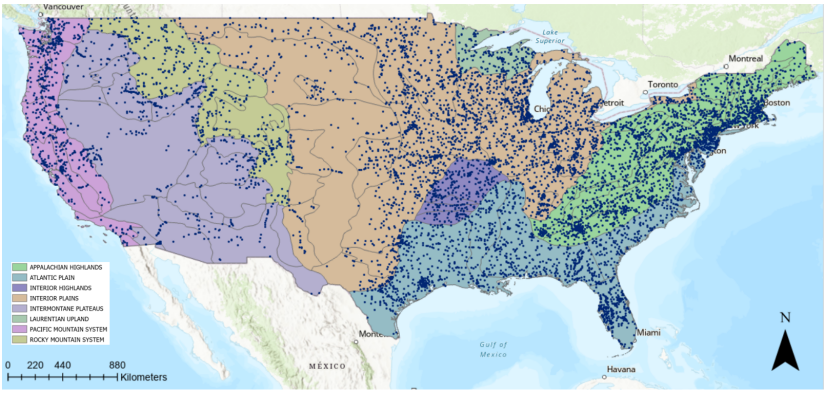


Figure 4. *Physiographic Divisions* [2]*and median width locations(blue points)*

Table 7. *Regional Curve equation performance*

River geometry	R^2	NSE	% Bias
Depth	0.12^1	-3.32^1	0.00^1
Width	0.42^1	0.38^1	0.00^1

[1] Log applied

It is important to note that the regional curve equations were developed specifically to provide a bankfull estimation of depth and width. Considering our dataset supplies target variable values that are not necessarily "bankfull", this analysis does not produce a consistent comparison to the models described in this study. However, it does strongly suggest that our modeling results are considerably more robust compared to the premier hydraulic geometry estimation for the CONUS. Moreover, the development of models to estimate channel geometry for multiple streamflow conditions may provide supplementary information over static estimations.

6. Conclusion

This study showed the potential of applying large scale datasets and ML for estimating channel geometry. The multilinear regression models have strong interpretability because they provide equations explaining the effect of each predictors on the model output. These models are computationally efficient, which is important for real-time applications. However, linear models are sensitive to data distribution and transformation, so the applicants need to have statistical knowledge for data trimming and transforming. Tree-based models, on the other hand, are not sensitive to data transformation, but easily get overfitted. In the case of channel

geometry, in which both target variables and covariates have many missing values, or outliers that are hydrologically justifiable, using Tree-based models is highly recommended. DL-based models, Neural Networks and CNN-based models are versatile tools, depending on the architecture and available computation power, and they could be designed for different purposes. These models, however, have lots of hyperparameters affecting the prediction performance significantly. Generally, there is no specific rule for fine-tuning these parameters. CNN-based variants take the advantage of processing images for extracting relevant features, so they could be used to better benefit from remote sensing data as supplements for compensating data deficiency, in the absence of predictors. All models in this study were developed based upon time dependent discharge plus the static attributes generally available nation-wide, via NHDPlus, satellite imagery, etc. As such, they have the potential to be deployed with forecasted streamflow (i.e. NWM forecasted discharge) to estimate the channel depth and width for the entire CONUS.

Supplementary Materials: https://github.com/lmanuel2/SIHydroInfo2.2022

References

1. Ghanghas, A.; Dey, S.; Merwade, V. Evaluating the Reliability of Synthetic Rating Curves for Continental Scale Flood Mapping. *J. Hydrol.* **2022**, *606*, 127470, doi:10.1016/j.jhydrol.2022.127470.
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 J. Am. Water Resour. Assoc.* **2017**, *53*, 903–928, doi:10.1111/1752-1688.12540.
3. Canova, M.G.; Fulton, J.W.; Bjerklie, D.M. USGS HYDRoacoustic Dataset in Support of the Surface Water Oceanographic Topography Satellite Mission (HYDRoSWOT) **2016**.
4. US EPA, O. NHDPlus (National Hydrography Dataset Plus) Available online: https://www.epa.gov/waterdata/nhdplus-national-hydrography-dataset-plus (accessed on 20 July 2022).
5. Brackins, J.; Moragoda, N.; Rahman, A.; Cohen, S.; Lowry, C. The Role of Realistic Channel Geometry Representation in Hydrological Model Predictions. *JAWRA J. Am. Water Resour. Assoc.* **2021**, *57*, 222–240, doi:10.1111/1752-1688.12865.
6. Raney, A.; Cohen, S.; Moragoda, N. Application of Deep Learning to Estimate Mean River Cross-Sectional Depth. **2021**, H22G-04.
7. Lin, P.; Pan, M.; Allen, G.H.; de Frasson, R.P.; Zeng, Z.; Yamazaki, D.; Wood, E.F. Global Estimates of Reach-Level Bankfull River Width Leveraging Big Data Geospatial Analysis. *Geophys. Res. Lett.* **2020**, *47*, e2019GL086405, doi:10.1029/2019GL086405.
8. Mitter, E. Gridded Soil Survey Geographic (GSSURGO-10) Database for the Conterminous United States - 10 Meter.
9. CGIAR-CSI Global Aridity Index and Potential Evapotranspiration Climate Database V3. *CGIAR-CSI* **2019**.
10. Shen, C. A Transdisciplinary Review of Deep Learning Research and Its Relevance for Water Resources Scientists. *Water Resour. Res.* **2018**, *54*, 8558–8593, doi:10.1029/2018WR022643.
11. Dawson, C.W.; Wilby, R.L. Hydrological Modelling Using Artificial Neural Networks. *Prog. Phys. Geogr. Earth Environ.* **2001**, *25*, 80–108, doi:10.1177/030913330102500104.

Chapter 4

Large scale prediction of channel roughness coefficient using machine learning

Md Abdullah Al Mehedi[1], Shah Saki[2], and Krutikkumar Patel[3]

[1]Villanova University; *mmehedi@villanova.edu*

[2] University of Connecticut; s*hah.saki@uconn.edu*

[3] Texas A&M University - Kingsville; *krutikkumar.patel@students.tamuk.edu*

Academic Advisors: Virginia Smith, *Villanova University*[1]; Emmanouil Anagnostou, *University of Connecticut*[2]; Adnan Rajib, *Texas A&M University Kingsville*[3]

Summer Institute Theme Advisors: Sagy Cohen, University of Alabama, sagy.cohen@ua.edu; Chaopeng Shen, Penn State University, cshen@engr.psu.edu

Abstract: The channel roughness is one of the most critical parameters in river hydraulics and fluvial flooding. It is spatio-temporally dynamic and subject to several external controlling factors. In the last few decades, the roughness coefficient was estimated using empirical equations. Accurate estimation of spatially explicit roughness coefficient incorporating multiple geographic and environmental influencing factors could lead to better flood and hydrological predictions. In this study, we develop a novel data-informed framework to predict channel roughness in the stream networks of the Contiguous United States (CONUS) using observed data. Several open-source datasets, including a large dataset of acoustic current doppler profiler (ADCP) surveys are used to retrieve the features for the analysis. Based on the quasi-steady-flow assumptions and flow velocities measured by ADCP, for the first time, a national-scale estimate of channel roughness values for thousands of reaches throughout CONUS reaches is conducted. Machine Learning (ML) based regressors are applied to predict channel roughness as a function of a set of influencing factors. The outcomes of the study show the applicability and robustness of the ML models in a large-scale deployment with a few data-scarce regions. The model evaluation (R^2) of Random Forest (RF) and Multilayer Perceptron (MLP) with spatial cross-validation are found to be 0.709 and 0.639 which depict satisfactory performance. The Feature Importance scheme reveals the rank of features with river discharge, total drainage area, forest, and agricultural land covers and Normalized Difference Vegetation Index (NDVI) being the top five influential features. The proposed framework reveals the hidden linkage among the influencing factors and channel roughness and a significant step towards resolving the complexities of channel roughness and flooding with a large multi-dimensional dataset in a computationally efficient manner.

1. Motivations

Channel roughness is a highly influential factor in hydro-morphological processes, streamflow prediction and flood inundation mapping. It is governed by several geographic and environmental features (e.g., size of the riverbed materials, channel features, plant type) [1]. Extensive experimentation and efforts are required for discretizing the computational domain, model calibration, validation, and uncertainty analysis in data-scarce areas in physics-based models. This may lead to significant inefficiency in large-scale deployment. However, in the age of big-data, ML-based approaches have a high potential to unravel the linkage among the channel roughness and the geomorphologic influencing factors. They may also be used to predict roughness coefficient as a function of those factors with satisfactory performance in large-scale [2–5].

2. Objectives and scopes

To our knowledge, no previous research has been conducted to predict roughness using ML-based models at a continental scale. Readily available channel roughness data can be of a major aid for decision makers and forecasters from the National Water Center (NWC) to prepare Flood Inundation Model (FIM). The major objectives of this study based on the motivations above are to 1) Develop a robust ML-based framework that predicts the spatially explicit channel roughness coefficient as a function of a collection of geographic and environmental features. 2) Identify and rank the most influential features for the roughness coefficient estimation.

3. Previous Studies

In river hydraulics, the channel roughness coefficient plays a crucial role in flood discharge prediction, velocity distribution and designing hydraulic structures. Several previous studies investigated the influence of channel velocity and geometry on the channel roughness using physics-based methods (e.g., using Chezy's *C* and Darcy- Weisbach's *f*). However, no linkage among the other geographic and environmental influencing factors from the external environment was established in that approach. Various components from the surrounding environment have the potential to influence channel roughness coefficient e.g., geometry of the channel, slant of the bed, bed harshness and other parameters of the stream. Mohanta et al. [6] proposed width ratio, relative depth, sinuosity, channel bed slope, and meander belt width ratio as the input parameters for ML-based models e.g., support vector regression and multivariate adaptive regression spline (MARS). Cowan investigated the irregularity of the channel geometry, longitudinal sinuosity, and natural/man-made obstructions as the inputs to predict Manning's roughness coefficient in meandering channels [7]. Due to the dynamics in the flow depth, cross-section and longitudinal geometry and slope, the complexities of the roughness coefficient distribution in a meandering channel can be significantly high [8]. Overstreet et al. [9] estimated the channel roughness considering the aspect ratio, viscosity, slope of the bed, and sinuosity as inputs. Pradhan et al. [10] established a formulation to calculate the roughness coefficient using a dimensional analysis for the meandering channel, incorporating relative depth, width ratio, longitudinal channel slope, and sinuosity as input parameters. However, previous research works were severely constrained due to the data retrieving and storage capabilities. Due to the data revolution, it is significantly easier to collect, extract and process different types of data these days. Cloud-computing platforms offer a source of data with basic preprocessing tools and conversion with high computational power. In recent years, ML-based predictive analysis has been effectively utilized to unravel the complexities of river hydraulics [11]. This study is centered on creating an ML-based regressors for predicting channel roughness coefficient (Manning's n) as a function of a set of external geographic and environmental factors.

4. Methodology

The methodology workflow is divided into three main categories and seven subcategories . The procedure began with the data collection and preprocessing. The datasets are collected from several open sources and spatially joined. In the Exploratory Data Analysis (EDA), the dataset is explored to make a framework for required feature conversion. In Feature Engineering (FE), the variables are transformed and scaled to obtain the best model performance. Once the best feature engineering approach is identified, ML-based regressors are trained, tested, and validated using a set of error matrices. The model performance is further enhanced through the extensive hyperparameter optimization.

4.1. Study Area and Workflow

Figure 1 shows the regional distribution of the USGS HYDRoSWOT sample data points across CONUS. The corresponding predictors are obtained using the NHDPlus v2 Flowline dataset, and global scale raster from GEE.

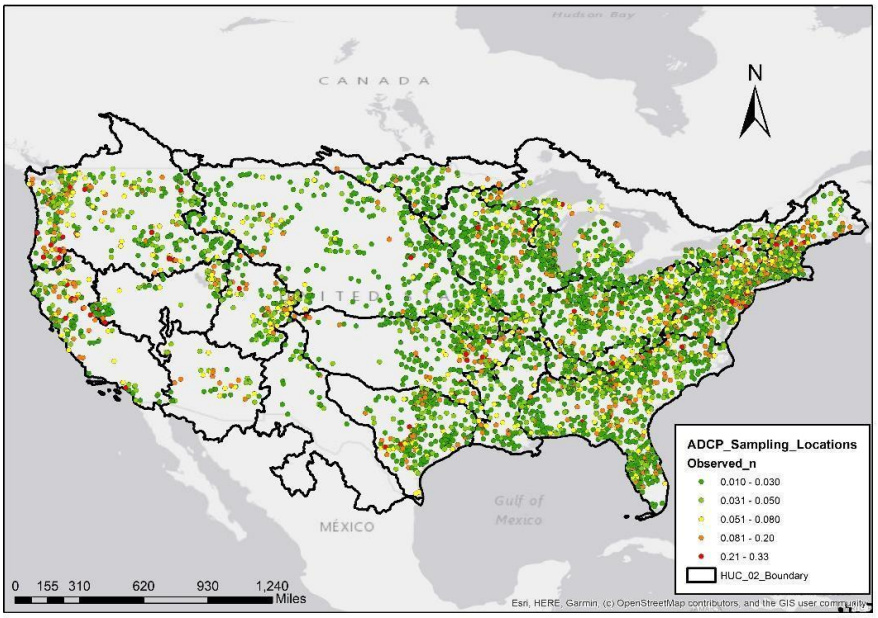


Figure 1. *Entire CONUS area is chosen as the study domain.*

4.2. Data Collection and preprocessing

The raw data are collected (**Table 1**), from the National Hydrography Dataset (NHDPlus V2), the USGS HYDRoSWOT Dataset, satellite images and contributions from many global scale research initiatives. The schematic of data processing is shown in **Figure 2**. The HYDRoSWOT sampling events are used as the reference point for retrieving the velocity (v) and hydraulic radius (R) (**Figure 2**). Eventually, the slope (S) is retrieved from the NHDPlus dataset for calculating the roughness coefficient which is computed analytically using Manning's equation (Equation 1) to prepare the observed/calculated channel roughness coefficient (n). The spatial distribution of the observed n is illustrated in **Figure 2**. With a buffer of 50 m, geospatial approaches are used to spatially join HYDRoSWOT sampling locations with NHDPlus Flowlines. Selected contributing predictors (**Table 1**) along with the observed channel roughness are used to train/test the ML models.

$$n = \frac{1.49\, R^{2/3} S^{1/2}}{v} \tag{1}$$

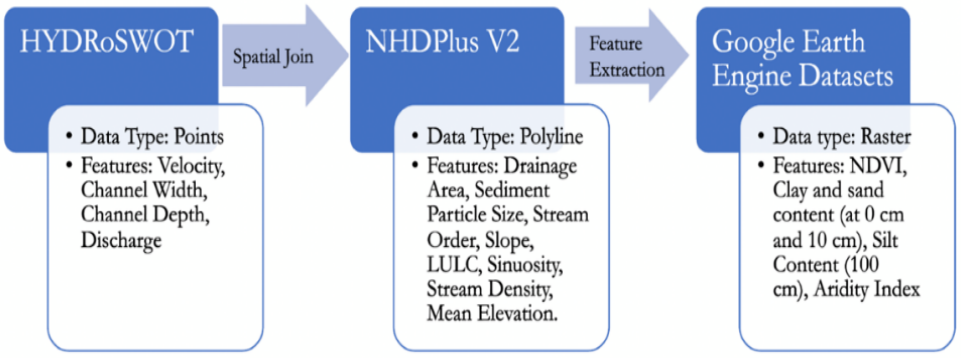


Figure 2. *Schematic of the data preprocessing step.*

The MOD13Q1 V6.1 product provides NDVI which is referred to as the continuity index to the existing National Oceanic and Atmospheric Administration-Advanced Very High-Resolution Radiometer (NOAA-AVHRR) derived NDVI. Sand and clay content data are stored in % (kg/kg) at 6 standard depths (0, 10, 30, 60, 100 and 200 cm) at 250 m resolution. The Aridity Index represents the ratio between precipitation and ET_0. Thus, rainfall exceeds vegetation water demand (aggregated on an annual basis). Global Aridity Index (Global-Aridity_ET_0) Version 2 dataset provides high-resolution global raster climate data, related to evapotranspiration processes and rainfall deficit for potential vegetative growth.

Table 1. *Data sources with the types and resolution*

Data	**Source**	**Data Type**	**Coverage**
Velocity, Channel Width, Channel Depth, Discharge	USGS HYDRoSWOT Canova et. al. 2016 [12]	Point	CONUS
Drainage Area, Stream Order, Slope, Mean Elevation.	National Hydrography Dataset	Polylines	CONUS
LULC, Sinuosity, Stream Density.	Wieczorek et al., 2018 [13]	Polylines	CONUS
Channel bed particle size (D_{50})	Abeshu et al., 2021 [14]	Polylines	CONUS
NDVI	MOD13Q1 V6.1	Raster	Global / 250 m
Sand, Clay, Silt	Hengl et al., 2018 [15]	Raster	Global / 250 m
Aridity Index	Trabucco et al., 2019 [16]	Raster	Global / 250 m

4.3. ML-based regressors

A set of ML-based regressors is chosen for the predictive analysis e.g., (RF), Extreme Gradient Boosting (XGB), K-Nearest Neighbors (KNN) and Multi-layer Perceptron (MLP). Random forest (RF) is an ensemble learning method for regression operated by constructing a collection of multiple decision trees when training the model. A multilayer perceptron (MLP) is a fully connected type of Feed-Forward Neural Network. XGBoost is an extreme gradient boosting regression and is a distributed gradient-boosted decision tree ML algorithm. KNN regression is a non-parametric method that approximates the association between predictors and the target variable by taking the average of the observations in the similar neighborhood based on a distance function.

4.4. Error Matrices

Three standard performance assessment methods to evaluate the predicted roughness and draw conclusion are the coefficient of determination(R^2) Root Mean Square Error (RMSE), and the Mean Average Error (MAE). Root Mean Square Error (RMSE) is the standard deviation of the prediction errors. RMSE is a measure of the spread of the prediction error with respect to the line of best fit. Lower RMSE value reflects better model prediction performance with less error. The third evaluation metric MAE depicts the arithmetic mean of the absolute errors between the predicted and observed roughness coefficient. Like the RMSE, lower values indicate better model performance.

4.5. Hyperparameter optimization and Relative Feature Importance

Hyper-parameters are not directly learnt from the ML models. In scikit-learn they are passed as a set of arguments to the constructor of the ML model classes. Searching framework involves an algorithm (RF/MLP), a parametric space, a method for searching or sampling the candidates of hyperparameters, a cross-validation process and a scoring function. In this study, the Grid Search CV (GSCV) scheme is considered for both RF and MLP. GSCV exhaustively takes all the hyperparameter combinations. The relative feature importance of the predictors is studied by analyzing the Permutation Feature Importance (PFI) technique in the computational domain. In PFI, the impact of shuffling the values of a feature (e.g., NDVI) over the target variable, roughness coefficient is quantified to observe the response in output variables due to the change in input variables. The score of the error matrix (R^2) derived from the observed and predicted values of the channel roughness because of shuffle in the predictors provides the score of relative feature importance.

4.6. Spatial K-folds Cross-Validation (CV)

Spatial K-folds Cross-Validation (CV) is performed for the ML regressors in predicting the channel roughness (**Figure 3**). The entire study domain is clustered into 18 folds (hydrologic regions) as the spatial autocorrelation among the nearby locations may lead to bias and wrong model evaluation if the models are evaluated considering the entire study area only. In the random sampling for train/test split, there is a possibility of taking the cells out in the study domain for the training set which are neighbors to the cells taken out for the test set. Consequently, those features in the train and test set are no longer independent, invalidating the evaluation of the ML classifiers. Therefore, the entire study domain is grouped into several folds to prepare individual train/test sets to compute the error matrix. Finally, the average of all the values of the error matrix is computed to show the model performance. Eighteen different splits are introduced to isolate the test set with an independent fold. In each split, the rest of the seventeen folds are used to train the model.

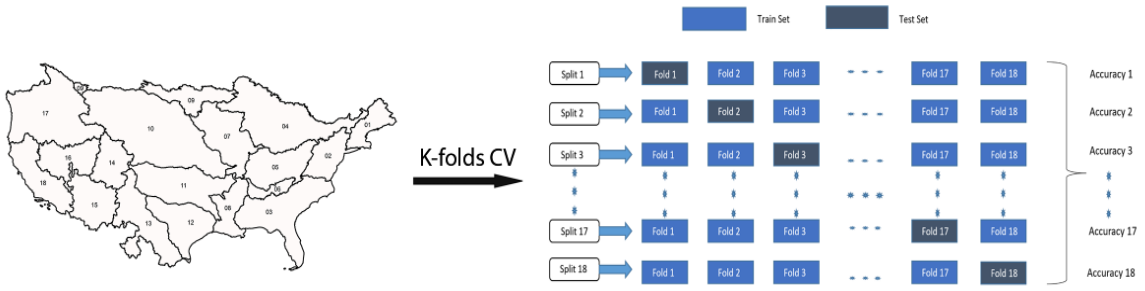


Figure 3. *Spatial K-folds CV with a set of ten train/test splits.*

5. Results and Discussion

5.1. Predicted vs Observed Roughness

After successful training using the derived channel roughness data, the ML models are deployed to predict the roughness for testing. Predicted roughness values are plotted against observed roughness values in a scatter plot to show the visual representation of the best tree and neural network-based regression model, RF and MLP performance in Figure **4**. RF regressor is run with the selected hyperparameters (number of trees = 100, minimum number of samples required to split an internal node = 10, minimum number of samples required to be at a leaf node = 1, number of features to consider when looking for the best split = 'sqrt', maximum depth of the tree = 50, Whether bootstrap samples are used when building trees = True, randomness of the bootstrapping of the samples = 0) after optimization. For MLP the selected hyperparameters are the size of hidden layers = (100, 50), activation function = (ReLU, Sigmoid), solver = Adam, alpha = 0.0001, learning rate = constant. The R^2 values for the training set of the RF and MLP algorithm are 0.734 and 0.672, respectively. Both models show poorer performance with lower R^2 (0.698 and 0.648) in the testing phase with the unseen dataset. R^2 provides a statistical measure of how spread/near the predicted roughness values are to the fitted regression line (black lines) in **Figure 4**, where the observed and predicted roughness values are identical.

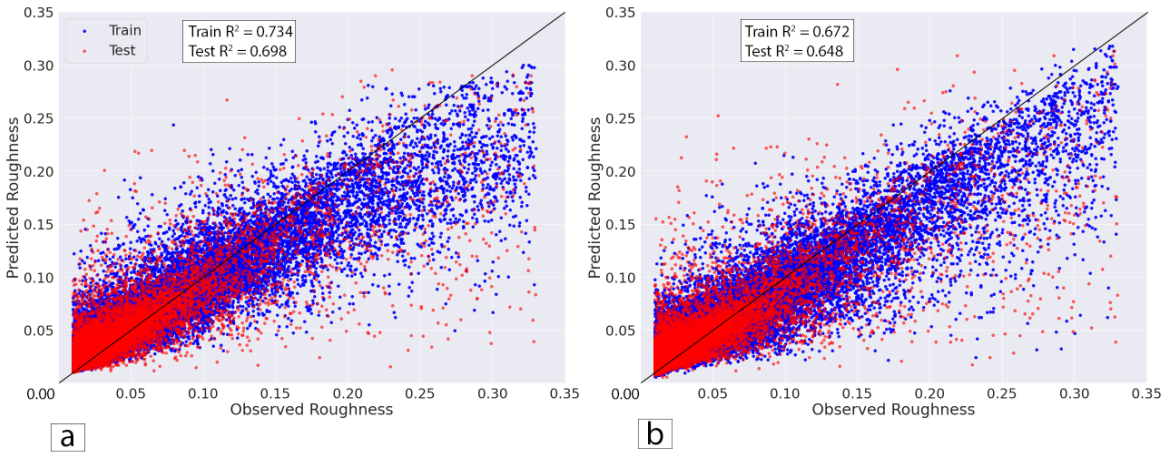


Figure 4. *Scatterplots of the observed and predicted channel roughness from (a) RF and (b) MLP regressor*

A comparative study on the model performance with the XGB and KNN is tabulated (**Table 2**). The RF algorithm outperforms MLP, XGB and KNN in predicting channel roughness with a R^2 score of 0.698 followed by 0.656 (XGB), 0.648 (MLP) and 0.587 (KNN). RMSE and MAE scores of the models also showed satisfactory performance from the models. The computational time required for the MLP is found to be approximately fifteen times higher

than the tree-based regressors, RF and XGB. The computational time for KNN is also higher (approximately five times) with poorer performance compared to all other regressors (**Table 2**).

Table 2. *Performance comparison of ML regressors*

Model	R^2	RMSE	MAE
Random Forest (RF)	0.698	0.007	0.005
Multilayer Perceptron (MLP)	0.648	0.021	0.017
XGBoost (XGB)	0.656	0.011	0.017
K-Nearest Neighbors (KNN)	0.587	0.073	0.088

Model performances are further validated using spatial k-folds CV with 18 hydrologic regions (**Figure 5**). The distribution of the individual model accuracy of the ML regression in the spatial K-folds CV is presented in **Figure 5** using boxplots. Median values of the model accuracy (R^2) are illustrated with the red lines with the values 0.709 (RF), 0.639 (MLP), 0.677 (XGB) and 0.628 (KNN). The range of the accuracy scores of the RF algorithm is 0.618-0.724 and is found to be the best ML regressor among all others. The median score of the RF model is 0.709 and is the highest score followed by the XGB, MLP and KNN. Overall, RF, MLP, and XGB showed satisfactory performance, as the range of the model accuracy of all regressors is 0.506-0.724. However, As the dataset comprises the temporal variability of channel roughness at multiple locations, it may result in overfitting in the model prediction. The model performance may alter with the point locations with a representative unique (e.g., median) roughness coefficient in the train set.

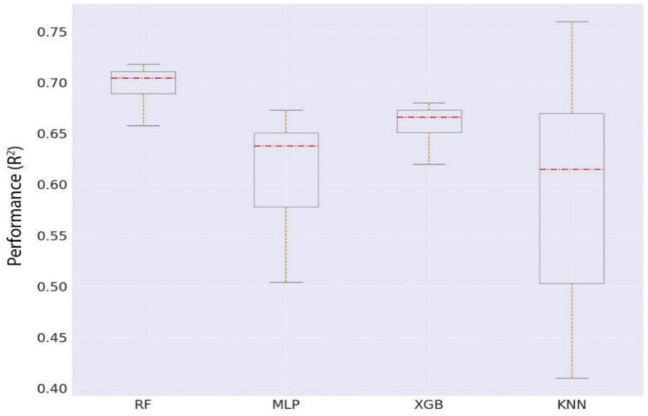


Figure 5. *Distribution of the model accuracies in the spatial CV for the ML regressors*

Spatial variation of the predicted roughness at the HYDRoSWOT locations (a) and the difference between the observed and predicted roughness (b) are illustrated in **Figure 6**. The overall range of the spatially explicit predicted roughness varies from 0.011-0.33. Higher roughness values are found in the southeast and pacific west regions. Comparatively lower values were predicted in the Midwest region, particularly near the Mississippi river. The difference map in **Figure 6(b)** shows the test (red border) and train regions (black border).

The green points indicate the values which are within +1 and –1 standard deviation from the observed roughness coefficient and the red points indicate the points which are more than 1 standard deviation. Overall, the train dataset has better prediction than the test dataset.

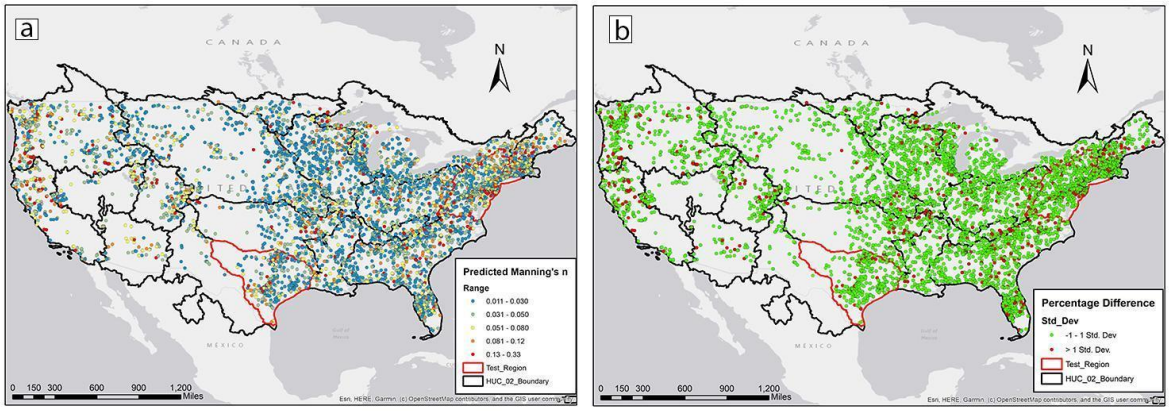


Figure 6. *Spatial map of the predicted channel roughness (a) and error (difference in observed and predicted roughness) (b).*

5.2. Feature Importance

Permutation Feature Importance (PFI) approach is used to determine the impact of the features on the ML-based regression based on the change in the R^2 value as a numerical indicator. From PFI analysis, channel discharge is found to have the highest importance in predicting channel roughness compared to the other features. The importance scores are shown in **Figure 7 (a)** to illustrate the significant response of channel roughness predicted from the ML based regression model due to the change in the predictors. Top five most influential features are overserved river discharge, total drainage area, forest, and agricultural land covers and NDVI. As the channel discharge and velocity shows high temporal variation and bivariate correlation illustrated in the **Figure 7 (b)** with the roughness compared to other features, the shuffle signal of discharge series may lead to highest response in the change in r^2 values i.e., high feature importance.

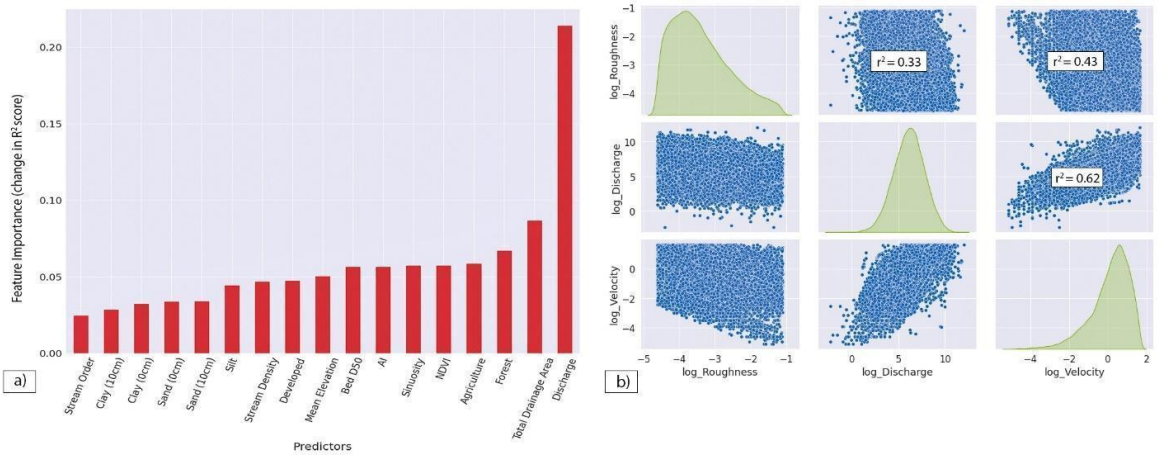


Figure 7. *Rank of influencing factors according to their Feature Importance (a) and bivariate correlation plot of log-transformed discharge, velocity, and roughness (b).*

6. Conclusion

Data-informed ML approaches provide a path to circumvent the complexities of geographic and environmental influencing factors and the computational burden of physics-based models.

In addition, obtaining channel roughness with model calibration is highly expensive and inefficient computationally at a large scale. A novel ML-based framework is developed to predict channel roughness in this study. Several features are retrieved and combined from NHDPlus V2, HYDRoSWOT and satellite imageries. A set of supervised regressors e.g., RF, MLP, XGB and KNN showed satisfactory performance. Spatial K-folds CV was performed to validate the model's performance to generalize to the unseen areas. The performance score for CV showed satisfactory performance for all the spatially segmented blocks (hydrologic regions). In addition, the study reveals the influencing predictors for channel roughness in chronological order through feature importance task. However, data scarcity of unique locations may result in poor model performance in several regions. Temporal variation of channel roughness at multiple locations alleviates the model performance in representing static roughness distribution over the large regions. Including discharge from the HYDRoSWOT dataset increases the model performance while limiting the model applicability in the entire domain. The performance of the models can be improved by including additional parameters with higher resolutions. Model validation with previously calibrated channel roughness at the study locations may enhance the model applicability. Further, other ML regressors such as Gaussian process regression, Bayesian regression and histogram-based gradient boosting can be explored for higher model performance. The proposed framework illustrates a promising ground and potential to deploy the predicted channel roughness to prepare a flood inundation mapping for a larger scale.

7. References

1. Ding, Y; Zhang, X;, He, Z; Lu, C; Bao, S. Sedimentary environment of a dammed lake buried in the modern riverbed of the Yalong River during the Last Glacial Maximum and its implication for fluvial geomorphic evolution. *Geomorphology*. **2021** Apr 1; 378:107588.
2. Kitsikoudis, V; Sidiropoulos, E; Iliadis, L; Hrissanthou, V. A Machine Learning Approach for the Mean Flow Velocity Prediction in Alluvial Channels. *Water Resour Manag*. **2015** Sep 1; 29(12):4379–4395.
3. Zhouyayan, L; Duque, FQ; Grout, T; Bates, B; Demir, I. Comparative Analysis of Performance and Mechanisms of Flood Inundation Map Generation using Height Above Nearest Drainage. **2022 Apr 8**; https://doi.org/10.31223/X59K9G
4. Melancon, A. Machine Learning Classification of Inundation following Hurricane Florence (2018) via L-Band Synthetic Aperture Radar and Ancillary Datasets. Master's Thesis, The University of Alabama in Huntsville; **2021**.
5. Zurqani, HA; Post, CJ; Mikhailova, EA; Cope, MP; Allen, JS; Lytle, BA. Evaluating the integrity of forested riparian buffers over a large area using LiDAR data and Google Earth Engine. Sci Rep. **2020 Aug 24**; 10(1):14096.
6. Mohanta A; Patra, KC; Sahoo, BB. Anticipate Manning's Coefficient in Meandering Compound Channels. Hydrology. **2018** Sep; 5(3):47.
7. Ren, H; Hou, Z; Duan, Z; Song, X; Perkins, WA; Richmond, MC; et al. Spatial Mapping of Riverbed Grain-Size Distribution Using Machine Learning. Front Water. **23 November 2020**. 2:551627 https://doi.org/10.3389/frwa.2020.551627
8. Haddadchi A; Booker, DJ; Measures, RJ. Predicting riverbed substrate cover proportions across New Zealand. CATENA. **2018 Apr 1**; 163:130–46.
9. Overstreet, BT; Legleiter, CJ. Characterizing channel change along a multithread gravel-bed river using random forest image classification. **2012 Dec 1**; 2012:H31E-1172.
10. Pradhan, A; Khatua, KK. Assessment of Roughness Coefficient for Meandering Compound Channels. *KSCE J Civ Eng*. **2018 May 1**; 22(5):2010–22.

11. Manessa, MDM; Kanno, A; Sekine, M; Haidar, M; Yamamoto, K; Imai, T; et al. Satellite-Derived Bathymetry Using Random Forest Algorithm And Worldview-2 Imagery. *Geoplanning J Geomat Plan.* **2016 Oct 25**; 3(2):117.
12. Canova, M.G; Fulton J. W.; and Bjerklie D.M. USGS HYDRoacoustic dataset in support of the Surface Water Oceanographic Topography satellite mission (HYDRoSWOT): U.S. Geological Survey data release, **2016**, http://dx.doi.org/10.5066/F7D798H6
13. Schwarz, GE.; Jackson, SE.; Wieczorek, ME. Select Attributes for NHDPlus Version 2.1 Reach Catchments and Modified Network Routed Upstream Watersheds for the Conterminous United States. *U.S. Geological Survey*; **Published 2018, Revised 2021**, https://doi.org/10.5066/F7765D7V
14. Abeshu, G.W.; Li, H.Y.; Zhu, Z.; Tan, Z.; Leung, LR. Median bed-material sediment particle size across rivers in the contiguous US. Earth Syst Sci Data. **2022 Feb 24**; 14(2):929–42.
15. Tomislav, H.; Sand content in % (kg / kg) at 6 standard depths (0, 10, 30, 60, 100 and 200 cm) at 250 m resolution (v0.2). **2018**; [Data set] Zenodo. https://doi.org/10.5281/zenodo.2525662
16. Trabucco, A.; Zomer, R; Global Aridity Index and Potential Evapotranspiration (ET0) Climate Database v2. **2019**; [Dataset] figshare. https://doi.org/10.6084/m9.figshare.7504448.v3

Chapter 5

QuiCFIM, a quick GIS-based combined flood inundation mapping framework

Ankit Ghanghas[1], Mark Wang[2], and Mohamed Abdelkader[3]
[1] Purdue University; *aghangha@purdue.edu*
[2]The University of Texas at Austin; *mark.wang@utexas.edu*
[3] Stevens Institute of Technology; *mabdelka@stevens.edu*

Academic Advisors: Venkatesh Merwade, *Purdue University*[1]; Paola Passalacqua, *The University of Texas at Austin*[2]; Marouane Temimi, *Stevens Institute of Technology*[3]

Summer Institute Theme Advisors: Celso Ferreira, George Mason University*, cferrei3@gmu.edu;* Kyle Mandli, Columbia University, *kyle.mandli@columbia.edu*

Abstract: Coastal regions are prone to compound flooding events, which are primarily driven by a combination of extreme precipitation, storm surge, and high tides. With a rise in the likelihood of compound flooding due to climate change and anthropogenic activities, it has become essential to accurately estimate the depth and extent of compound flooding in these regions. The accuracy of the flood depth and extent in turn dictates the selection of mitigation strategies and emergency measures needed to reduce losses. This study introduces a rapid and simple GIS-based method called Quick Combined Flood Inundation Mapping (QuiCFIM) for compound fluvial-coastal Flood Inundation Mapping (FIM). QuiCFIM combines inland and coastal FIM to enhance flood depth and extent estimations in coastal regions. This kind of FIM post-processing is necessary in continental-scale water forecasting operations such as the NextGen National Water Model. To this end, the method delineates regions as hydrologic, coastal or transition zones and estimates the inundation depth in each zone. The study utilizes the HAND-SRC based inland FIM, and c-HAND based coastal FIM as inputs to QuiCFIM. However, QuiCFIM is model-agnostic and can accept inundation results from any model. This is invaluable within the NextGen framework where outputs from multiple models will be chosen and combined to produce accurate end products. The method's accuracy was then tested for the compound flooding event during Hurricane Florence in 2018. The estimated flood depth and extent were validated against ground observations and satellite images, respectively. The results establish that QuiCFIM is capable of enhancing flood depth estimation in transition zones and providing a better estimate of the spatial coverage of the inundated areas. However, it should also be noted that the accuracy of QuiCFIM relies on the quality of input FIM.

1. Motivation

Coastal regions in the United States are home to about 40% of its population, including around 8.6 million Americans living in areas which are susceptible to coastal flooding. It is estimated that by 2050, a majority of US coastal areas will be threatened by 30 or more days of flooding every year [1]. The interior region (inland) flooding is mostly driven by extreme precipitation, resulting in pluvial and fluvial floods. On the other hand, floods in coastal regions are more complex and can be caused by extreme rainfall, storm surge, or a combination of the two occurring in close succession or simultaneously, resulting in compound flooding [2–5]. With a rise in sea level [6] and the increasing probability of tropical storms bringing in heavy precipitation and high storm surges [4], the need to understand compound flooding in coastal zones is higher than ever before.

In recent years, large-scale operational flood forecasting systems have been significantly improved by the development of better hydrological models. For instance, recent advancements in the National Water Model (NWM) have significantly improved the nation's flood resilience and mitigation capabilities in flood-prone inland regions. However, the NWM's ability to provide accurate flood inundation maps in coastal areas is limited. The NWM does not account for backwater in streams during periods of high storm surge, and the HAND-SRC (Height Above Nearest Drainage – Synthetic Rating Curves) [7], which are used to convert NWM discharge to flood inundation maps, are not reliable in coastal areas [8–10]. In contrast, NOAA's (National Oceanic and Atmospheric Administration) SLOSH model (Sea, Lake and Overland Surge from Hurricanes) can provide accurate and reliable flood maps based on storm surge. However, these maps do not incorporate fluvial flooding which occurs as a result of overtopped riverbanks during heavy rainfall. While efforts are under way to couple operational coastal models to the NWM, generating flood inundation maps resulting from different physical processes and operational systems remains a challenge in transition areas.

In accordance with NOAA's Office of Water Prediction (OWP) strategic vision, the development of Flood Inundation Mapping (FIM) tools would enhance the forecast of flood inundation in coastal zones by incorporating FIM from terrestrial and coastal models. Thus, addressing FIM for compound flooding events would support forecasting systems and provide more accurate information for decision-making and emergency management.

2. Objectives and Scope

This study aims to combine the best of both coastal and fluvial flood modeling worlds. The study proposes a novel and simple method called Quick Combined Flood Inundation Mapping (QuiCFIM) which delineates regions as hydrologic, coastal, or transition zones. Then it generates combined flood inundation maps (FIM) from precomputed and independent fluvial and coastal flood maps at any given timestep. Subsequently, the study validates the combined flood maps generated for Hurricane Florence in the Neuse River basin (**Figure 1**) by comparing the extent with the flood extent obtained from satellite imagery and comparing the inundation depth with the high water marks (HWM) for inland regions and with peak water levels (PWL) for coastal regions.

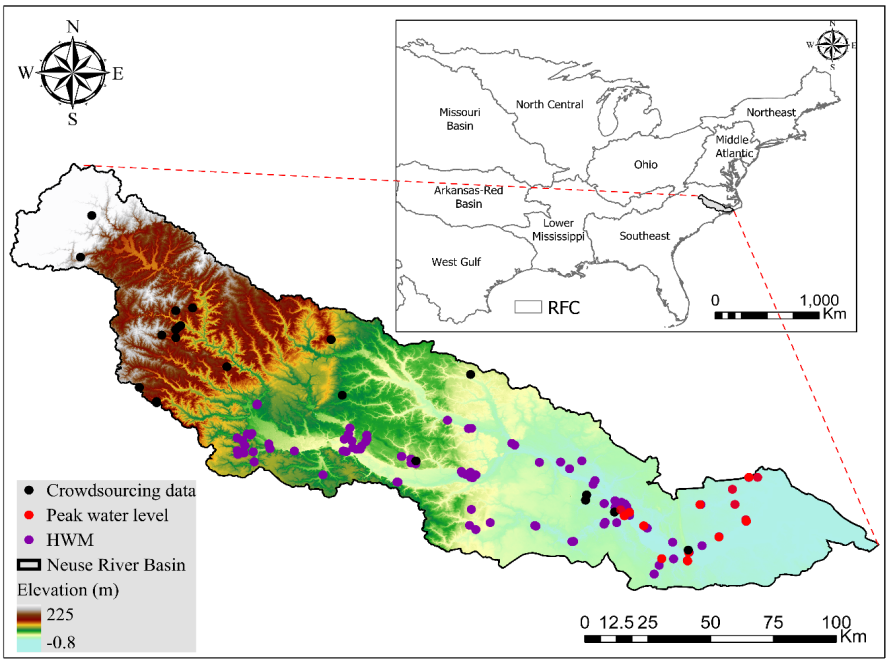


Figure 1. *Study area map.*

Hurricane Florence made landfall on the southeast coast of North Carolina at 11:15 UTC on September 14th, 2018. In parts of eastern North Carolina, the storm caused significant freshwater flooding as well as storm surge flooding. In North Carolina, the maximum storm surge inundation heights caused by Florence are estimated to be 8 to 11 feet above ground level along the Neuse River and its tributaries. In the southeastern part of North Carolina, the rainfall totals from Florence exceeded 20 inches. The peak rainfall and flooding occurred on September 19th, 2018. Although peak surge and peak flooding did not occur simultaneously, the combined flood did occur on the day of the storm's landfall when the precipitation rate was 2 inches per hour [11].

3. Previous Studies

Coastal communities have been vulnerable to flooding caused by the combined effects of excessive rainfall and storm surge inundation [12]. Coastal surges and intense rainfall interact in complex ways and have a significant impact on overland flooding. Several studies have investigated such events by employing statistical [5,13,14] and numerical [12,15] modeling approaches.

In the context of flood risk assessments, statistical methods have demonstrated that the relationship between extreme rainfall and storm surge can be statistically significant and that the interaction between terrestrial and coastal flooding can be critical in determining flood risk along the coast [5]. The combined hazard and exposure of storm surge and inland flooding has been quantified using deterministic models. Thompson and Frazier (2014), for instance, incorporated deterministic and probabilistic models to identify flood hazard extents. The study integrated storm surge, rainfall-runoff, and sea-level rise simulations within a geographic information system framework to determine the effects of compound flooding scenarios [15]. Bilskie and Hagen (2018) developed a novel approach to defining flood zones in coastal areas by combining rainfall excess and storm surge simulations. The proposed approach provided insight into estimating water levels and supporting flood risk assessments in flood transition zones [12]. Previous studies have demonstrated that incorporating hydrologic and hydrodynamic models results in a more accurate estimation of water depths and compound flooding extent. Nevertheless, coupling such models may not always be feasible or straightforward. This study proposes a simple method for combining inland and coastal inundation maps and estimating water levels and flood extent for compound flooding events.

4. Methodology

The proposed combined flood mapping scheme is versatile and accepts independent inland and coastal flood inundation maps generated by any suitable method, such as HAND-SRC, HEC-RAS, ADCIRC, SCHISM, or any other method. However, the current scheme is limited to having independent inland and coastal maps, which implies that the inland map should be a result of hydrological dominant process and the coastal map be a result of storm surge dominant process. In the current study, inland flood maps were produced using the HAND-SRC method applied to NWM's retrospective streamflow, and coastal flood maps were produced using the c-HAND method. The combined flood maps were then generated with QuiCFIM using the HAND-SRC inland flood map and the c-HAND coastal flood map and then validated against high water marks and Sentinel-1 Synthetic Aperture Radar (SAR) imagery. The following sections describe the methodology for each of these steps.

4.1 Inland Flood Mapping

The inland flood maps were generated by applying the HAND-SRC method to NWM retrospective version 2.1 modeled streamflow discharges using the HAND-SRC method [7,8]. HAND-SRC is a conjunction of using HAND based Relative Elevation Model (REM) for

topography and using REM based rating curves known as the Synthetic Rating Curves (SRC) to convert NWM discharge to a stage height. HAND value for any grid cell on the digital elevation model (DEM) is simply the height above the nearest stream grid cell it drains to. Given a stage height *d* at any stream cell, all the grid cells that drain to this cell and with a HAND value less than *d* will be flooded. The difference between the HAND value and *d* is the inundation depth at that grid cell. SRC is used to convert the NWM discharge to stage height, which is subsequently used to calculate flood inundation and depth using HAND. Detailed information about HAND-SRC and its implementation can be found in Zheng et al (2018) and Garousi-Nejad et al (2019) [7,8].

4.2 Coastal Flood Mapping

Coastal flood inundation was generated with a modification of the HAND method called c-HAND [16]. This is an approximate GIS-based approach to coastal flooding. combined flood mapping as described below can use inundation generated from any model— c-HAND was chosen due to its computational efficiency. The ocean water surface elevation is assumed to be a horizontal surface and is extrapolated to the terrain. A static equilibrium assumption was made approximating coastal inundation extent and depth, which allowed for rapid calculation. The ocean surface elevation and terrain elevation are both known relative to a common vertical datum. c-HAND is different from a simple bathtub approach in that it has an additional requirement of ocean connectivity [17]. In order for a DEM cell to have coastal inundation, it must satisfy two constraints: (1) it must have a terrain elevation lower than the ocean water level, and (2) it must be hydrologically connected to the ocean. For condition (2), the coastal inundation is separated into discrete connected regions. All regions not hydrologically connected to the ocean are eliminated, leaving one discrete region inundated at the coast.

4.3 Combined Flood Mapping

The study proposes a simple ratio-based method to generate combined flood maps for coastal regions. For regions where storm surge inundation depth (S) is greater than rainfall excess inundation depth (R), QuiCFIM classifies these regions as Coastal Zones (C), and the combined inundation depth (RS) equals the inundation depth from storm surge S for the region. Riverine flow in coastal areas typically follows a subcritical regime and assuming the storm surge region as an infinite reservoir, the intersection of river reach with a coastal flooded area would result in an M1 water profile. This M1 profile forms the basis of assuming RS equal to S for the coastal region. The region where the ratio of S to R is less than α, where α is the transition parameter, is classified as Hydrologic Zone (H) and RS equals R for this region, assuming that the effect of storm surge on compound flooding in Hydrologic Regions is significantly less [12]. QuiCFIM classifies the region in between the Coastal and Hydrologic Zone as the Transition Zone (T), where $\alpha < \frac{S}{R} \leq 1$. RS in the Transition zone is defined as per Equation (1) below.

$$RS = \frac{S}{R} * max(R,S) + (1 - \frac{S}{R}) * (R + S) \quad (1)$$

Here max (R, S) represents the maximum inundation depth in the overlapping region which presents the lower bound for RS and (R+S) represents the super positioned [13] inundation depth and presents the upper bound for RS. RS in the transition zone is a weighted average of the maximum overlapping depths and super-positioned inundation depth. This prevents the overprediction of inundation depth [12]. To summarize, the combined flood inundation depth is calculated using Equation (2) below where "IF" delineates C, T, H and the value preceding

"IF" condition is used in case the condition is satisfied. The study analyzes and presents the best values for α in subsequent sections.

$$RS = \{IF\,[S/R > 1, S]\ ;\ IF[S/R \leq \alpha, R]\ ;\ IF[\alpha < S/R \leq 1, (S/R \times max(R,S) + (1 - S \quad (2)$$

4.4 Evaluation Metrics

The Percent Bias (PB) was employed to assess the degree of agreement and variation between the estimated and observed water depths. The PB (Equation (3)) was calculated using water depths from high water marks and inundation depths generated for different storm surges and inland flooding conditions during Hurricane Florence. The PB is used to investigate whether the estimated water depths are being under- or over-predicted compared to the ground observations.

$$PB = 100 \times \left(H_e - H_o\right)/H_o \quad (3)$$

On the other hand, the Proportion Correct (PC), the Bias Ratio (B), the Probability of Detection (POD), and the Critical Success Index (CSI) statistics were used to evaluate the accuracy of the generated combined flood maps to estimate inundation extent. Here, the flood extent verification metrics were calculated against flood extent from remote sensing observations. The validation process consists of generating a confusion matrix that represents the status of the grid cells for the estimated and reference flood maps. For the estimated flood map, the inundated and non-inundated grid cells are indicated by Positive and Negative, respectively. Further, True and False were used to describe whether the estimated grid cell status matches with the reference grid cell. Thus, the agreement between the estimated and the reference inundation extent can be assessed using the following equations:

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

$$B = \frac{TP+FP}{TP+FN} \quad (5)$$

$$POD = \frac{TP}{TP+FN} \quad (6)$$

$$CSI = \frac{TP}{TP+FP+FN} \quad (7)$$

TP represent inundated pixel both in reference map and estimated map, TN denote dry pixel in both reference map and estimated results, FP represents inundated grid in the estimated map but not in reference map, and FN denote inundated pixel in reference map but not in estimated map [18]. The reference data that is used to validate the flooding extent is Sentinel-1 SAR imagery. The flooded areas were identified by applying the method of radiometric threshold [9]. The method consists of discriminating the "flood" and "non-flood" pixels by means of a representative backscattering coefficient threshold value.

5. Results

The study uses Hurricane Florence in the Neuse River Basin as a case study and compares the flood depth and inundation extent corresponding to the peak surge event during Hurricane Florence, where water levels were affected by both storm surge and inland flooding. The evaluation metrics mentioned in section 4.4 were applied to evaluate the accuracy of the proposed combined FIM approach. Validation is intended to examine if the compound FIM enhances the estimation of flood depth and extent in coastal transition zones.

5.1 Flood extent sensitivity analysis

The delineation of hydrologic and transition flood zones in QuiCFIM is based on the transition parameter α. Thus, it is vital to assess the key factor/transition parameter (α) influencing the variation of the flood zones' extents. **Figure 2** presents the variation of flood zone percentage to the change in α from 0 to 1. An increase in alpha reduces the extent of the transition zone with 0 leading to max transition zone extent and 1 resulting in zero transition zone extent. Subsequently, the higher transition zone extent results in a greater inundation extent. The current study area exhibits only a small variation in flood zones for alpha greater than 0.4. However, it should be noted that the α may vary based on the characteristics and location of the study area. Thus, the obtained results indicate that the α can be used as a calibration parameter for QuiCFIM to develop a more sophisticated and accurate inundation model.

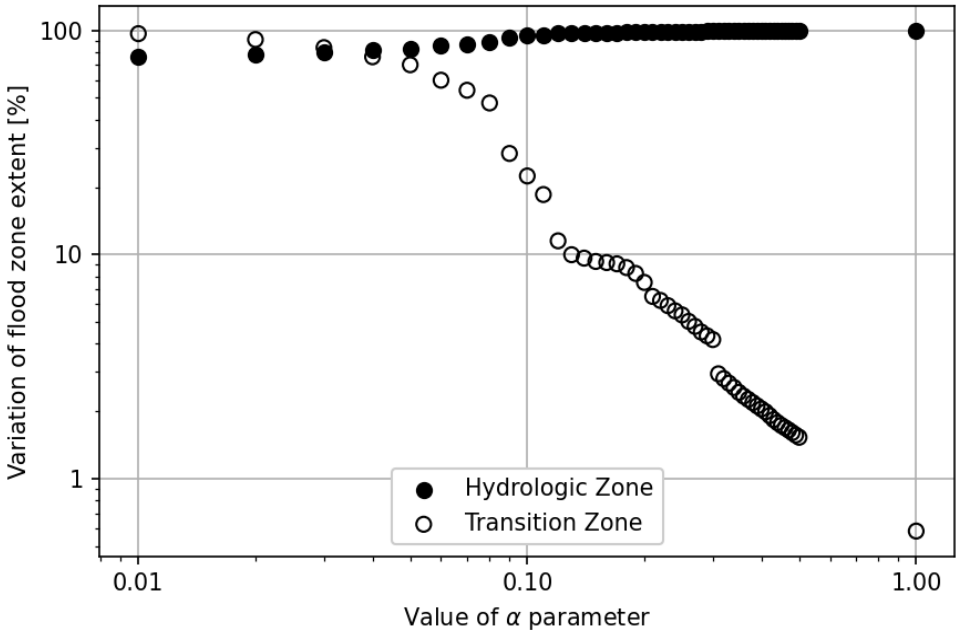


Figure 2. *Variation in the extent of flood zones as a function of the transition parameter*

5.2 Assessment of flood depth

Figure 3a compares the water depth for inland and combined FIM to the HWM and shows significant improvement in PB when using combined FIM instead of inland FIM. However, PB does not indicate any clear patterns of over- or underestimation. It is important to note that the combined FIM achieved lower PB values and a narrower spread of the PB. Thus, the proposed combined FIM method is capable of reducing the error associated with estimating flood depth in coastal transition zones using only riverine FIM.

On the other hand, the comparison of water depth from coastal and combined FIM to the PWL observations in coastal regions (**Figure 3b**) does not present significant variation in PB.

Both the coastal and combined FIM overestimated the observed PWL observation with similar variations in the PB values. The inundation in coastal areas is caused by storm surges and water levels inherited from the coastal model dominate, which results in little or no improvement in PB when using combined FIM in place of Coastal FIM.

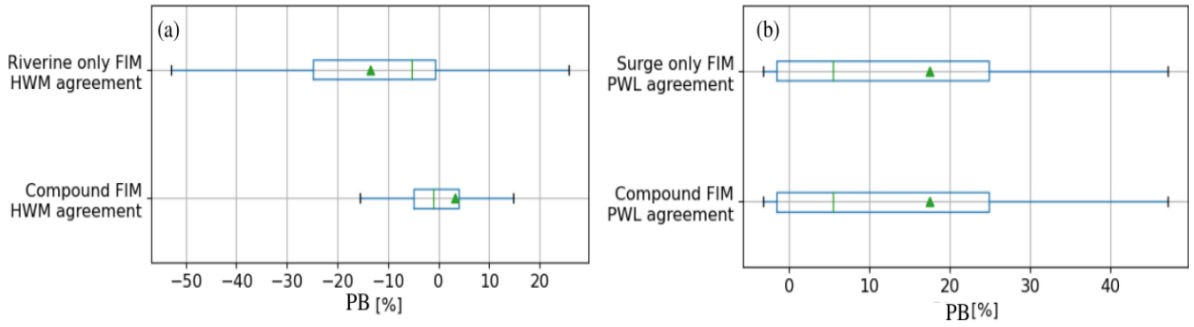


Figure 3. *a and b Comparison of flood depths*

5.3 Assessment of flood extents

Table 1 summarizes the performance of flood extent estimation using coastal, inland, and compound FIM when compared to SAR derived flood extent. The results show that the compound FIM produced the highest agreement between the estimated and the SAR inundation map. It is important to note that for the studied event, flooding caused by storm surge was more significant which resulted in coastal FIM outperforming the inland FIM. The implementation of the compound FIM further results in a more accurate flood extent.

Table 1. *Metrics for assessing the flood extent. Bolded results indicate higher agreement with the reference data*

Metrics	Coastal FIM	Inland FIM	Compound FIM
Proportion Correct	**0.94**	0.91	**0.94**
Bias Ratio	**1.37**	0.09	1.4
Probability of Detection	0.83	0.04	**0.84**
Critical Success Index	0.53	0.04	**0.54**

The PC results indicate that the applied FIM methods concur with the reference inundation map. However, the PC metric includes both TP and TN pixels and does not accurately reflect the capability of the compound method to detect flooding areas, as the study area is dominated by dry pixels (True Negatives). On the other hand, analysis of the bias ratio indicates that the coastal FIM and compound FIM overestimate the extent of inundation (B>1). It was also observed that the compound FIM has no significant effect on POD and CSI metrics. This lack of significant effect can be attributed to the fact that the case study event did not result in a significantly large transition zone and the compound FIM predominantly mimics the better of the two Coastal or Inland FIM.

6. Conclusion

QuiCFIM is model-agnostic and can accept inundation results from any model, numerical, conceptual, or empirical, so long as it provides inundation extent and depth across a DEM. The method's independence from any particular inundation modeling technique makes it a good candidate for post-processing large scale results from multiple models with standard outputs. One candidate for such an application is the NextGen framework being developed by the National Water Center, in which the Basic Model Interface (BMI) is implemented to model each step of the hydrologic cycle in a standard way. This study's combined flood mapping technique is attractive within the NextGen framework because it can rapidly generate approximate combined flood inundation maps by post-processing simulation results from HAND, HEC-RAS, SCHISM and any other inundation model that is implemented.

Supplementary Materials: https://github.com/ankitghanghas/flood_are_here_si_2022

References

1. Kopp, R.E.; Horton, R.M.; Little, C.M.; Mitrovica, J.X.; Oppenheimer, M.; Rasmussen, D.J.; Strauss, B.H.; Tebaldi, C. Probabilistic 21st and 22nd Century Sea-level Projections at a Global Network of Tide-gauge Sites. *Earth's Futur.* **2014**, *2*, 383–406, doi:10.1002/2014EF000239.
2. Hunt, J.G.R. Inland and Coastal Flooding: Developments in Prediction and Prevention. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **2005**, *363*, 1475–1491, doi:10.1098/RSTA.2005.1580.
3. Ray, T.; Asce, M.; Stepinski, E.; Sebastian, A.; Bedient, P.B.; Asce, F. Dynamic Modeling of Storm Surge and Inland Flooding in a Texas Coastal Floodplain. *J. Hydraul. Eng.* **2011**, *137*, 1103–1110, doi:10.1061/(ASCE)HY.1943-7900.0000398.
4. Wahl, T.; Jain, S.; Bender, J.; Meyers, S.D.; Luther, M.E. Increasing Risk of Compound Flooding from Storm Surge and Rainfall for Major US Cities. *Nat. Clim. Chang.* **2015**, *5*, 1093–1097, doi:10.1038/nclimate2736.
5. Zheng, F.; Westra, S.; Sisson, S.A. Quantifying the Dependence between Extreme Rainfall and Storm Surge in the Coastal Zone. *J. Hydrol.* **2013**, *505*, 172–187, doi:10.1016/J.JHYDROL.2013.09.054.
6. 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 252* **2002**, *25*, 149–164, doi:10.1007/BF02691304.
7. Zheng, X.; Tarboton, D.G.; Maidment, D.R.; Liu, Y.Y.; Passalacqua, P. River Channel Geometry and Rating Curve Estimation Using Height above the Nearest Drainage. *JAWRA J. Am. Water Resour. Assoc.* **2018**, *54*, 785–806, doi:10.1111/1752-1688.12661.
8. Garousi-Nejad, I.; Tarboton, D.G.; Aboutalebi, M.; Torres-Rua, A.F. Terrain Analysis Enhancements to the Height Above Nearest Drainage Flood Inundation Mapping Method. *Water Resour. Res.* **2019**, *55*, 7983–8009, doi:10.1029/2019WR024837.
9. Ghanghas, A.; Dey, S.; Merwade, V. Evaluating the Reliability of Synthetic Rating Curves for Continental Scale Flood Mapping. *J. Hydrol.* **2022**, *606*, 127470, doi:10.1016/J.JHYDROL.2022.127470.
10. Michael Johnson, J.; Munasinghe, D.; Eyelade, D.; Cohen, S. An Integrated Evaluation of the National Water Model (NWM)-Height above Nearest Drainage (HAND) Flood Mapping Methodology. *Nat. Hazards Earth Syst. Sci.* **2019**, *19*, 2405–2420, doi:10.5194/NHESS-19-2405-2019.
11. 2022 Atlantic Hurricane Season Available online: https://www.nhc.noaa.gov/data/tcr/

12. Bilskie, M. V.; Hagen, S.C. Defining Flood Zone Transitions in Low-Gradient Coastal Regions. *Geophys. Res. Lett.* **2018**, *45*, 2761–2770, doi:10.1002/2018GL077524.
13. Coles, S.G.; Tawn, J.A. Modelling Extreme Multivariate Events. *J. R. Stat. Soc. Ser. B* **1991**, *53*, 377–392, doi:10.1111/J.2517-6161.1991.TB01830.X.
14. Samuels, P.G.; Burt, N. A New Joint Probability Appraisal of Flood Risk. *Proc. Inst. Civ. Eng. Water Marit. Eng.* **2002**, *154*, 109–115, doi:10.1680/WAME.2002.154.2.109.
15. Thompson, C.M.; Frazier, T.G. Deterministic and Probabilistic Flood Modeling for Contemporary and Future Coastal and Inland Precipitation Inundation. *Appl. Geogr.* **2014**, *50*, 1–14, doi:10.1016/J.APGEOG.2014.01.013.
16. Wang, M.K. Near Real-Time Coastal Flood Mapping. Master's Thesis. *University of Texas at Austin*. **2022**.
17. Rucker, C.A.; Tull, N.; Dietrich, J.C.; Langan, T.E.; Mitasova, H.; Blanton, B.O.; Fleming, J.G.; Luettich, R.A. Downscaling of Real-Time Coastal Flooding Predictions for Decision Support. *Nat. Hazards* **2021**, *107*, 1341–1369, doi:10.1007/S11069-021-04634-8.
18. Cha, Y.; Park, S.S.; Kim, K.; Byeon, M.; Stow, C.A. Water and life from snow: A trillion dollar science question. *Water Resources Research*. **2014**, 5375–5377, doi:10.1002/2013WR014979.Reply.

Chapter 6

Coupling coastal and hydrologic models through BMI and Next Generation National Water Model Framework in low gradient coastal regions of Galveston Bay, Texas, USA

Ebrahim Hamidi[1], Hart Henrichsen[2], Abbie Sandquist[3], and Hongyuan Zhang[4]

[1] University of Alabama; *shamidi1@crimson.ua.edu*

[2] Brigham Young University; *hart@byu.net*

[3] University of Nevada Reno; *asandquist@nevada.unr.edu*

[4] Coastal Carolina University; *hzhang@coastal.edu*

Academic Advisors: Hamed Moftakhari, *University of Alabama*[1]; Daniel Ames, *Brigham Young University*[2]; Scott Allen, *University of Nevada Reno*[3]; Shaowu Bao, *Coastal Carolina University*[4]

Summer Institute Theme Advisors: Celso Ferreira, George Mason University, *cferrei3@gmu.edu,* Kyle T. Mandli, Columbia University, *kyle.mandli@columbia.edu*

Abstract: It is important to understand flooding in highly populated coastal regions, especially as the severity of extreme flood events are projected to increase. Coupling inland and coastal models can help us better represent flooding in coastal regions. This study develops the first coastal BMI to couple a coastal model (GeoClaw) with a Next Generation National Water Model (Ngen NWM) hydrologic model to expand the initial capability of the Ngen NWM to interact with coastal models. The coupling process is demonstrated for Hurricanes Harvey and Ike in a watershed that terminates in Galveston Bay. A GeoClaw model has been successfully run through the Ngen Framework with the GeoClaw BMI. The coupling process has been successfully documented, and demonstrated with data outside the Ngen Framework. Additional analysis is dependent upon further development of the Ngen Framework resources and a GeoClaw plugin to allow the model to accept time-series discharge data.

1. Motivation

Low gradient coastal regions, such as the Mid-Atlantic Coast and the South Atlantic Gulf Coast, are susceptible to damaging and catastrophic flooding from both large overland precipitation events and storm surges from tropical storms and hurricanes. Understanding what areas along the coast may be vulnerable to each type of flooding, or compound flooding from both coastal surge and intense rainfall, under different storm conditions is important to inform local decision makers and flood response teams about potential flood risk in those locations.

Recent studies suggest that sea level rise will lead to a greater chance of extreme flooding in the coming decades, doubling every five years into the future [1]. Recent studies have also suggested that tropical cyclones will have more intense rainfall [2]. Combined projected effects of sea level rise and increased rainfall from extreme storms in coastal watersheds amplifies the importance of modeling coastal flooding as a coupled process.

The Next Generation National Water Model Framework (Ngen Framework), a modular model engine, is designed for coupling different models. The Ngen Framework allows domain

scientists to run various hydrological models on small catchments across the United States. Each model can interact with the Ngen Framework through a Basic Model Interface (BMI). The BMI creates standardized variable names and functions to communicate between models and with the Ngen Framework. BMIs have been built for several hydrological models, including the Conceptual Functional Equivalent model (CFE), which is roughly equivalent to the model run by the existing National Water Model [3]. While multiple models can run in the Ngen Framework, there is currently no BMI for a model that includes ocean hydrodynamics. Such a coastal model is essential to consider hydrodynamic and atmospheric forces and to simulate storm surge. For coastal flood events, flow from rivers impacts water levels near shore, and should be included in coastal models. Likewise, storm surge will backup riverine water levels and impact hydraulic routing models. Therefore, it is necessary to run coupled models to more accurately represent water levels in low gradient coastal regions experiencing compound flooding. With the generation of a coastal BMI, a coupled approach to coastal flooding can be created within the Ngen Framework.

2. Objectives and Scope

Prior to this study, the lack of a BMI for a coastal model represented a gap in the functionality of the Next Generation National Water Model (Ngen NWM) for coastal regions. We developed a novel BMI for the GeoClaw coastal modeling program able to run within the Ngen Framework. This enables one-way coupling between inland hydrologic models, such as CFE, and GeoClaw coastal models. The GeoClaw BMI and coupling ability can improve the performance of the Ngen NWM in low-gradient coastal regions under flooding conditions.

The scope of this study is limited to evaluating watersheds that drain into Galveston Bay for two storm events: Hurricane Harvey and Hurricane Ike. We include comparisons of the surface water with and without streamflow discharge in the GeoClaw coastal model. In the future, the GeoClaw BMI could be run in other coastal regions, with other storms, or with different inland hydrologic model outputs within the Ngen Framework.

3. Previous Studies

In 2012, a new term "compound event" was first introduced in the Intergovernmental Panel on Climate Change (IPCC) report [4]. It means a combination of events that may not themselves be extremes but lead to an extreme event or impact when combined. Coastal flooding can be an example of a compound event. Both inland fluvial and coastal processes contribute to the occurrence of coastal flooding. The number of coastal compound floods has increased significantly in the past few decades [5-7]. This new view of coastal compound flooding motivated a series of studies about coupled coastal flood modeling, which involves coupling inland rainfall runoff models with coastal surge models. Inland hydrologic models focus on the estimation of water fluxes through different hydrological processes. This method is well-suited for catchment hydrology. Coastal models mainly focus on the detailed water flow at one-, two-, or three-dimensions, which is necessary to model storm surge and coastal hydrodynamics. Running a coastal model coupled with a hydrological model can provide a better representation of water levels in areas susceptible to compound coastal flooding.

The major solution to couple coastal models with inland river systems is to set the upstream boundary condition of the coastal model according to river channel discharge values. The discharge values can be obtained either from observation gauges or from an inland hydrologic model. This one-directional data-passing process is referred to as one-way coupling [8]. Currently, the one-way coupling method for coastal modeling is dominant [9,10].

4. Methodology

4.1 Study Area

We tested the coastal model coupling process for watersheds near Galveston Bay in Texas, USA. This area is selected as it has a low elevation gradient near the coast, it is highly populated, and it has been affected by numerous extreme storms including Hurricane Harvey in 2017 and Hurricane Ike in 2008. Hurricane Ike had extreme storm surge levels, and Hurricane Harvey had more extreme rainfall amounts. We evaluate the relative areas flooded by storm surge, rainfall, and compound flooding under both storm conditions.

Three Galveston Bay watersheds were used for analysis: Cedar Bayou, Double Bayou, and Oyster Bayou (**Figure 1a**). These regions each have well-defined outlets that drain directly to the ocean, and a contributing watershed area that extends inland from the coast far enough to capture a potentially highly dynamic flood zone, but small enough to keep the hydrologic model relatively simple. These watersheds were pulled from the Hydrofabric, which is a network of small catchments, stream reaches, and catchment properties across the US that are associated with the Ngen NWM.

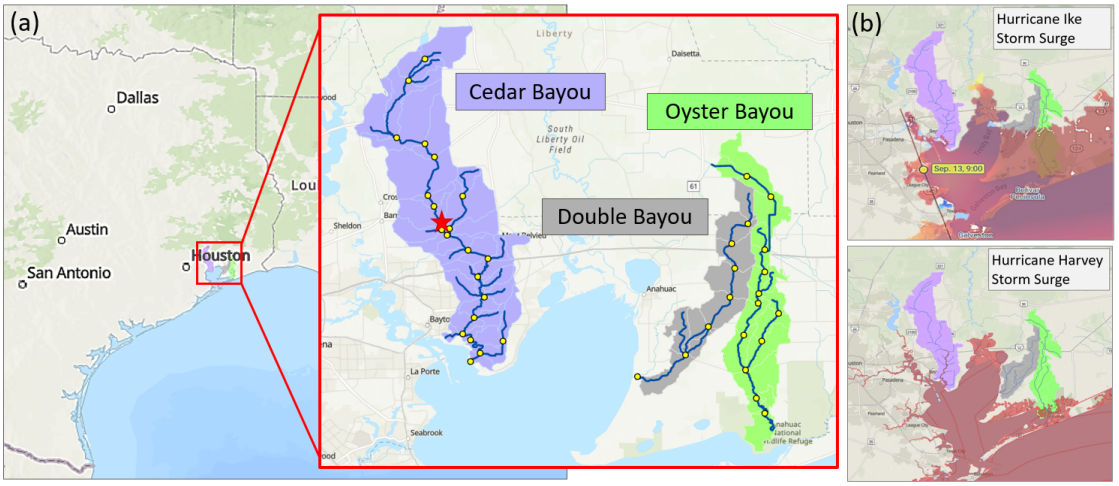


Figure 1. *(a) Study area catchments and flowlines as defined by the Ngen NWM Hydrofabric[11]. (b) Storm Surge Flood Maps for Galveston Bay area for Hurricane Ike (top), from images published by the National Weather Service [12]) and Hurricane Harvey (bottom), from ADCIRC model published by CERA [13]). The simulated storm-surge flooded area is shaded in transparent red. The red star marks the location where discharge is calculated and input into GeoClaw for the coupled model.*

4.2 Development of GeoClaw BMI

GeoClaw, the coastal model selected for this study, solves depth-averaged shallow water partial differential equations using high-resolution finite volume numerical methods. This model takes the advantage of adaptive mesh refinement to deal with storm and tsunami prediction problems. The adaptive mesh-refinement only computes detailed features where and when they are necessary for a given storm, significantly reducing computing requirements. The GeoClaw model is an open-source Fortran code with a Python code interface to facilitate setting up a geophysical model and visualizing the results [14]. This model is well-suited to generate a BMI for use with the Ngen Framework since it is open-source, has a consistent structure, and can generate multiple simulations within the time constraints of the project timeframe.

In addition to the existing model, we used a plugin to allow GeoClaw to accept streamflow discharge as an input. This plugin is developmental, and does not necessarily result in generating numerically stable models. We ran the coupled model with this plugin to demonstrate the inland-coastal flood model coupling ability.

To interact with the Ngen Framework, we built a BMI wrapper for GeoClaw. Building into the Ngen Framework has several advantages. First, it creates a platform to use any other existing BMI's to interact with the model in any supported programming language. We use the CFE BMI to generate discharge for GeoClaw, making a coupled model. Second, it provides the ability to interact with GeoClaw in a standardized interface. This interface simplifies the process needed to update or modify parameters while running timesteps in GeoClaw. Third, the integration into BMI provides a functionality to create new scenarios. The BMI runs models one checkpoint at a time, allowing us to quickly create new scenarios by changing inputs or model parameters in between timesteps.

Since we use more than one Hydrofabric catchment to determine stream discharge, the Ngen Framework requires interacting with the routing software called T-Route. While catchment discharge can be calculated using existing BMIs, multi-catchment routing and discharge is calculated by the T-Route module. T-Route takes individual catchment discharges and connects them to downstream catchments using transformation functions. Unlike a BMI, T-Route only runs after the hydrologic models have finished running. To overcome this, we created a script that reads the T-Route output file (a time-series of discharge) and sets it as the input for GeoClaw through the GeoClaw BMI. This script, paired with the discharge plug-in for GeoClaw, completes the interface for our coupled model. Figure 2 shows the basic conceptual interfaces for our coupled model.

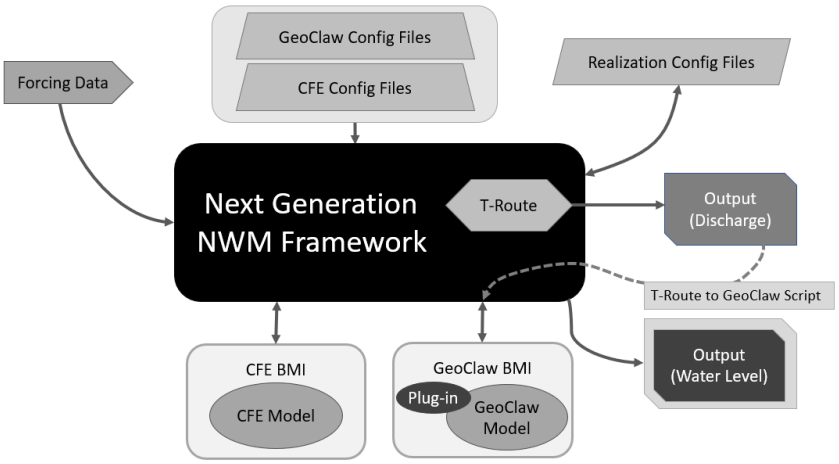


Figure 2. *Diagram of coupled model interfaces with Ngen Framework. Final output is the surface water level.*

The GeoClaw model reads the discharge value at points along the rivers in each watershed upstream of the storm surge. The red stars in Figure 1a mark the locations where discharge is added in the coupled GeoClaw model.

4.3 Model Simulations

To test the coupling process, we run three models on the study area: (1) the inland hydrologic model (CFE), (2) the coastal model (Geoclaw), and (3) the coupled coastal and inland models (GeoClaw with a discharge plugin, using peak flow output data from CFE). Detailed documentation for the set up of the models can be found at the links in the Supplementary Materials section.

To evaluate the influence of discharge on the coupled model, we place several observation points in GeoClaw along the stream in each watershed, and output a plot of water level over time at each point. We compare the plots from the coastal and coupled GeoClaw models at observation points to show variation in the water level. If the water level at an observation point is the same for the models run with and without discharge, that point is likely only influenced by inland or coastal flooding, but not both. If the water level does vary between the

coupled and coastal models, then that observation point is likely influenced by compound flooding.

4.4 Data Sources

For CFE, Daily precipitation and daily mean temperature values for one year leading up to each hurricane event are from Parameter-elevation Relationships on Independent Slopes Model (PRISM) and gridMET respectively, both downloaded from ClimateEngine [15]. Catchment delineations, flowpaths, and attributes used for the CFE model are obtained from the Hydrofabric using the hyAggregate R package [11], shown in Figure 1a.

The GeoClaw model uses Automated Tropical Cyclone Forecast (ATCF) hypothetical data [16]. For topography inputs, it uses Vol. 5 - Western Gulf of Mexico Coastal Relief Model (CRM) 3-arc second spatial resolution topobathy data, and ⅓-arc second spatial resolution for Galveston Bay from NOAA National Centers for Environmental Information (NCEI).

5. Results

5.1 GeoClaw Results

The GeoClaw model simulates Hurricanes Ike (3 days before to 1 day after the approximate landfall time (7 am Sep. 13, 2008)) and Harvey (3 days before to 4 days after the approximate landfall time (6 am Aug. 26, 2017)). A summary of results are shown in Figures 3 and 4.

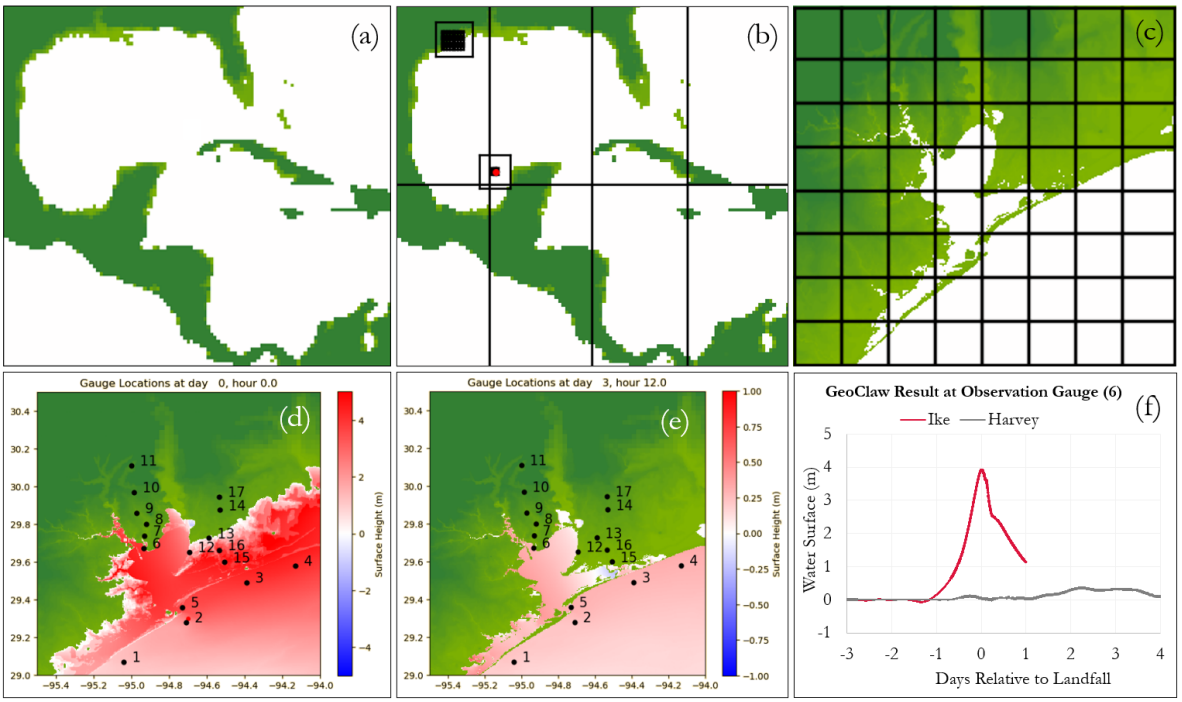


Figure 3. *(a) Entire Geoclaw model domain, (b) illustration of adaptive mesh refinement around the eye of the hurricane, (c) constant refinement of our study area, (d) snapshot of Hurricane Ike model, (e) snapshot of Hurricane Harvey model, and (f) water level at gauge 6 for hurricanes Harvey (red), and Ike(blue).*

Figure 3 shows the entire Geoclaw domain (a), the refined mesh around the study area ((b) and (c)), the surface water level represented by a colorbar across our study area for each storm ((d) and (e)), and surface water level plotted at observation gauge 6 for both storms (f). The surface water level reaches above 4 meters for Hurricane Ike, and less than 0.5 meters for Hurricane Harvey. Figure 4 validates GeoClaw results with four private gauges during Hurricane Ike [17] and certifies the agreement between the numerical model and gauge data. Additional detailed results are provided at the link in the Supplementary Materials section.

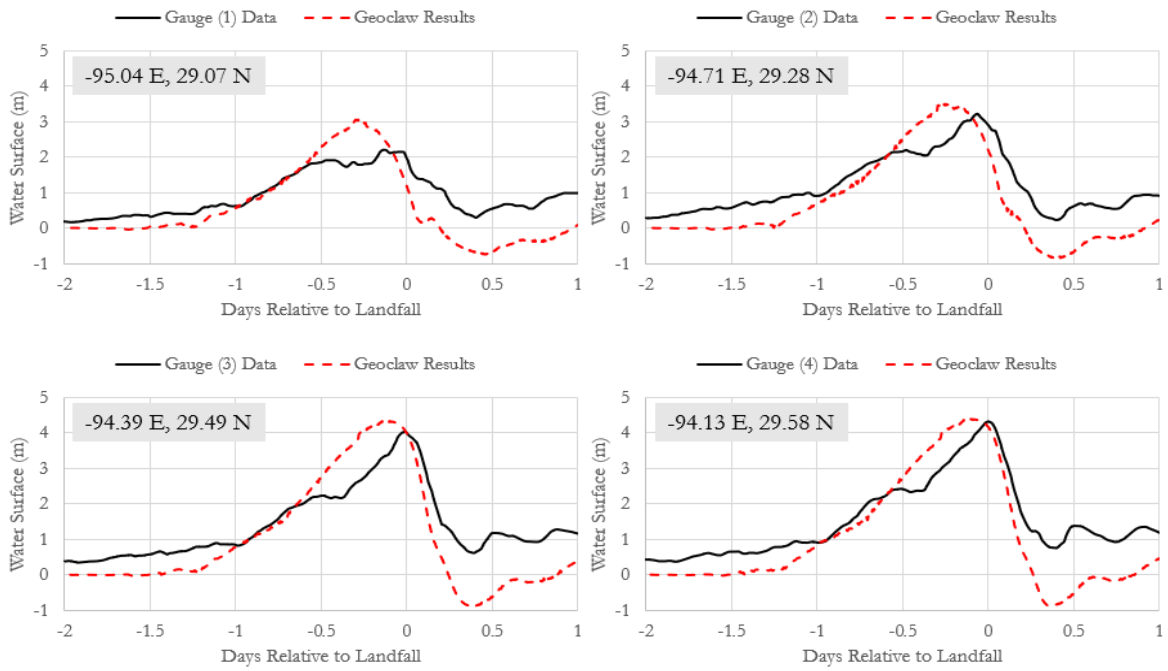


Figure 4. *GeoClaw water level output for Hurricane Ike compared to data from private gauges [17]. The black numbered points in Figure 4 show the location of the gauges.*

5.2 CFE Results

When attempting to run CFE in the Ngen Framework, errors arose due to inconsistencies between versions of Ngen resources. As a substitute, we ran CFE for each watershed as a single catchment in a python environment outside the Ngen Framework. Figure 5 shows the results as a discharge time series for each storm and watershed. Cedar Bayou has a larger contributing area to the runoff point (runoff points are shown by the red stars in Figure 1a), and therefore has a greater discharge volume.

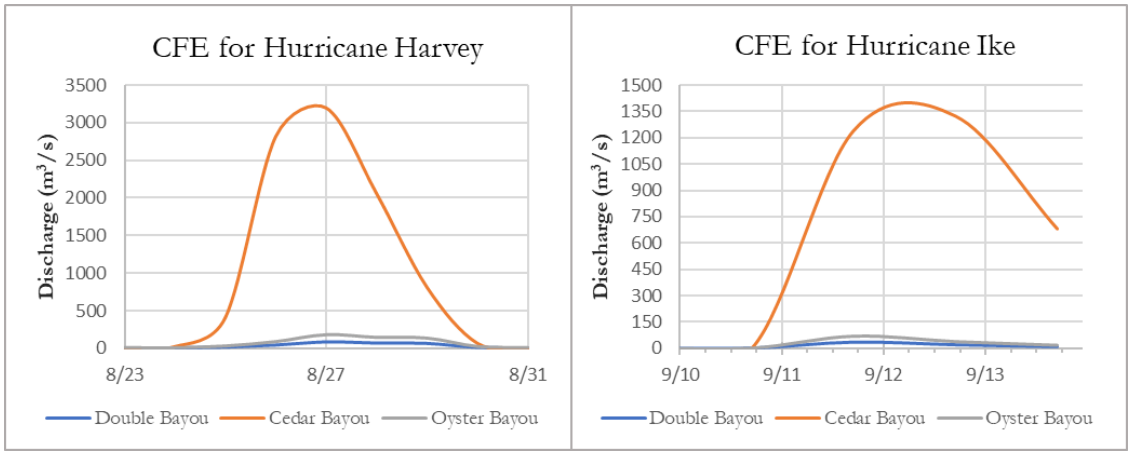


Figure 5. *Output runoff data from CFE for each watershed for Hurricanes Harvey (left) and Ike (right).*

In the near future, once the variable name and other inconsistencies between the Ngen Framework resources (T-Route, ngen code, and the Hydrofabric) are resolved, CFE with routing can be run within the Ngen Framework. Once the GeoClaw plugin that accepts time-series discharge inputs is running, these CFE with routing results can be passed to GeoClaw directly through the BMI.

5.3 Coupled Results

The coupled model was run with a single discharge value in time, matching the peak discharge value calculated by the CFE model. In the future, once the GeoClaw plugin has been improved, GeoClaw will be able to accept the time-series discharge values from CFE. Figure 6 shows an example comparing the coastal model (no discharge included) and the coupled model (with a constant discharge input near observation point 9 in Cedar Bayou of 48200 cfs and 113300 cfs for Hurricanes Ike and Harvey, respectively) at observation gauge 6. Additional results are available at the links in the Supplementary Materials section.

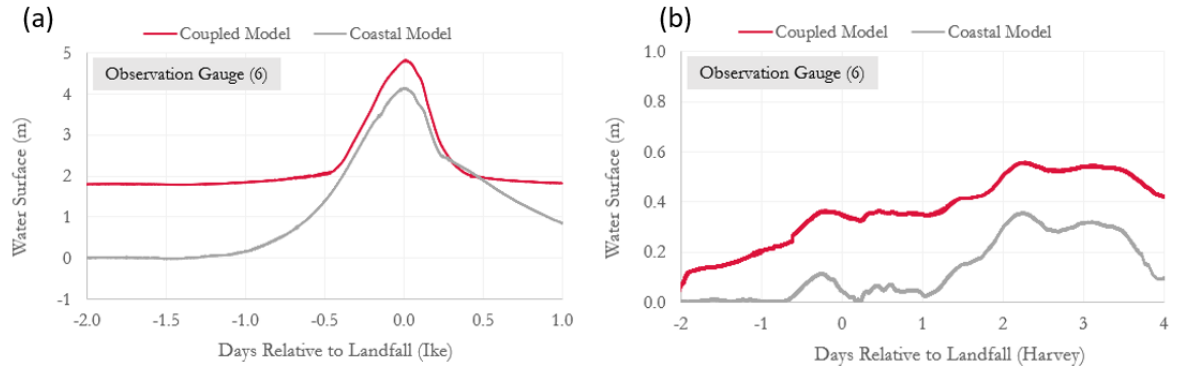


Figure 6. *Plots of water level at observation gauge 6 for the coupled and coastal models for Hurricanes Ike (a) and Harvey (b).*

6. Conclusion

We have established an initial interaction between an inland hydrological model (CFE) and a coastal hydrological model (GeoClaw) through a BMI in the Next Generation National Water Model Framework. The preliminary results show river discharge is a non-negligible factor in modeling coastal flooding. This establishment of a coastal model BMI provides an example for modular coupling of coastal and hydrologic models in future coastal flood studies.

This research could be expanded by coupling the GeoClaw BMI with other inland hydrological models, or running the models with other storms or in different regions. More accurate modeling work through this coupling scheme is dependent upon further development of the GeoClaw discharge plugin and the Ngen Framework resources. As the Ngen Framework evolves, it would be helpful to keep compound coastal flooding applications in mind to preserve the capability of the Ngen NWM to couple coastal and hydrologic models. In the future, this capability could be expanded to include two-way coupling, in which inland models could accept information about water level from coastal models. This would require changing the routing scheme to be bidirectional. The coupling process would also benefit from having the routing information accessible via a BMI, which is not currently the case in the existing Ngen Framework.

Supplementary Materials:

Detailed results of all three models (CFE, GeoClaw, and Coupled) are hosted on Hydroshare at https://www.hydroshare.org/resource/379b4c8c663c460d87c246641dc5cea2/. Code for the GeoClaw BMI, an example of how to run the GeoClaw model in the Next Framework, and an example of how to run CFE outside the Ngen framework can be found at https://github.com/henrichsen/GeoClaw_BMI.

References

1. Taherkhani, M., et al. Sea-level rise exponentially increases coastal flood frequency. *Sci Rep* **2020,** 10, 6466. Available online: https://doi.org/10.1038/s41598-020-62188-4 (accessed on 27 June 2022).

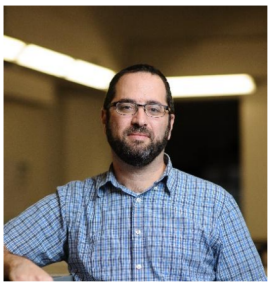
2. Biasutti M.; Sobel A.H.; Camargo S.J.; Creyts T.T. Projected changes in the physical climate of the Gulf Coast and Caribbean. *Climatic Change* **2012**, 112, 819–845. Available online: https://doi.org/10.1007/s10584-011-0254-y (accessed on 26 June 2022).
3. Ogden, F.L. Justification of and Parameter Estimation for a Conceptual Functional Equivalent (CFE) Formulation of the NOAA-NWS National Water Model (version 2.1) **2021**.
4. Field, C.; Barros, V.; Stocker, T.; Dahe, Q. (Eds.). Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation: Special Report of the Intergovernmental Panel on Climate Change. *Cambridge: Cambridge University Press* **2012**, doi:10.1017/CBO9781139177245.
5. Bevacqua, E.; et al. Higher probability of compound flooding from precipitation and storm surge in Europe under anthropogenic climate change. *Science Advances* **2019,** 5(9). Available online: https://doi.org/10.1126/sciadv.aaw5531 (accessed 26 June 2022).
6. Wahl, T.; Jain, S.; Bender, J.; Meyers, S.D.; Luther, M. E. Increasing risk of compound flooding from storm surge and rainfall for major US cities. *Nature Climate Change* **2015**, 5(12), 1093–1097. Available online: https://doi.org/10.1038/nclimate2736 (accessed on 26 June 2022).
7. Wdowinski, S.; Bray, R.; Kirtman, B.P.; Wu, Z. Increasing flooding hazard in coastal communities due to rising sea level: Case study of Miami Beach, Florida. *Ocean & Coastal Management* **2016**, 126, 1–8. Available online: https://doi.org/10.1016/j.ocecoaman.2016.03.002 (accessed on 26 June 2022).
8. Santiago-Collazo, F.L.; Bilskie, M.V.; Hagen, S.C. A comprehensive review of compound inundation models in low-gradient coastal watersheds. *Environmental Modelling & Softwar*e **2019**, 119, 166–181. Available online: https://doi.org/10.1016/j.envsoft.2019.06.002 (accessed on 26 June 2022).
9. Muñoz, D. F. et al. Inter-Model Comparison of Delft3D-FM and 2D HEC-RAS for Total Water Level Prediction in Coastal to Inland Transition Zones. *JAWRA* **2022**, 58(1), 34–49. Available online: https://doi.org/10.1111/1752-1688.12952 (accessed 26 June 2022).
10. Ye, F.; et al. Simulating storm surge and compound flooding events with a creek-to-ocean model: Importance of baroclinic effects. *Ocean Modelling* **2020**, 145, 101526. Available online: https://doi.org/10.1016/j.ocemod.2019.101526 (accessed 26 June 2022).
11. Johnson, M. NextGen Hydrofabric. *NOAA* **2022**. Available online: https://mikejohnson51.github.io/hyAggregate/ (accessed 6 July 2022).
12. NOAA's National Weather Service. Hurricane Ike Storm Surge. *Tropical Weather* **2016**. Available online: https://www.weather.gov/lch/ikesurge (accessed 26 June 2022).
13. CERA. (n.d.). Storm Surge - Wave - Compound Flood Guidance. *Coastal Emergency Risks Assessment*. Available online: https://cera.coastalrisk.live/ (accessed 6 July 2022).
14. Berger, M.J., George, D.L.; LeVeque, R.J.; Mandli, K. The GeoClaw Software for Depth-Averaged Flows with Adaptive Refinement. **2010**. Available online: https://doi.org/10.48550/ARXIV.1008.0455 (Accessed 10 July 2022).
15. Daudert, B.; et al. ClimateEngine. *app.climateengine.com*. Available online: https://app.climateengine.com/climateEngine (accessed 7 July 2022).
16. Ftp.nhc.noaa.gov. 2022. Index of /atcf. Available online: https://ftp.nhc.noaa.gov/atcf/ (accessed on 5 July 2022).
17. Kennedy, A.B.; et al. Origin of the Hurricane Ike Forerunner Surge: HURRICANE IKE FORERUNNER. *Geophysical Research Letters* **2011,** 38, no. 8. Available online: https://doi.org/10.1029/2011GL047090. (Accessed 8 July 2022).

Appendix

Team Leaders

Fred Ogden,
Chief Scientist
National Water Center

Jonathan Frame,
Senior Hydrologist
Cloud to Street



Sagy Cohen,
Associate Professor
University of Alabama

Chaopeng Shen,
Associate Professor
Penn State University

Kyle Mandli,
Assistant Professor
Columbia University

Celso Ferriera,
Associate Professor
George Mason University

Summer Institute Course Coordinators and the Executive Director of CUAHSI

Arash Modaresi Rad, Jerad Bales, and Emily Deardorff,

Summer Institute Students

Chapter 1: Automated decision support for model selection in the Nextgen National Water Model

Francisco Haces-Garcia, Lauren A. Bolotin, Qiyue Liu, Mochi Liao

Chapter 2: Data assimilation of USGS streamflow data in the Nextgen Framework

Zachariah J. Butler, Fitsume T. Wolkeba, Kenneth O. Ekpetere, and Motasem S. Abualqumboz

Chapter 3: Data-driven approaches for estimating river channel geometry

Zahra Ghahremani, Shuyu Chang, Laura Manuel, and Mohammad Erfani

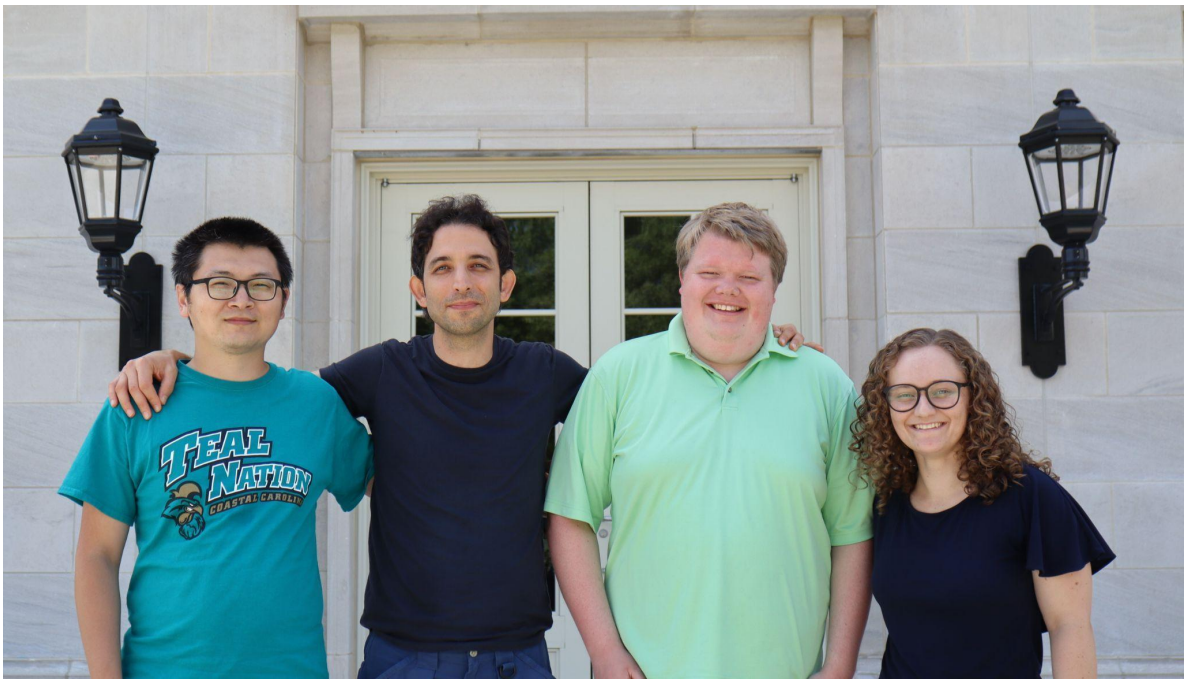
Chapter 4: Large scale prediction of channel roughness coefficient using machine learning

Shah Saki, Krutikkumar Patel, and Md Abdullah Al Mehedi

Chapter 5: QuiCFIM, a quick GIS-based combined flood inundation mapping framework

Mark Wang, Ankit Ghanghas, and Mohamed Abdelkader

Chapter 6: Coupling coastal and hydrologic models through BMI and Next Generation National Water Model Framework in low gradient coastal regions of Galveston Bay, Texas, USA



Hongyuan Zhang, Ebrahim Hamidi, Hart Henrichsen, and Abbie Sandquist