# Analysis of Flood Drivers Contributions to Compound Flooding Using Coupled Modeling and Machine Learning


**Javed Ali[1], Sadaf Mahmoudi[2], Farnaz Yarveysi[3], Samuel Daramola[4]**

[1]University of Central Florida; *javed.ali@ucf.edu*

[2]University of Alabama, Tuscaloosa; *smahmoudikouhi@crimson.ua.edu*

[3]University of Alabama, Tuscaloosa; *fyarveysi@crimson.ua.edu*

[4]Virginia Tech, Blacksburg; *samueldaramola@vt.edu*

**Academic Advisors:** Thomas Wahl, *University of Central Florida*[1]; Hamed Moftakhari, *University of Alabama*[2]; Hamid Moradkhani, *University of Alabama*[3]; David Muñoz, *Virginia Tech*[4]

**Summer Institute Theme Advisor:** Kyle T. Mandli, Columbia University, *kyle.mandli@columbia.edu*



**Abstract:** This study aimed to understand the drivers of compound flooding in New York City (NYC) using advanced hydrological and hydrodynamic modeling techniques and machine learning methods. We employed the National Water Center's Conceptual Functional Equivalent (CFE) model for hydrologic modeling of eight delineated watersheds within the NYC area, and for hydrodynamic modeling, we used GeoClaw, incorporating river discharge data from CFE. By coupling these two modeling methodologies, we simulated nine historical storm events, determining the combined effects of precipitation, storm surge, and river discharge as drivers of compound flooding. To quantify their relative contributions, we trained and validated three machine learning models – Random Forest, Support Vector Machine, and Multi-Layer Perceptron – on the modeled flood depth data. Our findings demonstrated that storm surge was the main cause of compound flooding in NYC, with precipitation also playing an important role. Interestingly, river discharge doesn't have as much impact on these flooding patterns. These results, supported by historical data, have profound implications for urban planning, disaster management, and policy-making in NYC, providing a solid foundation for developing targeted strategies to mitigate compound flooding.


## 1. Motivation

Flooding, a phenomenon resulting from drivers, such as heavy rainfall, storm surge, and river discharge, poses significant risks to coastal communities in the United States [1]. The severity and frequency of these events are expected to increase due to climate change, necessitating a comprehensive understanding of the contributing drivers and their interactions [2]. Historically, the study of flooding events has often been compartmentalized, with separate analyses conducted for different types of flooding such as fluvial, pluvial, and coastal contributions. However, in many instances, these events do not occur in isolation but rather in combination, leading to compound flooding. The drivers of these events, including storm characteristics, sea-level rise, and land-use changes, can interact in complex ways to exacerbate flooding impacts. In New York City (NYC), for instance, the observed rise in compound events has been attributed to a shift toward weather patterns characterized by storm surges that coincide with increased precipitation [3]. Traditional modeling approaches have struggled to accurately capture these interactions due to their inherent complexity and the high dimensionality of the problem. However, recent advances in computational power and the development of sophisticated coupled models that integrate atmospheric, oceanic, and hydrological processes have opened new possibilities for studying compound flooding [4]. Despite these advances, significant challenges remain in analyzing compound flooding events. Machine learning, with its ability to handle high-dimensional data and capture complex patterns [5], offers a promising tool for improving our understanding and analysis of compound flooding.

## 2. Objectives and Scope

This study aims to assess the relative contributions of various flood drivers that may have contributed to compound flooding in Manhattan, New York City, during various historical storms. To achieve this objective, a coupled model that integrates both hydrological and hydrodynamic processes is implemented. The hydrological component of the model utilizes the National Water Center's CFE (Conceptual Functional Equivalent) model within the NextGen Framework, while the hydrodynamic component employs the GeoClaw numerical model. By coupling these models, the study simulates the impact of hurricanes and tropical storms that have affected New York City in recent years. The outputs of the coupled model provide insights into the flood drivers associated with each simulated event. Additionally, the hydrodynamic modeling component allows for the estimation of flood depths across the study area. To further investigate the contribution of each flood driver to the flood depth at the tract-level resolution, machine learning algorithms were employed. By leveraging various machine learning algorithms, the study aims to determine the relative importance and contribution of each flood driver in different parts of Manhattan. This integrated approach combining hydrologic and hydrodynamic coupled modeling, and machine learning techniques will provide valuable insights into the relative contributions of flood drivers in Manhattan, ultimately enhancing our understanding of flood processes and facilitating effective flood management and mitigation strategies in urban coastal areas.

## 3. Previous Studies

GeoClaw is a two-dimensional hydrodynamic model that incorporates adaptive mesh refinement capabilities, primarily designed for simulating shallow earth-surface flows involving water-wave propagation and inundation, including scenarios such as tsunamis, storm surges, and general overland flooding [6]. High-resolution finite volume methods are employed in GeoClaw to address geophysical flow problems. Spero et al. [7] compared GeoClaw with HEC-RAS for modeling the 1976 Teton Dam failure. The evaluation of GeoClaw's suitability for dam failure modeling was based on its ability to accurately depict the extent of inundation and the arrival times of flood waves. The study found that the 2D GeoClaw dam-break model produced results that reasonably aligned with historical gauge records, field observational data, and HEC-RAS results. The model demonstrated stability and relatively low computational costs. While GeoClaw has been predominantly utilized in dam failure, tsunami, and geo-hazard studies, there are limited examples of its application in flood modeling. The use of GeoClaw in coastal flooding and storm surge simulations is not extensively documented. Hence, in this study, we propose to employ GeoClaw and validate its performance against observational data from NOAA tide gauges to assess its efficacy in simulating coastal flooding events. The selection of the most significant features plays a crucial role in pattern recognition systems. Nowadays, when examining the combined effects of multiple variables, researchers often employ machine learning techniques to obtain relevance scores [8]. For instance, Yarveysi et al. [1] used a machine learning algorithm to objectively assign weights to variables contributing to the overall estimated vulnerability, thus reducing subjectivity in determining the impact of various social, economic, and infrastructural factors on vulnerability. Similarly, Opoku et al. [9] utilized five supervised machine learning algorithms to predict depression. Through the permutation importance method, they were able to identify influential behavioral markers in the prediction of depression.

## 4. Methodology

### 4.1 Data and Study Area

This study focuses on analyzing historic storms that occurred between 2007 and 2019 and impacted New York City (NYC) (Figure 1a). We selected eight catchments that cover the whole study area (Figure 1b). The selected time period aligns with the capabilities of the hydrofabric AORC (Atmospheric Oceanic Reanalysis and Characterization) data utilized in this study.

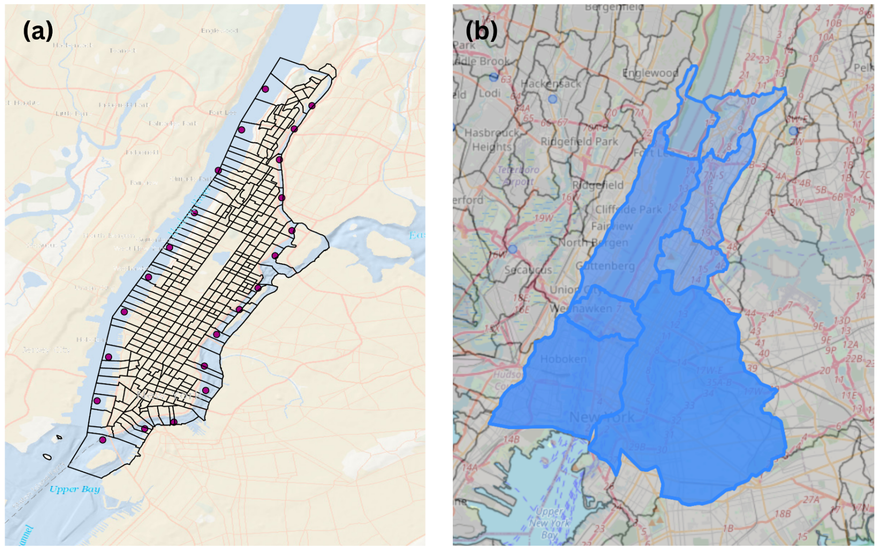


**Figure 1:** (a) Study area: Manhattan, New York City, and (b) eight catchments used in the study area.

Table 1 presents the information about the storms, and their ATCF (Automated Tropical Cyclone Forecasting) names. All 9 storms included in the analysis are simulated using the CFE model, and the maximum discharge associated with each storm is subsequently used as an input parameter for the GeoClaw hydrodynamic model. The storm data utilized in GeoClaw, including the storm's eye latitude, longitude, pressure, wind speed, and radius, can be obtained from the following repository: https://ftp.nhc.noaa.gov/atcf/archive/. By incorporating the storm data from these selected events, this study aims to comprehensively examine the hydrological and hydrodynamic processes associated with historic storms in NYC, providing valuable insights into flood dynamics and contributing to enhanced flood risk management strategies.

**Table 1:** Information about the storm events, their respective landfall dates, the dates of impact on NYC, the hour of landfall, and their ATCF names

| Storms | Date of Landfall | Date of impact in NYC | Hour of landfall | ATCF data |
|---|---|---|---|---|
| Tropical Storm Barry | 2-Jun-2007 | 5-Jun-2007 | 02:00 | AL022007 |
| Hurricane Hannah | 6-Sep-2008 | 6-Sep-2008 | 07:20 | AL082008 |
| Hurricane Irene | 28-Aug-2011 | 28-Aug-2011 | 13:00 | AL092011 |
| Hurricane Sandy | 29-Oct-2012 | 29-Oct-2012 | 23:30 | AL182012 |
| Hurricane Arthur | 4-Jul-2014 | 4-Jul-2014 | 08:00 | AL012014 |
| Tropical Storm Jose | 19-Sep-2017 | 20-Sep-2017 | 00:00 | AL122017 |
| Tropical Storm Philippe | 28-Oct-2017 | 30-Oct-2017 | 22:00 | AL182017 |
| Hurricane Dorian | 6-Sep-2019 | 7-Sep-2019 | 12:30 | AL052019 |
| Hurricane Ogla | 27-Oct-2019 | 27-Oct-2019 | 03:00 | AL172019 |

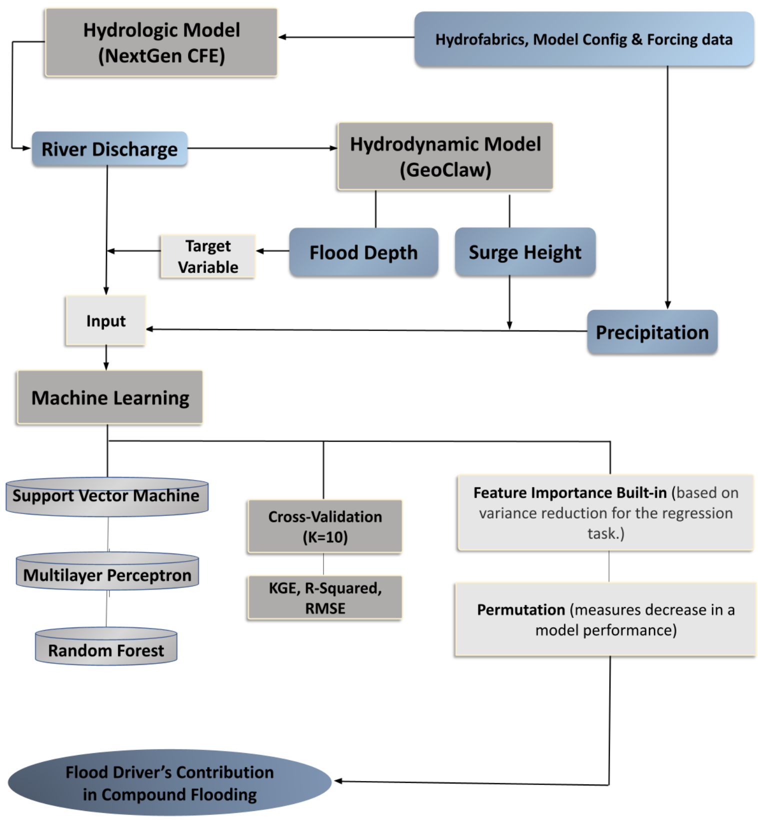


**Figure 2:** Flowchart showing the methodology employed in this study.

### 4.2 Hydrologic Modeling using the National Water Center's CFE Model

The National Water Center's Conceptual Functional Equivalent (CFE) model, a conceptual rainfall-runoff model designed to simulate rainfall conversion into runoff, was utilized in this study to understand the volume of water flowing into rivers and streams post-rain events. The initial stage of hydrologic modeling with the CFE entails obtaining and processing HydroFabric data pertinent to the eight delineated watersheds within the study area. These data are secured from the Amazon S3 Bucket and processed to construct parameter configuration files. These configuration files, pivotal for the operation of the CFE model and the Simple Logical Tautology Handler (SLoTH) within the NextGen framework, consist of model default parameters, specific formulations, detailed input and output paths, simulation time steps, and initial conditions, along with other settings relevant for the precise modeling of the hydrological system under study. Following this, basin-averaged forcing inputs are generated, each designed to correspond with distinct storm event time periods within the individual watersheds. These inputs, drawn from AORC v1.0 kerchunk header files, are specifically prepared for integration with NOAA's advanced Next Generation (NextGen) Water Resource Modeling Framework. Finally, the CFE model is executed within the NGEN framework using the 'ngen' command, which consolidates the positional arguments for running the model. More information on this process can be found in the GitHub repository.

### 4.3 Hydrodynamic Modeling using GeoClaw

*i) Setting up the model for validation:* The model was set up according to the descriptive specifications in the clawpack repository and our model repository on GitHub.

*ii) Incorporating river discharge:* The subroutine "src2" script, written in Fortran, was used to integrate the river discharge into the model. It sets certain geographical bounds for the river source and computes the river's discharge in cubic meters per second for cells that fall within the river source area.
*iii) Bias Correction:* A maximum bias correction is a statistical approach employed in hydrological and meteorological modeling to address systematic biases in modeled variables. The methodology of maximum bias correction [10] involves identifying the maximum bias between the modeled and observed data, followed by the application of a correction factor to mitigate or eliminate the bias. The primary objective of maximum bias correction is to enhance the accuracy and reliability of modeled outputs.

### 4.4 Machine Learning Approach

The outcome of coupled modeling (including precipitation, storm surge, and river discharge) for the nine historical storms was used as input features to the various ML algorithms tested here. These algorithms are trained and validated using the flood depth estimates from the GeoClaw hydrodynamic model, split 80/20 for training/validation. The depth estimates are obtained at 20 computation points randomly selected around Manhattan. We tested the applicability and performance of three different ML algorithms namely Random Forest (RF), Support Vector Machine Regression (SVMR), and Multi-Layer Perceptron (MLP). To select the best-performing algorithm, we consider various evaluation metrics (i.e., KGE and RMSE). RF's inherent feature importance mechanism was used for the regression task, while the permutation importance method was utilized for SVMR, and MLP to calculate feature importance. K-fold cross-validation (K=10) was employed to verify the algorithm's performance on new data, ensuring stability in accuracy estimates.

### 4.5 Evaluation Metrics

Four evaluation metrics were used to assess the effectiveness of the machine learning and the hydrodynamic models.: the Nash-Sutcliffe efficiency (NSE) [11], gauging the predictive accuracy and reliability of the model; the Kling-Gupta efficiency (KGE) [12], providing an overall assessment of the model's ability to replicate observed values in terms of timing, magnitude, and variability; the coefficient of determination (R-squared), representing the proportion of variance in the dependent variable that can be attributed to the independent variables; and the root mean square error (RMSE), which assessed the accuracy and precision of the model by measuring the average discrepancy between predicted and observed values [13].

## 5. Results

### 5.1 Hydrologic Modeling with NextGen CFE Model

Figure 3 shows the river discharge for different catchments over time during Hurricane Irene (2011). The river discharge for each catchment generally increases over time, which is expected as the hurricane progresses. There is a notable peak in discharge rates for all catchments around August 28, which likely corresponds to the height of the hurricane. Catchment "cat-694852" consistently has the highest discharge rate throughout the event, reaching close to 120,000 cubic meters per second at its peak. Catchments "cat-694853" and "cat-694854" also show significant increases in discharge during the hurricane, with maximum values around 46,000 and 26,000 cubic meters per second, respectively. The rest of the catchments have lower overall discharge rates, with maximum values not exceeding 24,000 cubic meters per second. The high values of discharge could be attributed to the intense rainfall and strong winds associated with Hurricane Irene. These conditions can result in a significant increase in river discharge, as more water is transported into rivers and streams from direct rainfall and surface runoff. Results for other storm events are shown in the supplementary information.

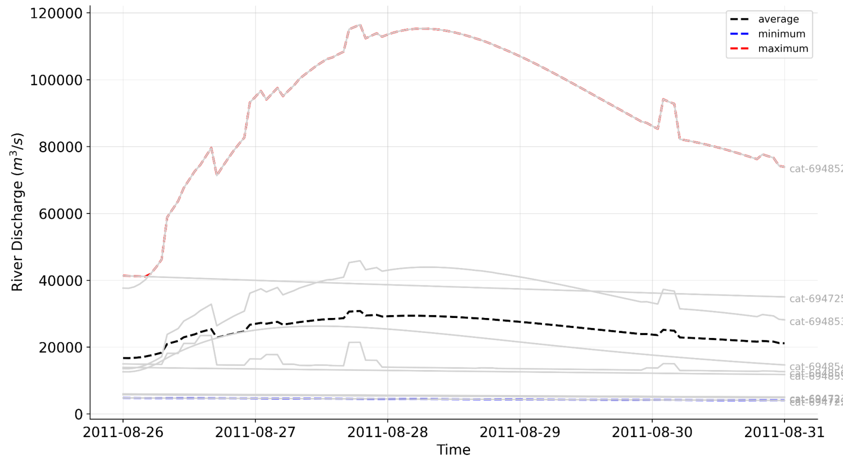


**Figure 3:** River discharge output from the NextGen CFE model during Hurricane Irene (2011) for different catchments in the NYC area.

## 5.2 Hydrodynamic Modeling using GeoClaw Model

In this study, the model's performance was validated in the context of two hurricanes, Irene (2011) and Sandy (2012). The results of Sandy are shown in the Supplementary Information. The performance of the GeoClaw model in capturing storm events exhibits notable proficiency in capturing the peak of storms. However, during the initial days leading up to the storm event, the model may exhibit variations. Consequently, to ensure a reliable validation process, the decision was made to focus solely on validating the model against the peak of the storms.

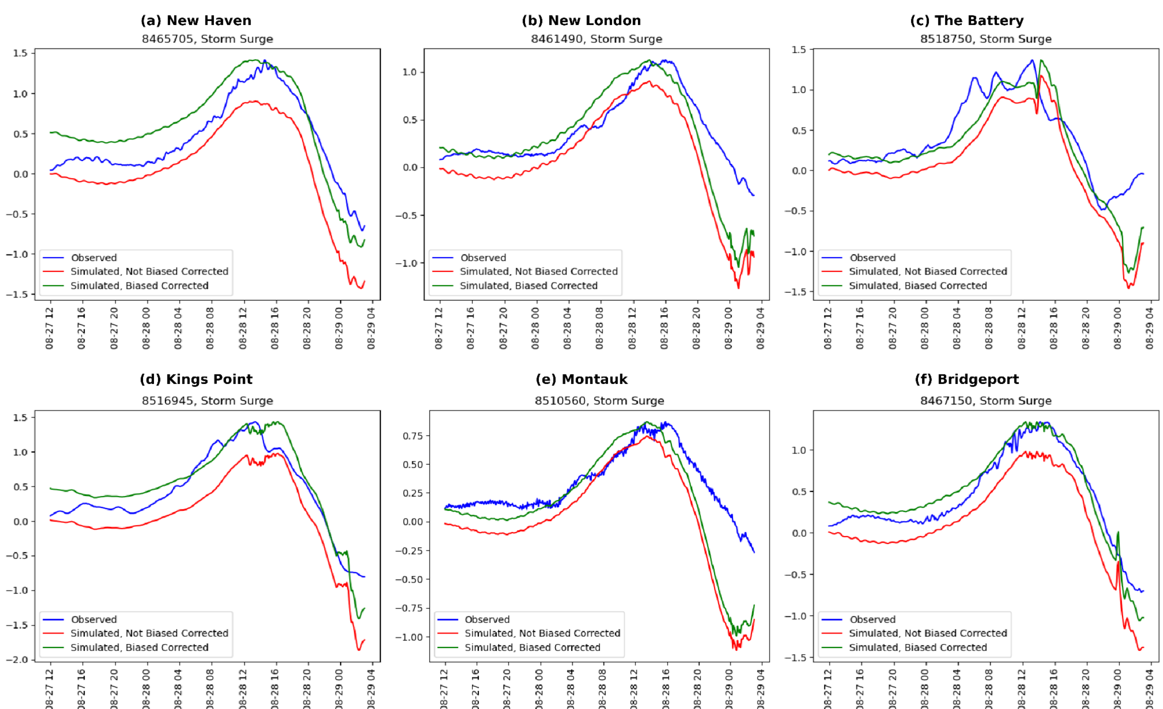


**Figure 4:** Comparison of observed, simulated, and maximum biased corrected simulated storm surges during Hurricane Irene (2011).

Figure 4 shows the comparison between the simulation and observation time series at the time of the storm's peak during Hurricane Irene, which provides evidence of GeoClaw's proficiency in capturing the peak of the storm. Moreover, evaluation metrics of the biased corrected and not biased corrected simulation results of Hurricane Irene are shown in Table 2. These findings highlight the station-specific effects of bias correction on the evaluation of the models used in the study.

**Table 2:** The evaluation metrics of simulated storm surges during Hurricane Irene; the bold ones are after bias correction.

| Hurricane Irene's Evaluation Metrics | | | | | |
|---|---|---|---|---|---|
| Station Name | Station ID | NSE | KGE | RMSE | R Squared |
| The Battery | 8518750 | 0.204⇒**0.559** | 0.313⇒**0.594** | 0.43⇒**0.32** | 0.204⇒**0.559** |
| Kings Point | 8516945 | 0.515⇒**0.847** | 0.344⇒**0.737** | 0.391⇒**0.219** | 0.515⇒**0.847** |
| Montauk | 8510560 | -0.188⇒**-0.45** | 0.109⇒**0.077** | 0.296⇒**0.327** | -0.188⇒**-0.45** |
| Bridgeport | 8467150 | 0.647⇒**0.917** | 0.428⇒**0.835** | 0.292⇒**0.146** | 0.647⇒**0.917** |
| New Haven | 8465705 | 0.612⇒**0.685** | 0.342⇒**0.568** | 0.309⇒**0.278** | 0.612⇒**0.685** |
| New London | 8461490 | 0.353⇒**0.299** | 0.324⇒**0.478** | 0.291⇒**0.303** | 0.353⇒**0.299** |

### 5.3 Machine Learning

We examined a variety of performance metrics to determine which machine learning (ML) algorithms are most suited to the specific objectives of this project. We found that MLP is the best-performing algorithm (Figures 5a, 5b, and 5c). This algorithm, with a median KGE of 0.82, and a median RMSE of 0.17 m, demonstrates significantly better performance than SVM and slightly better than RF. Results from the feature importance analysis implemented using the trained MLP (Figure 5e) showed storm surge as the predominant factor contributing to the impacts of compound flooding around Manhattan, with an average relative importance score of 0.53. Local precipitation, although the secondary contributor, still holds substantial influence, exceeding storm surge in certain areas and averaging an importance score of 0.29. River discharge, however, has a relatively limited contribution to flood severity around Manhattan, with an average score of 0.17. These findings align with existing literature. For instance, comprehensive studies of compound historic floods in the lower Hudson River (such as [14]) concur that the river flow's contribution to flood variability is negligible to extremely limited. Thus, our analysis reaffirms these observations, providing valuable insights for future flood mitigation efforts in the Manhattan area. For further details, please refer to Supplementary information.

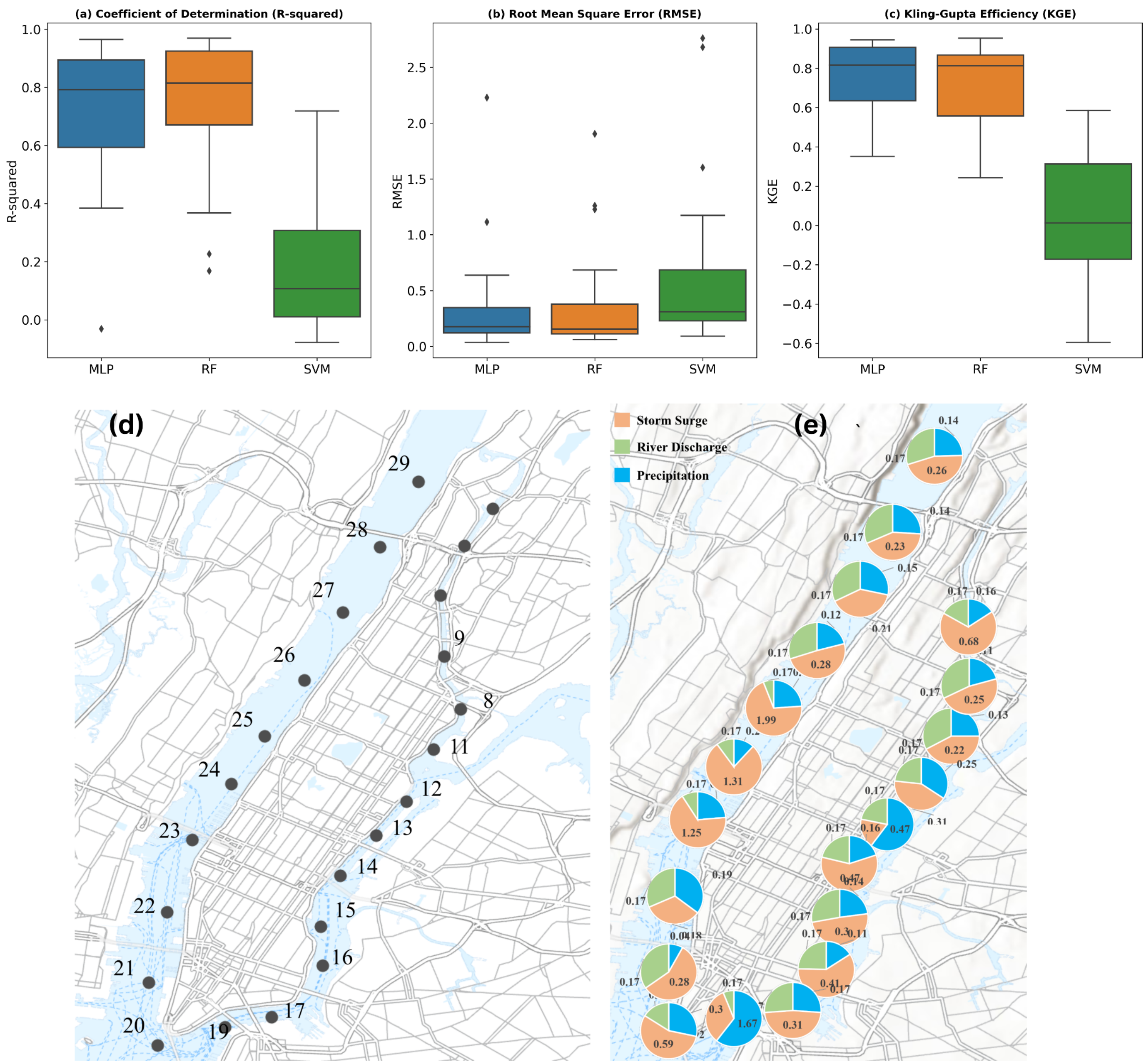


Figure 5: Performance metrics of various ML algorithms (a) R-squared, (b) RMSE, (c) KGE, (d) study locations (numbered), (e) the relative importance of different flood drivers at different locations in the Manhattan area.

## 6. Conclusion

In this study, we conducted a thorough analysis of nine historic storms that caused significant flooding impacts in Manhattan, New York City. We developed a coupled hydrologic-hydrodynamic modeling framework that uses NextGEN CFE for hydrologic modeling to simulate runoff generated during major storms and GeoClaw to estimate water level variation in the study area based on the fluvial flux received from the NextGEN-CFE model and other forms of reanalysis data such as ocean and atmospheric boundary forcing. The main output of the coupled modeling framework was the calculation of water levels at different locations throughout Manhattan, which serve as the target variables for the machine learning (ML) algorithms implemented within the study. By incorporating the intensity of fluvial, pluvial, and coastal drivers as input features, these sophisticated ML algorithms allowed us to analyze the contribution of each of these drivers to the overall water level variability at different locations across the study area during compound flooding events. The study's findings underscore that storm surge emerges as the determinant factor in understanding the complex dynamics of compound flooding in Manhattan. Moreover, local precipitation emerges as a secondary, yet significant, contributor. Contrarily, river discharge does not demonstrate a substantial role in influencing the dynamic variability of the flooding regime within the study area.

**Supplementary Material:** The supplementary information can be found here. The computational code implemented in this study is accessible at our open-access GitHub repository https://github.com/javedali99/si2023-compound-flooding

## References

1. Yarveisi, F., Alipour,A., Moftakhari, H., Jafarzadegan, K., Moradkhani, H.(2023). Block-Level Vulnerability Assessment Reveals Disproportionate Impacts of Natural Hazards across the Conterminous United States. Nature Communications. DOI: 10.1038/s41467-023-39853-z.
2. Ming, X., Liang, Q., Dawson, R., Xia, X., & Hou, J. (2022). A quantitative multi-hazard risk assessment framework for compound flooding considering hazard inter-dependencies and interactions. *Journal of Hydrology*, *607*, 127477. https://doi.org/10.1016/j.jhydrol.2022.127477.
3. Lacy, H. S., DeVito, A., & De Nivo, A. C. (2013). Geotechnical Aspects of Three Storm Surge Barrier Sites to Protect New York City from Flooding. In Storm Surge Barriers to Protect New York City: Against The Deluge (pp. 134-149). https://doi.org/10.1061/9780784412527.010.
4. Olbert, A. I., Moradian, S., Nash, S., Comer, J., Kazmierczak, B., Falconer, R. A., & Hartnett, M. (2023). Combined statistical and hydrodynamic modelling of compound flooding in coastal areas—Methodology and application. *Journal of Hydrology*, *620*, 129383. https://doi.org/10.1016/j.jhydrol.2023.129383
5. Dargan, S., Kumar, M., Ayyagari, M. R., & Kumar, G. (2020). A Survey of Deep Learning and Its Applications: A New Paradigm to Machine Learning. *Archives of Computational Methods in Engineering*, *27*(4), 1071–1092. https://doi.org/10.1007/s11831-019-09344-w.
6. Berger, M. J., George, D. L., LeVeque, R. J., & Mandli, K. T. (2011). The GeoClaw software for depth-averaged flows with adaptive refinement. Advances in Water Resources, 34(9), 1195–1206. https://doi.org/10.1016/j.advwatres.2011.02.016.
7. Spero, H., Calhoun, D., & Schubert, M. (2022). Simulating the 1976 Teton Dam Failure using Geoclaw and HEC-RAS and comparing with Historical Observations (arXiv:2206.00766). arXiv. http://arxiv.org/abs/2206.00766.
8. J. M. Fontana, M. Farooq and E. Sazonov, "Estimation of feature importance for food intake detection based on Random Forests classification," 2013 35th Annual Internabltional Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Osaka, Japan, 2013, pp. 6756-6759, doi: 10.1109/EMBC.2013.6611107.
9. Opoku Asare K, Terhorst Y, Vega J, Peltonen E, Lagerspetz E, Ferreira D. Predicting Depression From Smartphone Behavioral Markers Using Machine Learning Methods, Hyperparameter Optimization, and Feature Importance Analysis: Exploratory Study. JMIR Mhealth Uhealth 2021;9(7):e26540.
10. Teklu T. Hailegeorgis and Knut Alfredsen, "Regional flood frequency analysis and prediction in ungauged basins including estimation of major uncertainties for mid-Norway", Journal of Hydrology: Regional Studies, vol. 9, pp. 104-126, 2017.
11. Nash, J.E. and Sutcliffe, J.V. (1970) River Flow Forecasting through Conceptual Model. Part 1—A Discussion of Principles. Journal of Hydrology, 10, 282-290.
12. Gupta, H. V., & Kling, H. (2009). On the evaluation of hydrological models: towards a measurable paradigm. Hydrology and Earth System Sciences, 13(12), 1869-1880.
13. Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. Geoscientific Model Development, 7(3), 1247–1250. https://doi.org/10.5194/gmd-7-1247-2014.
14. Orton, P., Georgas, N., Blumberg, A., & Pullen, J. (2012). Detailed modeling of recent severe storm tides in estuaries of the New York City region. Journal of Geophysical Research: Oceans, 117(C9). https://doi.org/10.1029/2012JC008220