# Abstract

Floods pose a significant global risk to our society, economy and environment; these hazards are increasing due to climate change, sea-level rise and land cover change. Mitigating these risks require reliable predictive models for disaster preparedness and mitigation. While physically-based models provide robust predictions, they demand extensive computational resources. Morphology-based models, though computationally efficient, struggle in coastal watersheds with complex oceanic and hydrologic interactions. Machine learning (ML) models offer an alternative framework, yet challenges remain in their transferability, and uncertainty quantification.

The main goal of this dissertation was to evaluate the efficiency of ML-based models for hindcasting flood characteristics in coastal watersheds. The main objectives were to: **(1)** evaluate the ML model transferability across different flood events and coastal watersheds, **(2)** assess the incorporating high-water marks (HWMs), topobathy data and their uncertainties in improving the performance of ML models, and **(3)** explore the performance of ML models in highly regulated coastal watersheds where hydraulic infrastructure influences water movement. Computationally efficient, flexible, and transferable models were developed using regression-based ML algorithms, customized loss functions, and geospatial analyses.

The research focused on multiple large coastal watersheds in the northeastern and southeastern United States, using recent hurricanes as case studies. The main findings included:

1. The ML models demonstrated strong performance in hindcasting maximum flood depths for the event they were trained for in natural and regulated coastal watersheds.
2. In natural watersheds, the models showed satisfactory transferability across flood events with a slight degradation in their performance. That is, ML models can be used for hindcasting flood depths for events in the same watershed with satisfactory performance. The transferability was, however, not satisfactory in regulated coastal watersheds. The transferability was dependent on the characteristics of the flood event such as rainfall, storm track, storm surge height, wind speed and antecedent conditions.
3. Incorporating the uncertainty of flood observations substantially improved the performance and transferability of the hindcast model.
4. In regulated coastal watersheds, the inclusion of hydraulic infrastructure features (e.g., levees, pumps and spillways) improved the ML model performance when the uncertainty of flood observations was incorporated in the model.
5. The ML model should be cautiously applied to unseen data, beyond the events and watersheds that they were trained for.

This research contributed to the development of computationally efficient flood hindcast models that assist decision-makers in protecting infrastructure and communities. Multiple areas for future research were also discussed. By improving disaster resilience and preparedness, it aims to reduce flood risks in coastal areas.