Final Project Report

Catherine Jackson (ccj3)

Tue., Apr. 30

1 Introduction

1.1 Storm Surge Variability and Literature Review

Hurricanes represent an ever-present threat to US coastlines, and these events have incredible potential to damage infrastructure and threaten lives. In fact, "Billion-Dollar Weather and Climate Disasters" (2024) found that, when analyzing data from 1980 to 2024, damages from tropical cyclones and severe storms account for 51.8% and 17.3% of climate disaster-related costs respectively. These severe storms and their associated surge events must be better understood in order to protect communities from their damages.

However, both the frequency and characteristics of these storms and their associated surges is made more complex due to spatial dependencies. Needham and Keim (2012) sought to create a storm surge database based on historical activity. They argue that this database can be used by communities to better protect themselves based on historical data. They found that the Central and Western Gulf Coast is particularly vulnerable to hurricanes, both in terms of increased frequency and magnitude, as well as storm surge. Comparatively, the East Coast experienced less frequent and smaller events. Furthermore, Xu et al. (2010) hoped to better understand how the incorporation of multi-scale simulations affected surge predictions, but their results also demonstrated that differences in coastal tophography and bathymetry led to large variations in storm surge events. Islam et al. (2021) found similar results. After including information about translational speed and coastal geometry, the predictions from their surge index became significantly more accurate. In this way, Needham and Keim (2012) demonstrates that there are significant variations along the US coastline based on hurricane magnitude and frequency, and the associated storm surge, due to important climate variations. In addition, Xu et al. (2010) and Islam et al. (2021) demonstrated that including information about coastal geometries and bathymetries is important to accurately modeling and predicting surge.

The image below, from Needham and Keim (2012), visually shows the differences in storm surge across the Southern US coastline based on historical information and data.

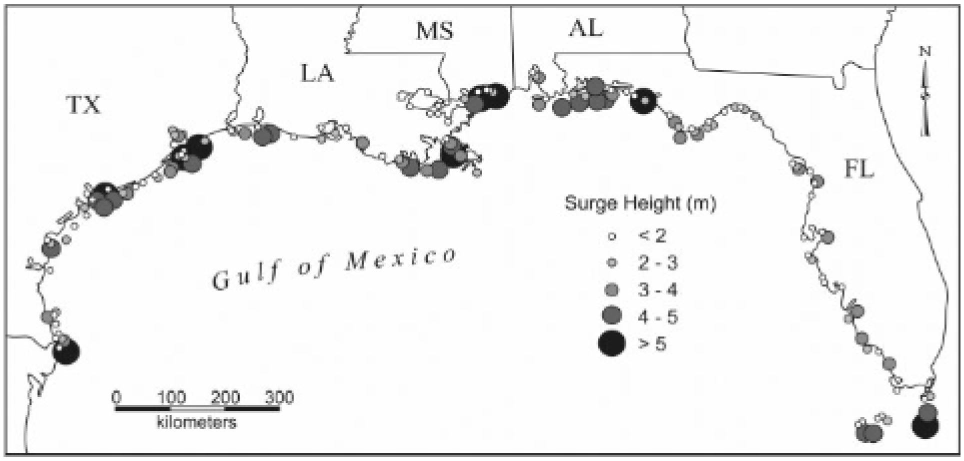

Figure 1. The location and height of the 195 peak storm surges along the US Gulf Coast identified in SURGEDAT.

In conclusion, these papers all demonstrate that physical differences in the coast as well as climate differences in the storms themselves mean that surge behavior is incredibly spatial. To accurately protect communities and analyze risk, the spatial trends must be understood and regional data must be included.

1.2 Problem Statement

In the original risk analysis code, a representative storm surge is randomly generated. This storm surge is represented with a GEV distribution defined by three parameters, each being generated from its own distribution. Most importantly, though, this function was used to apply regardless of the site of interest. The same function was used at many different points across the coastline. As discussed above, however, it is important to include spatial variability given that severe storms and associated surges are highly dependent on the location along the coast and nearby geometry and bathymetry.

The problem that must be addressed, then, is a methodology by which spatial information can be incorporated into the surge function such that the surge is more accurate given the physical characteristics and location of the site.

It is important to address this lack of spatial specificity in the current storm surge generation, within the context of climate risk analysis, as, even if the climate risk assessment methodology is appropriate and efficient, it is dependent on the quality of incoming data. For example, a project could have a statistically sound methodology of determining whether to raise a house given certain flood risks, but if the storm surge distribution used is incorrectly low, the model would incorrectly underestimate the true risk. Improving data inputs improves the decisions produced by this risk analysis tool.

For these reasons, the surge distribution should be as accurate as possible for the specific site of interest. The best methodology to do so is to adjust this distribution towards the observational data from the nearest gage.

1.3 Selected Feature

To address the spatial trends demonstrated in severe storms and associated storm surge, the surge distribution used in risk analysis must be specific to the site of interest. Therefore, the new addition to the decision-support tool must shift the distribution such that it better represents observed data at the nearest gage.

Consequently, this selected feature will be a modification of the code that provides a representative storm surge distribution sample. Note, first, that this code is found in lab code rather than in the HouseElevation.jl file. Consequently, these adjustments are made within the definition of the function itself.

When selecting the best methodology by which the representative distribution can be shifted such that it better represents the behavior at the specific site of interest, I took inspiration from Dr. Doss-Gollin's CEVE 543 class. Bayesian updating allows a distribution to be adjusted as new, observed data is produced. This means that the original surge distribution parameters can be updated using data that comes from the nearest gage, producing distributions more accurate given the physical location of the site.

Again, this improves the climate risk assessment by using Bayesian updating as a methodology to improve the input data. A model is only as reliable as the data it uses to make predictions, so improving the surge based on the actual physical location improves the decision outputs of this model.

2 Methodology

Inspiration for this methodology is drawn primarily from Doss-Gollin (2023). Slides from this course help describe the general process of Bayesian Updating. First, a project might have an original distribution, defined by parameters, each with their own probability distributions. In other words, one of the important parameters might be defined by a Normal distribution (e.g. mean of 5 and standard deviation of 0.5), but it is not a point value. Furthermore, the project might have a reasonable first guess as to what these distributions could be. However, as new data is produced, can be update the beliefs about these distributions so they are more accurate. This is the process of Bayesian Updating.

The slides also define Bayes' rule for distributions as:

$$p(\theta|y) \propto p(y|\theta)p(\theta)$$

Where:

1. $p(\theta|y)$ is the posterior distribution of the parameters given the data
2. $p(y|\theta)$ is the likelihood of the data given the parameters
3. $p(\theta)$ is the prior distribution of the parameters

In other words, the prior is the best guess before any additional data is produced. The likelihood tells us how likely the new observational data is given our beliefs within the prior. We can use the likelihood to update our new beliefs about the distribution.

However, explicitly calculating the posterior can be computationally challenging. Therefore, many models use sampling to reduce this computational load. This concept can be difficult to conceptualize. One analogy is to think of the posterior as a mountainous landscape. Higher peaks represent more likely values of the parameters. The goal is then to find the highest peak. Instead of solving analytically, the model can "walk around" the landscape to find the best values.

Markov Chain Monte Carlo includes many different methods to walk around the landscape, but in general the process is as follows:

1. Start at a random point in the parameter space
2. Propose a new point in the parameter space
3. Determine whether the new point should be accepted or rejected, dependent on the criteria of the specific method

2.1 Implementation

You should make your modifications in either the `HouseElevation` or `ParkingGarage` module. Detail the steps taken to implement the selected feature and integrate it into the decision-support tool. Include code snippets and explanations where necessary to clarify the implementation process.

2.2 Validation

As we have seen in labs, mistakes are inevitable and can lead to misleading results. To minimize the risk of errors making their way into final results, it is essential to validate the implemented feature. Describe the validation techniques used to ensure the accuracy and reliability of your implemented feature. Discuss any challenges faced during the validation process and how they were addressed.

3 Results

Present the results obtained from the enhanced decision-support tool. Use tables, figures, and visualizations to clearly communicate the outcomes. Provide sufficient detail to demonstrate how the implemented feature addresses the problem statement. Use the `#| output: false` and/or `#| echo: false` tags to hide code output and code cells in the final report except where showing the output (e.g.g, a plot) or the code (e.g., how you are sampling SOWs) adds value to the discussion. You may have multiple subsections of results, which you can create using `##`.

4 Conclusions

4.1 Discussion

Analyze the implications of your results for climate risk management. Consider the context of the class themes and discuss how your findings contribute to the understanding of climate risk assessment. Identify any limitations of your approach and suggest potential improvements for future work.

4.2 Conclusions

Summarize the key findings of your project and reiterate the significance of your implemented feature in addressing the problem statement. Discuss the broader implications of your work for climate risk management and the potential for further research in this area.

5 References

"Billion-Dollar Weather and Climate Disasters." 2024. National Centers for Environmental Information; https://www.ncei.noaa.gov/access/billions/summary-stats.

Doss-Gollin, James. 2023. "CEVE 543: Data Science for Climate Risk Assessment." Rice University; https://ceve543.github.io/.

Islam, Md. Rezuanul, Chia-Ying Lee, Kyle T. Mandli, and Hiroshi Takagi. 2021. "A New Tropical Cyclone Surge Index Incorporating the Effects of Coastal Geometry, Bathymetry and Storm Information." *Scientific Reports* 11 (16747). https://doi.org/10.1038/s41598-021-96256-0.

Needham, Hal F., and Barry D. Keim. 2012. "A Storm Surge Database for the US Gulf Coast." *International Journal of Climatology* 32: 2108–23. https://doi.org/DOI of the paper if available.

Xu, Hongzhou, Keqi Zhang, Jian Shen, and Yuepeng Li. 2010. "Storm Surge Simulation Along the u.s. East and Gulf Coasts Using a Multi-Scale Numerical Model Approach." *Ocean Dynamics* 60: 1597–1619. https://doi.org/10.1007/s10236-010-0321-3.