Final Project Report

Lucia Romero-Alston (lmr12)

Tue., Apr. 30

Introductions

Problem Statement

The goal of this project is to improve our flood induced house-elevation policy search model by tackling the effects of uncertainty in sea-level rise.

It is important to be able to assess the vulnerability of coastal communities with respect to flooding so that we can develop solutions that improve people's quality of living and protection.

Flooding is a primary hazard coastal communities have to face, as they are particularly exposed due to development within and the expansion of floodplains. Flooding is becoming more common, even in the absence of extreme weather events. This is detrimental to communities because consistent flooding can clog storm drains, flood streets, inundate and fatigue coastal infrastructure and ecosystems, and affect freshwater supplies .[1] One of the strongest drivers of this increase in frequency and intensity of flooding is human induced climate change. Climate change is having a significant effect on many of the factors that contribute to flooding, specifically within sea-level rise, and this flooding is having detrimental impacts on coastal communities. Even locally, more than a third of Harris county falls within a FEMA-designated floodplain, and according to experts, what is now considered a 500-year floodplain (a 1 in 500 chance of flooding each year) will soon be considered a 100-year floodplain due to the increase in the rate and intensity of flooding .[2] The intensity of climate hazards such as flooding will only increase in the absence of measures to reduce emissions. If this is the case, we must reduce the risk and vulnerability of exposed communities if we want to decrease the damage and destruction of coastal flood impacts. The decision-support tool we are building provides insight into risk management strategies that can be taken in order to reduce risk, where risk is characterized by damage costs. This mitigation strategy is the elevation of a home or building in a flood-prone area, which essentially decreases the effect of flooding on the building and as a result, the damages and resulting damage costs.

Selected Feature

As the climate warms and humans continue to emit greenhouse gasses into the atmosphere, the sea-level will continue to rise at alarming rates. According to the European Environment Agency, the global sea level has already risen about 21 cm since 1900 and the rate at which it is rising is accelerating .[3] This is largely due to the decrease in freshwater impoundment in the face of a

significant increase in ice sheet melt and thermal expansion .[4] Sea-level rise is a major contributor to coastal flooding around the world, and it is one of the elements factored into our policy search model.

In our analysis we must make a lot of assumptions in order to quantify the risk associated with flooding in the area of our home. This estimation of risk is particularly apparent in the definition of the states of the world (SOWs), or scenarios, over which we inspect our damages and costs. The components of our SOWs are sea-level rise, storm surge, and the discount rate, all of which are estimated to a certain degree due to uncertainty in climate change, future economies, and physical dynamics among other things. These estimations can lead to overconfidence in results that are not fully representative of real world happenings. So while it is impossible to perfectly model each component, we can implement more in depth analysis of the individual factors contributing to flooding. To stress the importance of the representation of the SOW components, if we underestimate the frequency or intensity of flooding, we will then underestimate the amount of damages induced, and then the cost of the damages, so in the end, the analysis of these components is the first step in quantifying risk.

In this project I will focus on the improvement of the modeling of one of the scenario components, sea-level rise. The current model is probabilistic and the parameters used to model sea-level rise are estimated and quantified without having their uncertainty accounted for. This is difficult because sea-level rise can be considered a "deep" uncertainty, meaning experts cannot agree on the probability distributions associated with the contributing variables .[5] The estimation of deeply uncertain values as is done in the model can lead to skewed results.

Instead, in this project sea-level rise will be considered using rejection sampling to produce a sample of parameters that are consistent with expert predictions of future sea-level rise, and better account for the uncertainty and the probabilities associated with the contributing parameters.

Literature Review

As previously mentioned, our house elevation decision-support tool is a probabilistic model of flooding and the resulting damages. The main factor that distinguishes this model from a robust decision making model is that it provides a single optimal solution with some quantity that serves as a ranking or comparison to the other possible solutions. This solution though, must be taken with a grain of salt because the downfall of probabilistic models is that they fail to truly consider uncertainty. Sriver et al. discuss the difference between a probabilistic model and a robust decision making model in their paper with respect to sea-level rise and the parameters with ranging uncertainties that contribute to this measurement .[6] The probabilistic analysis, like the one we are modeling, uses the best available information to create one single joint distribution for all uncertain parameters .[6] This means that it uses probability density functions for all of the input parameters instead of distinguishing between well characterizable and deep uncertainty values. This allows the model to search for an optimal strategy for the best-estimate distribution. Again, the downside is that these models do not fully characterize uncertainty, which is why robust decision making models are beneficial. These models almost work backwards, beginning with a policy under consideration and identifying the SOWs where the policy performs best. It uses a single

probability density function to represent well-characterized uncertainties, and then the model is run over an experimental design that is informed by the deeply uncertain factors .[6] The resulting database of model runs can be used to identify the conditions where the decision fails to meet some established criteria, and it will provide more information about things that have limited available information. Interestingly, robust decision making models make no statements about the likelihood of the scenarios and they are additionally beneficial because they establishe regular relationships between analyst and policymaker .[6]

To understand the effects of using a probabilistic or robust decision making model to represent sea-level rise, we must consider the types of uncertainty that influence this metric. As previously stated, sea-level rise is considered a "deep" uncertainty, and to zoom in on why, we must examine the drivers. There are three main sources of uncertainty when modeling sea-level rise. First, the projections depend on scenarios of external factors, such as global warming and emissions, and they often dominate outcomes. We want to be careful about what emission scenarios we choose because the ones used don't always represent absolute bounds and that can lead to overconfidence in our results. The second source is parameter uncertainty and internal variability, which consists of thermal expansion, melting land ice, and changes in ocean topography. The third source is model structure uncertainty, which is common because of the variety in the types of models and their inability to truly represent the real world .[6]

In our analysis we will be using the following Lempert quadratic sea level rise equation:

$$SLR = a + bt + ct^2 + c^* I\left(t - t^*\right)$$

I will expand on how this equation will be implemented in the "Methodology" section of this report, but in order to understand the uncertainties that contribute to sea-level rise we must know what parameters we are using in the calculation. The Sriver et al. paper nicely combines the different types of parameters and their respective uncertainties into a table.

There two sources that contribute to how we characterize uncertainty are how much evidence there is for something and how much people agree on it .[6] There is sufficient understood information about parameters a, b, and c, so they are considered well-characterized and can be comfortably represented using a joint probability distribution. On the other hand, there is little direct observational evidence for potential changes in system dynamics but there is some agreement on the upper bounds of how it will contribute to sea level rise in the next century, so the changes in system dynamics, characterized as c^* and t^*, are classified as "deep" uncertainties .[6]

The results in the papers on sea-level rise were enough to prove that sea-level is a significant factor in flood risk and that the acknowledgement of uncertainty can play a key role in results. According to Oddo et al. even small sea-level changes can increase flood risk by orders of magnitude. They found that while Van Dantizig's linear sea-level rise model predicted sea-level rise in the year 2100 to be ~0.7m, the beta-calibrated forecasts predicted it to be at ~1.2m. This is a significant outcome because it puts Van Dantzig's result below 85% of beta-calibrated forecasts .[5] This goes to show that the inclusion of proper uncertainty characterization in sea-level rise drastically affects the bounds to be analyzed. Applying these new values to the model increases variability

in expected model damages and the expected optimal total cost index by 11%, indicating that the model output was significantly sensitive to changes in sea-level .[5]

Methodology

Implementation

The primary methodology in this analysis of sea-level rise is the application of rejection sampling. This method is used when the goal is to take samples from a complicated distribution. In other words it is a means of approximating a sample from some distribution that is intricate and possibly not properly normalized. In order to sample this distribution we use a distribution that is somewhat similar but much easier to quantify, and run a sort of comparison between the two. This complex distribution, $p(x)$, is known as the target distribution, and the simpler distribution, $q(x)$, is known as the proposal distribution. The principal condition is that the proposal distribution must have the same, or larger, domain as the target distribution, or the support of $p(x)$ must be equal to or within the support of $q(x)$.[7] In the end the goal is to sample from the proposal distribution in order to simulate sampling from the target distribution .[8]

To create the proposal distribution we employ the Lempert quadratic sea level rise equation that was mentioned in the "Literature Review" section of this report.

$$SLR = a + bt + ct^2 + c^* I(t - t^*)$$

In this equation the parameters a, b, and c represent the well characterized process of thermosteric expansion as a quadratic polynomial. The parameter a is a constant that represents the initial sea level, the parameter b represents the linear trend in sea level rise, and the parameter c represents the non-linear trend in sea level acceleration which can be represented by a second degree parametric relationship .[6] This quadratic form has a reasonably good fit to observed historical changes of sea level globally for the past 200 years .[6] The parameter c^* represents the increase in the rate of sea level rise at some uncertain time t^* .[5] Both of these parameters are considered "deeply" uncertain because they account for the poorly understood process of abrupt sea level rise due to changes in ice flow dynamics. Using this equation and these parameters, we can calculate an ensemble of different x, or sea-level in the year 2100. The proposal distribution employs Julia's KernelDensity.jl model in order to represent the probability distribution of these different x values.

```
xq = let
    dfq = CSV.read("data/slr_oddo.csv", DataFrame)
    [Oddo17SLR(a, b, c, tstar, cstar) for (a, b, c, tstar, cstar) in eachrow(dfq)]
end


s = []
for i in 1:length(xq)
    slr_param = getindex(xq,i)
    slr_x = slr_param(2100)
    push!(s, slr_x)
```
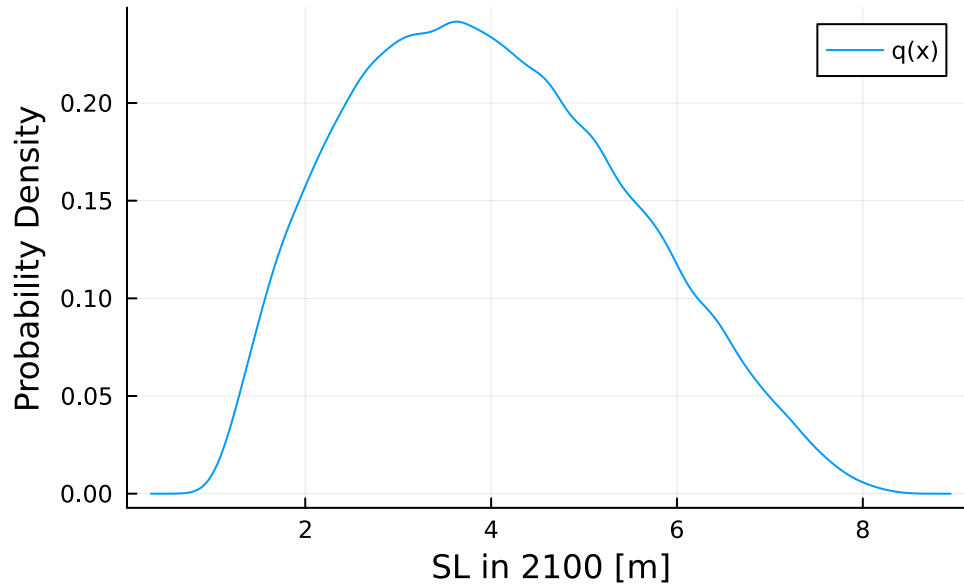
```
end

q = kde(float.(s))

plot(q.x, q.density, xlabel="SL in 2100 [m]", ylabel="Probability Density",
label="q(x)")
```


To produce the target distribution we will establish and rescale a beta distribution so that it is consistent with the domain of the probability density function for the sea-level rise predictions from Oddo et al. study .[5]

```
lb = 0.9973656603828569       # from q
ub = 8.28765880297302         # from q

function target_pdf(x)
    x_rescaled = (x-lb)/(ub-lb)
    return pdf(Beta(2,3),x_rescaled)
end
```

Now that we have both distributions, we must establish a scaling factor that ensures that the proposal distribution is always greater than or equal to the target distribution. This is important because if it is smaller, then there is part of the target distribution that will not be represented in the sampling process. To find this scaling factor we use the relationship:

$$k = max\left(\frac{p(x)}{q(x)}\right) \text{ .[8]}$$

This scaling factor can then be multiplied by the proposal distribution as demonstrated in the graphs below, where the second graph is that where the scaling factor is included.
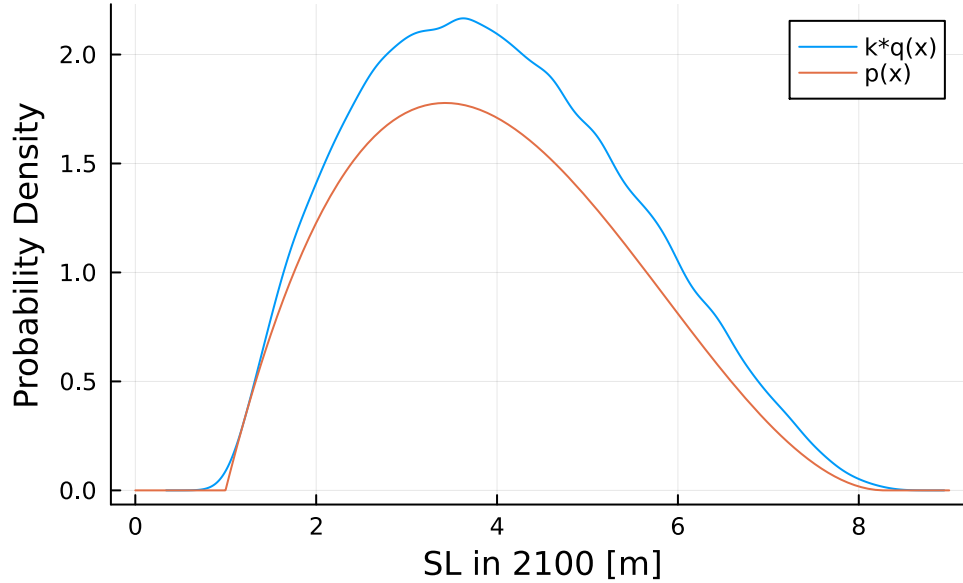
Because the proposal distribution is easier to quantify and analyze, we want to take a sample, s, and consider its likelihood under both distributions. If s is likely to occur in the target distribution but not the proposal distribution, then it is a value that we want to keep. This is because we know it is crucial to our target distribution, and there are lower odds that we will sample it again from the proposal distribution. If s is not likely to occur in the target distribution but is likely in the proposal distribution, we might want to discard it because we are likely to get it again and it is not important to our target distribution .[8] This can be modeled through the following:

$\frac{p(s)}{kq(s)} > u$

In this relationship u is a random value sampled from the normal distribution $U(0, 1)$. If the value of the fraction above is greater than the uniform random variable, then we accept the sample s. If this is not true, then we reject s and continue sampling .[8] This is a process that is done over a large number of iterations, and can be represented by a for loop and an if statement. In the end this produces an ensemble of values of sea level rise that are within the accepted range and appear as though they have been sampled from the target distribution. Using this ensemble we can produce a collection of parameter combinations from which we can sample to generate our states of the world.

```
function sample(size)
    xsample = []
    n = Normal(0.0, 1.0)
    for i in 1:size
```

```
        xs = rand(x)
        cs = rand(n)
        if (target_pdf(xs) / (k .* pdf(q,xs))) >= cs
            push!(xsample,xs)
        end
    end
    return xsample
end
```

Validation

The primary validation technique that I used was that the target distribution was the same beta distribution as the one used by Oddo et al. in their paper. This way, I would be able to compare my results to theirs, specifically, the way in which results changed before performing this rejection sampling procedure versus after implementing it. My results were validated, particularly in the extreme changes that will be explained in the "Results" section. In particular, they mentioned that even small sea-level changes can increase flood risk by orders of magnitude .[5] This was important to my results because the difference was so extreme that there was suspicion that they were due to error. The fact that other researchers had similar results strengthens my data and my results, and similarly, my argument that modeling uncertainty in parameters is important.

Validation and verification of results can be difficult for models like this that attempt to model the future, and things that cannot be experimented. This is a downfall using a probabilistic model to represent things which we do not know for certain. This inability to physically observe and experiment on our results is another source of uncertainty in flooding projection models, and another reason why the "optimal" result of this model must be considered as information, but not taken as law.

Results

What one would expect is that performing a more in-depth evaluation of the elements influencing the scenarios to account for uncertainty generates more accurate and more extreme results. The initial, general version of this house-elevation model had a minimum total cost of $80,637.53 and a minimizer of 10.12 ft in elevation. Comparatively, when we run our elevation model employing rejection sampling with only 1,000 iterations, the minimum total costs are $443,075.67 with a minimizer of 8.6 ft in elevation. And when we run this version with 10,000 iterations, there is a minimum total cost of $453,006.72 with a minimizer of 9.05 ft in elevation. In this analysis it is clear that our consideration of uncertainty and our focus on sea level rise did significantly increase the extremity of damages. This is apparent in the increase in minimum costs by an entire order of magnitude. What is interesting and unexpected is that the minimizing house elevation, or policy decision, decreased. In the general case this minimizing value was extremely high at 10.12 feet, but with the inclusion of sea-level rise analysis, this value dropped by over 10%. The following is a graph of the minimizing policy and the relationship between house elevation and the net present value. This graph is representative of the unchanged model
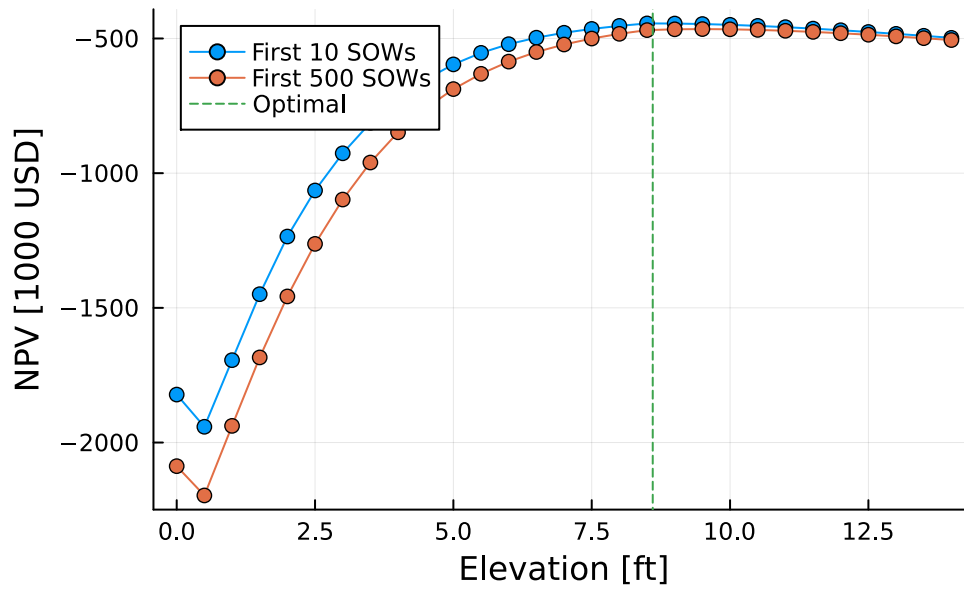
This graph is the same relationship but for the model that considers a more in depth analysis of sea-level rise over 10,000 iterations of rejection sampling.
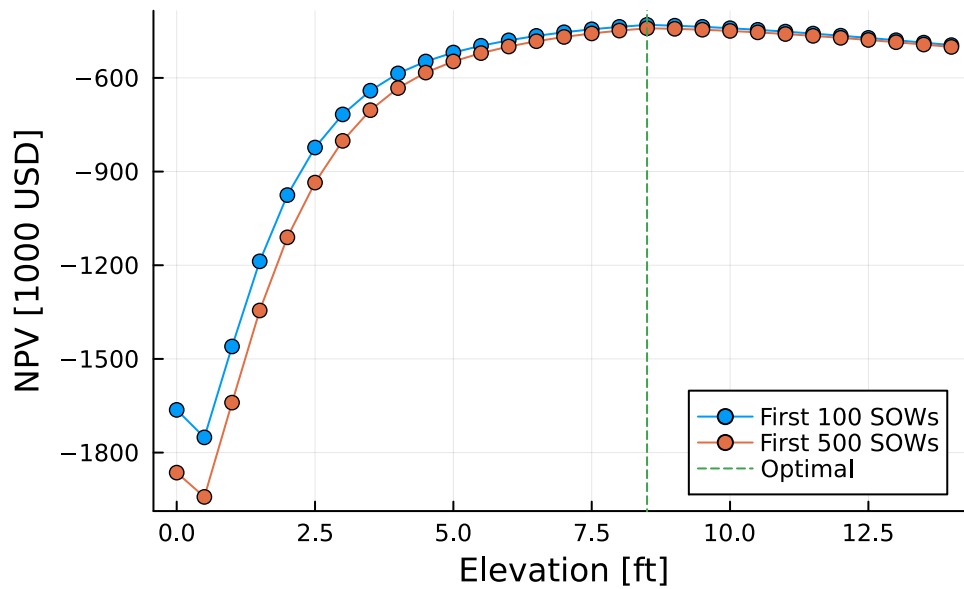
As you can see, the shape of the graphs are pretty similar, and the most drastic change is the predicted costs in damages. This goes to show that even if analysis of the uncertainties doesn't

change the policy decision in house elevation too much, this sort of analysis can be extremely informative about how significantly sea-level rise will affect flooding and damages.

Conclusions

Discussion

What is clear in the results of this analysis is that performing a more in-depth evaluation of the elements influencing the scenarios to account for uncertainty generates more accurate and more extreme results. This is crucial to consider in the context of climate risk assessment because it emphasizes the importance of modeling uncertainty and why we must not overlook it. It is very tempting to model a real world scenario using assumptions and estimations in order to obtain results that give you a specific value and a specific solution. The results of this are a reminder that it is risky to over-simplify a complicated analysis, which is generally the case when applying a probabilistic model. At the same time, it is a demonstration of the ability to be more considerate of uncertainty of parameter values within the probabilistic models, because we are able to compare it to our original model that did not employ rejection sampling for precision.

There were some limitations to this modeling approach in multiple areas. First, it is difficult to quantify certain values, and ensure that the uncertainty behind parameters is fully accounted for. Some deterministic ways of modeling uncertainties reduce complex systems to best guess estimates, but for some forms of analysis this is necessary to avoid overcomplicating the model. It is undesirable to overcomplicate probabilistic models, especially when they represent unverifiable or untestable events like future sea-level rise or future economic parameters. It is also difficult to quantify other factors that might contribute to the overall cost due to flooding. Some studies have tried to assign a tangible cost to the human life of flood victims which is highly variable and deeply uncertain .[5] There are also some limitations to rejection sampling, namely selecting the appropriate proposal distribution .[9] This is difficult because it affects the scaling factor, and the probability of acceptance is dependent on this .[7] So if the chosen proposal distribution is too far off from the target distribution, there can be a significant hindering of the efficiency of the sampling process. Moreover, for rejection sampling we must know the probability distribution function of the target distribution, which can be difficult to produce, and this type of sampling can be inefficient in higher dimensions .[9]

There are a few considerations that were left out of this analysis of flooding and resulting damages costs. Some of these are a result of them being overcomplicated, deeply uncertain systems, or simply information that is not easily accessible. For example, this analysis left out direct consideration of the effect of greenhouse gas forcing projections and regional variations in sea-level projections .[6] Another consideration is cognitive myopia which results from focusing on a single performance method, which in this case is minimizing costs, and so solutions are restricted because we did not explore the full set of tradeoffs between possible objectives .[5] This is something that can possibly be considered in future analysis by implementing a multiobjective framework.

Conclusions

One of the key findings in this analysis is that a deeper investigation and understanding of uncertainty in most aspects of a model can significantly improve it. Considering this informative model on house elevation decision analysis, the effect of only considering sea-level rise a little more was huge. Although the effect on the policy decision was a significant 10%, the change in model outcome that was truly astounding was the difference in the expected total costs in damages. Even under what the model predicts to be the best case-scenario, the cost is a whole order of magnitude larger than what was initially predicted. This is also a significant consideration because this is still under the conditions of being a probabilistic model, and so there are still many parameters that have been estimated and whose uncertainties have not been considered. Even with the more in depth analysis of sea-level rise, we still cannot perfectly predict future events in climate and its reactions.

At the rate that we are experiencing sea-level rise, and that we will continue to experience it, it is crucial to be able to understand the way it affects environmental hazards such as flooding, and as a result, the livelihood of those exposed to it. This project used a framework to more accurately represent future sea-level rise while considering uncertainty to a higher degree than general probabilistic models. This allows us to better represent hazard conditions and develop information on what risk management strategies we can implement to produce beneficial policies for at risk communities.

This is particularly important here in Houston where we experience a high volume and intensity of flooding relatively often. Modeling risk such as flooding is an opportunity to be mindful of equity in engineering and how our inputs and resulting policies affect different demographics. According to the Houston Chronicle, "Rice University's Kinder Institute for Urban Research reported that 26 percent of Harris County's rental units were affected during Hurricane Harvey, whereas 15 percent of homeowner-occupied units were. Similarly, 25 percent of Black residents and 21 percent of Hispanic residents of Harris County were impacted, while 13 percent of white residents were.["[2]"] This goes to demonstrate how flooding and environmental hazards disproportionately affect minority communities. This is important to consider in our analysis and our models because there are unintentional skews in our results that may be consequential in policy making. For example, the model we have produced might recommend the protection of a more expensive home over other less expensive homes based solely on how much they are worth, because the price of the home is a significant driver of our damage costs. If this is not considered at the time of policy making, it is likely to skew aid, funding, and other sorts of mitigation frameworks to wealthier areas and away from communities that have worse infrastructure and are in greater need of resources. Another skew in our model is that it does not consider the lower frequency and intensity of flooding in wealthier areas as compared to other neighborhoods with poor infrastructure.

There is potential for further research in the modeling of a house elevation analysis. A similar in depth analysis can be performed across all contributors to the SOW including the surge distribution and the discount rate. Moreover, an analysis of house elevation decisions in a robust decision making model may be beneficial to help understand how a given elevation height will perform

across more states of the world, and even more extreme cases. This will also allow for the analysis of the different heights under more accurate representations of uncertainty and sampling. It would be interesting to examine how this model can be altered to include regional fluctuations in sea-level, which would be beneficial for coastal communities who are looking to better understand how to protect themselves from unique geographical hazards or invest in unique coastal infrastructure projects IN the end there is a lot of work to be done in the modeling of climate hazards such as flooding, and it is important always to be mindful of how we extracting and administering our projects.

References

1. Silverstein, Kate. "Rising Seas, Flooding Coasts." Climate Central. September 26, 2023. https://www.climatecentral.org/climate-matters/rising-seas-flooding-coasts-2023.

2. Lloyd, Olivia, Wesley Ratko, and Alexandra Kanik. "Why climate change and urban sprawl could make flooding worse in Harris County." Houston Chronicle. January 25, 2023. https://www.houstonchronicle.com/projects/2023/flood-data-harris-county-historical/.

3. European Environment Agency. "Global and European sea level rise." European Environment Agency. January 15, 2024. https://www.eea.europa.eu/en/analysis/indicators/global-and-european-sea-level-rise?activeAccordion=.

4. O'Neill, Ian J., and Jane J. Lee. "Nasa-Led Study Reveals the Causes of Sea Level Rise Since 1900." Nasa. August 21, 2020. https://climate.nasa.gov/news/3012/nasa-led-study-reveals-the-causes-of-sea-level-rise-since-1900/.

5. Oddo, Perry C., Ben S. Lee, Gregory G. Garner, Vivek Srikrishnan, Patrick M. Reed, Chris E. Forest, and Klaus Keller. "Deep Uncertainties in Sea-Level Rise and Storm Surge Projections: Implications for Coastal Flood Risk Management." Risk Analysis 40, no. 1 (2017): 153-168. https://doi.org/10.1111/risa.12888.

6. Sriver, Ryan., Robert J. Lempert, Per Wikman-Svahn, and Klaus Keller. "Characterizing uncertain sea-level rise projections to support investment decisions." PLoS ONE 13, no. 2 (February 2018): 1-35.https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0190641.

7. Peng, Roger D. "Rejection Sampling." In Advanced Statistical Computing (May 2022). https://bookdown.org/rdpeng/advstatcomp/rejection-sampling.html.

8. Tae, Jake. "Rejection Sampling." GitHub. January 25, 2021. https://jaketae.github.io/study/rejection-sampling/.

9. Sachdeva, Kapil. "What is Rejection Sampling?" Medium. February 15, 2021. https://towardsdatascience.com/what-is-rejection-sampling-1f6aff92330d.