FLOODING IN BOSTON

HOW INCREASED FLOODING IN BOSTON WILL AFFECT OUR CITY

Gabriel Honigsberg, Dezhou Wang
CS504, Boston University, Spring 2019

As climate change continues to affect our geography, the city of Boston faces several noticeable threats. For the purposes of our project we looked specifically at flooding. Flooding in the past few years has already grabbed the attention of our local media, as storm surges flooded our coasts, covering our streets in ocean water. As of now, the flooding only lasts a few days, and direct damage is only realized close to the shore. However, forecasts show that as the earth warms we can expect these surges and regular flooding to spread deeper into the city, having a more significant and greater impact.

Data

The city of Boston (through open city data) provides several datasets describing future flooding. The geojson data describes two different types of flooding at three different levels. The levels are 9, 21, and 36 inches. The types are one precent annual flood and sea level rise at high tide. The one percent annual flood set is for a more extreme case when there is a storm surge that may just happen once per year. We used this flood shape for most of our analysis as it covered the most area and demonstrates the most significant damage possible from flooding.

We then had to gather data that we could use to better understand the areas in which increased flooding is expected. First we obtained a dataset we could use to segment the city into smaller areas. A natural fit for this was Boston's zoning-subdistricts dataset which split the city approximately 1,600 shapes representing the city's neighborhoods. Although each subdistrict varied in size, each generally represented a few blocks. We also grabbed data from Zillow, which we used to measure the value of real estate throughout different areas of the city. This was merged with the zoning data by checking to see if the coordinates provided by Zillow landed inside the shape described in the city of Boston geospatial zoning data. We performed this transformation through the use of a python library called shapely that helped select appropriate rows where that coordinate based information could be added.

We also wanted information on how busy areas inside the flood area are. We decided that traffic would be a good measure of this, as it represented how much streets are used throughout the day, independent of whether the neighborhood is residential or commercial. All in all, traffic data demonstrates how much an area is used. The traffic data we obtained came from the Massachusetts Department of Transportation (MassDot), and was downloaded from the public data portal. It was then uploaded to datamechanics.io so that we could pull it from there. Specifically, the MassDot data we decided to us described traffic flow at checkpoints throughout the city, such as ramps, intersections, and tolls. The data contained averages from the past three years for a number of different times of day. For our purposes we decided to focus on peak hours, as we are interested in worst case scenarios. We also added information on whether or not a traffic point was inside a flood area by using the coordinates provided in the MassDOT dataset. This manipulation was performed using the same methods as with the Zillow data.

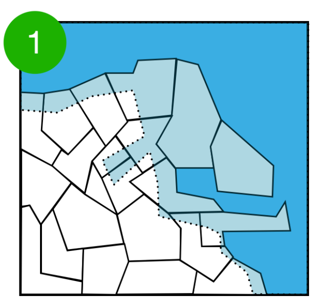
Statistical Analysis

In our statistical analysis of our data we wanted to look at two things. First, we wanted to know how significant the impact of flooding was going to be. The flooding only covered small parts of the city when considering the entire area that Boston covers. That said, flooding for the most part took place in critical areas of the city, meaning that it's impact is quite significant. To demonstrate this significance, we decided to run a correlation, showing the difference between the cities flooded areas, and not flooded areas, where value was based on traffic flow and real estate prices.

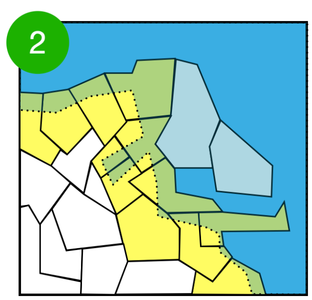
Optimization

As part of project we also developed an algorithm which we could use to optimize the use of limited resources that could be used in reaction to a flooding event. The idea behind the algorithm was to figure out where to create flood boundaries that could prevent flooding from spreading. The challenge was to find how to create boundaries that add up to the shortest distance (that meets the resource requirements) while preventing the most possible flood damage. We also had to consider what flood damage meant, as it is a more complex issue than simply how much area is under water. So, we once again used our traffic and real estate data to quantify how much value was being saved by each potential boundary.

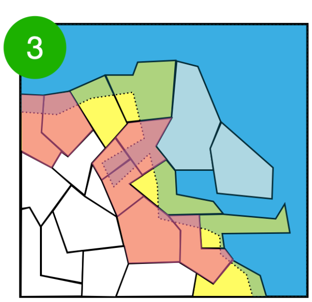
Next, we had to figure out a way to formulate the problem so that it could be solved by a constraint solver such as Z3. We defined the potential state spaces as all the different combinations of subdistricts, where each subdistrict also had an associated value based on its real estate prices and traffic flow at peak hours. The state space was a valid solution if it was possible to build a barrier to protect it using the limited resources which were available. In other words, this meant finding the perimeter of the combined polygons and subtracting the parts of those polygons that were outside of the flood zone. We also had to make sure we were only considering the areas within the subdistricts that were flooded, as the subdistricts along the flooding border weren't entirely flooded. The constraint solver's optimization then had to pick which one of these potential solutions maximized the total value of all the subdistricts included in that solution.



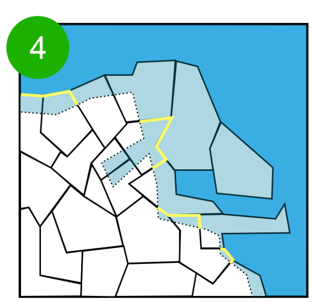
BUILD MAP WITH SUBDISTRICTS AND FLOODING



FIND SUBDISTRICTS ON LAND/FLOOD BORDER



DETERMINE COMBINATIONS OF POLYGONS THAT CREATE MOST VALUE FOR AREA WITH SMALLEST PERIMETER THAT ENTERS FLOOD AREA



GENERATE BORDER THAT FITS RESOURCE REQUIREMENTS AND HAS OPTIMIZED THE VALUE CREATED IN SAVED AREA

The algorithm is flexible, as it can be applied differently based on the city's needs. For example, they might want to consider population, age of population, or public health data. So, although we used traffic and property data, anyone can adjust those inputs easily, so long as the value can be associated with a certain area. Additionally, the outputs can be adjusted. In our case we thought that someone might want to build a wall or barrier to protect against flooding. But, a city with limited resources might decide to deploy a different strategy. Either way, the algorithm can consider any sort of geographic output that is desirable.

Conclusion

The most exciting part of this project was learning to understand spacial data, especially when considering it in terms of a constraint satisfaction or optimization problem. It was interesting also, to use different types of location based data (single points, and shapes) and needing to make decisions based on that. Overall, we believe most of our methodologies are applicable to much more than just flooding, traffic, and real estate. The principles we used are generalizable to any government that is facing impacts from climate change, and needs to make decisions based on how it values its areas, and the forecasts that describe how these are areas are going to be affected.