Evaluating Coastal Erosion

Danielle Lamb

DSC680-T301 Applied Data Science

Professor Catherine Williams

Business Problem

As coastal areas are becoming an increasingly popular place to live, more development moves closer to the coast. A result of this is the destruction of natural protective ecosystems from waves, hurricanes, and tsunamis. Because of this, there is a greater risk of coastal erosion which could damage coastal structures or could be disastrous in an extreme environmental event. The goal of this project is to evaluate the status of coastal areas into categories of low and high risk of erosion. From this, additional resources could be allocated to build protective structures such as jetties and sea walls.

Background

There are numerous natural ecosystems that provide immense protection of coastline from erosion by waves and more severe storms. With ever changing coastlines as a result of industrial development and tourism, these protective ecosystems are quickly being eradicated. In addition to increased coastal erosion, several species of plants and animals become more at risk for extinction as their habitats are destroyed.

Coral reefs are a warm water, high light ecosystem. These ecosystems feature both hard and soft corals which make up the main structure of the reef. Hard corals create texture on the bottom of the ocean, hills and mountains underwater. This disturbance allows waves to be broken up and does not allow the quick building of wave height. With the introduction of pollutants from industrial development near coastlines, harmful algal blooms and increased carbon dioxide in the water causes reef environments to die.

Mangrove forests are another tropical-focused ecosystem. “The world’s area of mangrove forests has been reduced by about 35% on a worldwide scale since the 1980s, and 2.1% of the existing worldwide mangrove area is lost each year,"(Valiela, Bowen, York, 2001) . The mangrove trees live within the coastal environment with their long, spindly, roots extending through the water into the sand. Because of the typical density of these trees, the wave action from the ocean is normally stopped before hitting the land-to-sea interchange. This means with higher and choppier waves in extreme weather events such as hurricanes, these damage causing waves are not able to make it to shore. Without this, these damaging waves can wash away beaches and cause damage to coastal structures.

Saltmarshes are an expansive ecosystem that is flooded and drained with ocean water by the tides. These ecosystems are explosive with diversity and are found on every coast of the United States. Most marshes include numerous species of reed-type plants in addition to shellfish such as oysters. This mixture provides the coastline with a combination of protection from waves and wind. "Salt marshes also protect shorelines from erosion by buffering wave action and trapping sediments. They reduce flooding by slowing and absorbing rainwater and protect water quality by filtering runoff," (NOAA, 2013).

Oyster reefs are a collection of oysters along the bottom of the coastline. One of the largest expanses on the east coast is within the Chesapeake Bay. Not only do the oysters provide similar physical protections as coral reefs, but also provide water filtration to reduce the overall amount of pollutants within the water column. Because of their building block like layering, the softer sediments below are protected from erosion by wave action and storms. "When oyster reefs are only used as a place to harvest commercial oyster meat, they can become degraded," (NOAA, 2021), and are no longer able to provide the same protections to coastal erosion.

These crucial coastal environments are being depleted at an alarming rate due to industrialization. Because of this, the cost of extreme weather events will continue to increase.

Protective measures can be put in place to protect existing developed coastline just as jetties, groynes, seawalls, and sand dune development. Although these methods work temporarily, by destroying more of these environments, the coast as a whole becomes a larger threat to washout by extreme weather events.

Data

Data for this project was found through the National Oceanic and Atmospheric Administration. Records from 1950 to 2021 feature information on severe weather events or unusual weather events. These records were taken by the National Weather Service (NWS). The portion of data for this project began in January 2000 and ends in December 2021. Initially, there were a total of 1,244,206 records over the 11 years with 51 features. The target of this data is property damage by dollar amount.

Methods

A total of 21 data files were combined and concatenated to create one cohesive dataframe from 2000 to 2021. Features were then selected based on need of the project. These features included date, location, type of weather event, and monetary amount of damage to property. Due to the nature of this project, locations strictly on the coast were chosen as well as weather events that related to coastal action such as flooding, hurricanes, and storm surges.

The target feature, property damage, was evaluated through a scatterplot and boxplot to visualize this data and begin outlier detection. Using the interquartile range method, outliers were removed. This method was used twice before all outliers were removed.

Once outliers were removed, data from year and damage_property feature was entered into the Gaussian Mixture model. The number of ideal clusters was determined to be 2, since there was a need for a low risk and high risk in addition to having only two peaks in the distribution. The cluster was added to the dataframe and plotted as a scatterplot. Those in the high risk cluster were pulled into a separate dataframe.

Assumptions

The gaussian mixture model assumes that the data is a combination of several normal distributions. Once cleaned, this dataset featured a bi-model distribution which could assume two normal distributions.

Analysis

With the majority of data points below the $100,000 mark, many highly damaging events were considered outliers. Once cleaned, 9 features were left with a total of 152 records. A total of 61 records were removed as outliers using the IQR method as seen in Figure 1. While these records were removed, because they fall above the high risk cluster, one can assume these areas could also benefit from additional coastline protective measures.

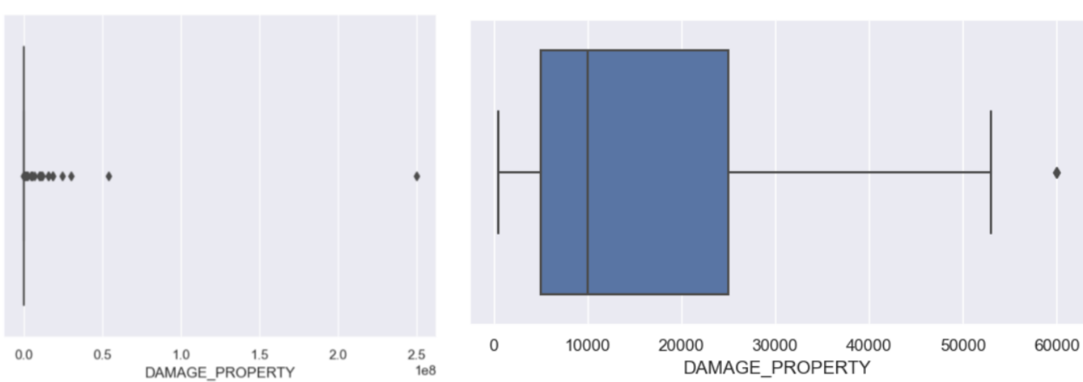

Figure 1: left; boxplot of original data, right; boxplot of cleaned data

Two clustering methods were considered after data cleaning, k-means clustering model and the Gaussian Mixture model (gmm). When considereing the k-means model, the assumptions are the clusters are spherical and of similar size. Looking at the distribution of the cleaned data, the point of the second speak was much lower than the fist, indicating that the size of the second cluster could be much smaller than the first (Figure 2).

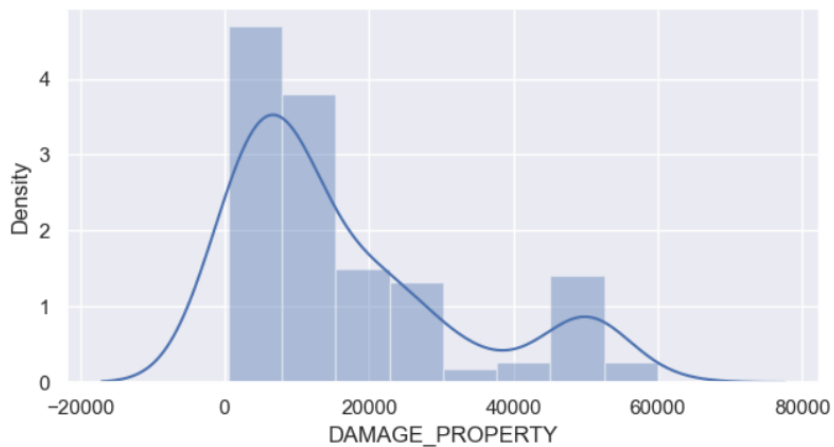

Figure 2: Histogram of cleaned data. One large peak at $10,000 with an additional small peak at $50,000

Looking at the scatterplots of the cleaned data, there was no obvious trend within the data over time (Figure 3). However, there is a large focus of data from zero to $30,000, and an additional smaller focus around $50,000. From the distribution plot (Figure 2), there was a suggestion of a large concentration of points near zero, and a smaller concentration of higher points, which Figure 3 confirms. These clusters are not as obvious when looking at subplot A, but is more apparent when property damage is on the X-axis in subplot B (Figure 3). With this confirming the smaller cluster size of the high risk group, the Gaussian Mixture model assumptions fit this data better than k-means.

A. B.

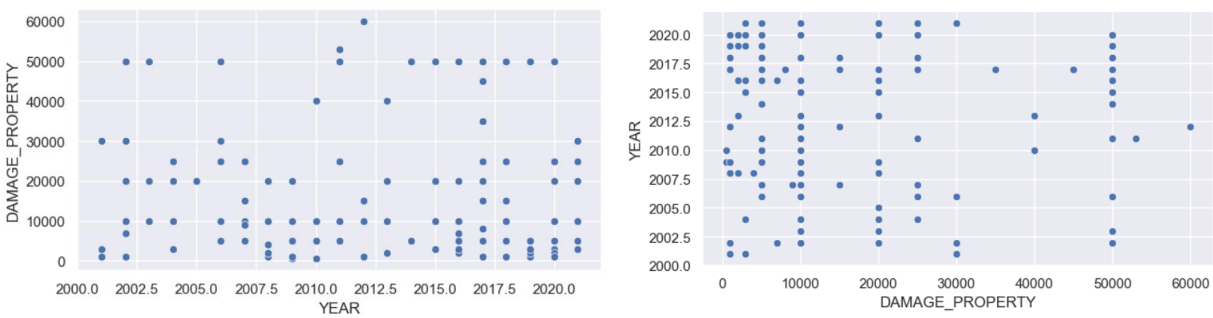

Figure 3: A. scatterplot of year and property damage in dollars. No trend is obvious. B. scatterplot of property damage in dollars vs. year, clear clusters near $15,000 and $50,000.

Using a n_components of two to get two clusters, the Gaussian Mixture model concluded that those points above $30,000 were high risk, and those below were low risk (Figure 4). With a Silhouette score of 0.73, this model and number of components define this data well.

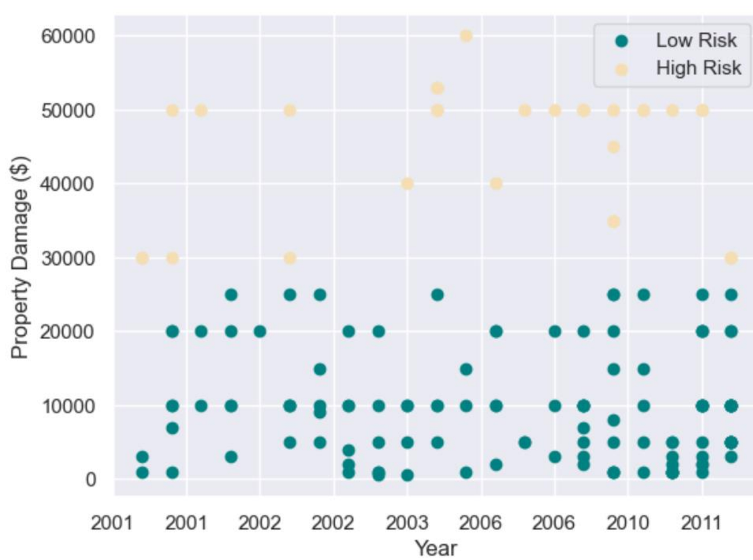

Figure 4: Scatterplot of cluster number over year and property damage. Model concludes those $30,000 and over are considered high risk.

The points in the high risk cluster were mainly found in Mississippi and New York. This indicates that both states could have the potential of improving coastal wetland health or adding man made coastline protection.

Conclusion

While no damage is the best case for extreme weather events, those events with damage above $30,000 were considered more high risk areas. These areas could be less protected along the coastline and potentially more developed with towns and cities. While the outliers of this study were removed for the clustering model, they all were outside the bounds on the higher end, so the high risk label could also be used for these observations. The most frequent weather event was flooding and flash flooding, indicating that areas in Texas, New York, Rhode Island, Mississippi, Florida, and Hawaii are all inadequately protected from flooding. Coastal wetlands play an important part in absorbing large amounts of water from the ocean and storms. If these states could focus on returning lost land to these ecosystems, flooding cost could decrease.

Limitations

This data only had a name associated with the weather event. Because of this, when evaluating which weather situations to choose, types such as flood and flash flood could not be determined if the original cause was ocean related. After cleaning, 10 years of data resulted in 152 observations. More data concentrated on coastal weather events would provide the model more data to be more accurate in determining the threshold of low and high risk.

Challenges

Filtering through the data was difficult. Out of over a million records, only about 215 were considered coastal and possibly ocean related. There was a need to adapt the plan moving forward after cleaning to involve clustering methods that do not need overly large amounts of data.

Future Uses

The future of this study could look at the differences between the damages through time. If additional measures are taken to ensure the protection of ecosystems that provide natural protection or the addition of man made structures, there is a potential for a decrease in damage cost along the immediate coastline.

Implementation Plan

Results from this study can help locate areas of the United States that are more at risk for damage from extreme weather events and coastal erosion. This allows for a better understanding of the severity of damage over the years and how it can be improved through natural and man-made structures. In addition, this study can describe which event types cause the most damage so that greater protection methods can be proposed for counties in need of better protection.

Literature Cited

NOAA. (2013, June 01). What is a salt marsh? https://oceanservice.noaa.gov/facts/saltmarsh.html

NOAA. (2014, February 18). Retrieved February 07, 2022, from https://www.ncei.noaa.gov/pub/data/swdi/stormevents/csvfiles/

NOAA. (2021, October 26). Oyster Reef Habitat. Retrieved January 30, 2022, from https://www.fisheries.noaa.gov/national/habitat-conservation/oyster-reef-habitat

VALIELA, I., BOWEN, J. L., & YORK, J. K. (2001). Mangrove forests: One of the world's threatened Major Tropical Environments. *BioScience*, *51*(10), 807. doi:10.1641/0006-3568(2001)051[0807:mfootw]2.0.co;2