Oceans Rise, Properties Fall

Alex De Mouy, Malcolm Evans, Jacob Kasner, Fernando Narbona

EXECUTIVE SUMMARY

In the last 60 years, sea level has risen 6 inches nationally, according to Sea Level Rise, a national nonprofit institution. Although a slow rate over that timeframe, the rate is expected to increase in the coming years. In the next 20 years, the seas will rise by another 6 inches nationally. Now referred to as sea level rise "ground zero," South Florida is known to have low elevation levels and a large coast. Sea levels in Florida are projected to be increasing by an inch every 3 years, greater than 1.5x the national projected increase.

As recent as 2019, 3.5 million people in Florida are estimated to be at risk due to coastal flooding. As the risks and rate of sea level rise increase over years, the number of people and buildings at risk due to flooding will increase substantially.

Within 10 years, sea level rise will affect every county in the Miami, Cape Coral, North Port, and Tampa Metro regions. This is an area of 22,456,309,262 square meters at risk, represented by $351,493,105,459 value of property, housing a population of 2,093,815. And the situation will only worsen in the next century if no action is taken. Infrastructurally, this will also cause overwhelming migration to Georgia, Texas, California, Tennessee, and Virgina, among others, with direct impact on each of those states' economic resources. Additionally, by 2021, 6 of 32 affected coastal counties will chronically be more than 90% inundated with water. In others words, sizeable parts of Florida will cease to exist if government and developers do not work together to build policies and practices that actively combat sea level rise, including but not limited to considerations of climate change policies and sea wall infrastructure development.

The models underpinning this project should continue to be expanded and applied to other coastal states to better understand regional and nationwide potential impact that may begin to warrant consideration for federal recommendation, upon similarly trended new findings. Additional data, such as carbon emissions, should also be incorporated into this model to specifically build climate change response recommendations by measuring how effective various climate change responses could be in addressing sea level rise holistically. Additional reporting on societal issues in relation to sea level rise should be also be considered to better understand potential for full displacement of people, causing them to become unhoused, or other such harm. Gentrification must also be considered and reported on, in this model to get a better understanding of the potential for full displacement derived from sea level rise. Supporting the continued work of this research could benefit proposals to larger governments and development offices, seeking to build the aforementioned policies and practices.

TABLE OF CONTENTS

CONTEXT

For hundreds of years, we have decided to sacrifice our environment in exchange for economic "progress." Years of unethical environmental practices have contributed to climate change that poses significant risks, including gradually rising sea levels, which will soon force existing coastal areas and their inhabitants to move inland. This will certainly affect the real estate market, one of the largest and most critical asset markets contributing to global wealth. In the United States, this encompasses approximately 250 cities at risk. How might we quantify the potential economic loss for the housing market by analyzing the consequences of predicted sea level rise (SLR)?

As land is invaded by rising sea levels, housing is made uninhabitable. Individuals and businesses, both those that develop and purchase this real estate, incur high risk. This also increases the number of individuals and families at risk of becoming unhoused.

In other parts of the world, people have already started to experience this and FiveThirtyEight writes that "climate change migrants are stuck in a legal and political gray zone. Our systems aren't designed for them. Changing those systems would be complicated, inconvenient, and expensive. Efforts to create frameworks to take them into account run into opposition."

94.7 million Americans, or 28.6%, live in coastal regions and therefore could migrate inland following climate disaster claiming their homes, so while these people are not climate refugees, landlocked states bordering coastal ones are not ready for a mass influx of citizens due to interstate migration.

"Our results show that what we do today will have a huge effect in 2300," Alexander Nauels, a researcher at the Potsdam Institute for Climate Impact Research, said in a news release. "Twenty centimeters is very significant; it is basically as much sea-level rise as we've observed over the entire 20th century. To cause that with only 15 years of emissions is quite staggering."

With regard to the current situation in America, Forbes asserted in late 2019 that "Many of these coastal areas have some of the most expensive properties in the world, including Miami Beach, the Hamptons, and Malibu Beach, among others. More than 99 percent of today's population in 252 coastal towns and cities would have their homes submerged."

America is clearly in danger of losing a massive chunk of real estate. Florida, in particular, is notorious for low elevation levels and many coastal cities. Our project seeks to identify the at-risk coastal areas and estimate total property loss across different points in time based upon projected SLR and incentivize developers and city policymakers to invest inland and rebuild consciously after flood damage. The time for action is now.

PROJECT DEFINITION

BUSINESS SOLUTION

Year after year, we see sea levels rising, primarily because of high carbon emissions levels, which will force existing coastal areas and their inhabitants to move inland.

In the last 60 years, sea level has risen 6 inches nationally, according to Sea Level Rise, a national nonprofit institution. Although a slow rate over that timeframe, the rate is expected to increase in the coming years. In the next 20 years, the seas will rise by another 6 inches nationally. Now referred to as sea level rise “ground zero,” South Florida is known to have low elevation levels and a large coast. Sea levels in Florida are projected to be increasing by an inch every 3 years, greater than 1.5x the national projected increase.

As recent as 2019, 3.5 million people in Florida are estimated to be at risk due to coastal flooding. As the risks and rate of sea level rise increase over years, the number of people and buildings at risk due to flooding will increase substantially.

SOLUTION SCOPE

As land is invaded by rising sea levels, housing is made uninhabitable. Individuals and businesses, both those that develop and purchase this real estate, incur high risk. The amount of individuals and families at risk of becoming unhoused also increases. In reviewing projected economic losses, state and city municipalities, as well as property developers, should be incentivized to take action.

Ultimately, building and city developers will be able to answer various questions such as:

- Where might houses be built to avoid rising seas? Should we invest in elevating houses right now?
- What monetary losses can we expect?
- How many tracts and counties will face almost complete flooding in the near future?

DATASET DESCRIPTION

DATASETS SUMMARY AND TABLE DESIGN

The following table summarizes the main characteristics of the datasets that we will take into account throughout the development of our project:

Fig. 1. Summary of the main characteristics of the datasets

Dataset Name	Features	Source	Table(s)
Cartographic Boundary Files - Shapefile	State, County, Tract, Land Area, Water Area	United States Census Bureau	M_county_tracts_population_housing_units.csv
Population by Tract	Population, Tract	usboundary.com	M_county_tracts_population_housing_units.csv
Housing Units by Tract	Housing Units, Tract	usboundary.com	M_county_tracts_population_housing_units.csv
County Codes	County,Jurisdiction,County FIPS,Place Feature-ID	Florida Department of Revenue	M_county_tracts_population_housing_units.csv
Redfin Metro Migration	County, Nearest_Metro, Latitude, Longitude, Destination City, Origin_City, Destination_State, Percent_Leavers	RedFin RealEstate Brokerage	metro_mig_percents.csv, M_cty_dest_pct.csv, cty_pct_pivot.csv
Flooding & Economic Impact	Population, Housing Value, Inundation, Property Tax, Number of Properties	Union of Concerned Scientists	M_community_impact_cty_nm, M_inundation_counties, county_inundation_pivot
U.S. Census Bureau County-to-County & State-to-State Migration	county-to-county/state-to-state statistics (plus, minus, net migration)	US Census Bureau	State_to_State_Migrations_Table_2019 copy.csv
National Elevation Dataset	Geographical Feature (Name, ID, Type), County, Elevation, Latitude and Longitude	USGS	FL_Features_20201101.txt, M_elevation_nb_ipynb, M_elevation.xlsx
American Community Survey - Housing Characteristics	Monthly Owner Costs, Gross Rent, Property Value, Housing Occupancy	United States Census Bureau	ACSDP52019Y.DP04_data_with_overlays_2020-12-23T150608.csv (10 files - going back to 2010)

SOLUTION ARCHITECTURE

TOOLS USED

- Google Data Studio
  - Data Blending
  - Interaction Between Charts
  - Geocharts & Bubble Maps
- Python
  - Pandas
  - Matplotlib
  - Scikit-Learn
  - Facebook Prophet
- Jupyter Notebook

RESULTS

DATASETS WRANGLING

The dataset wrangling process was divided into three phases: cleaning, enriching, and joining:

- Datasets Cleaning

  All data from the Union of Concerned Scientists used Excel to organize the information with separate sheets for each year and prediction level (high, medium, low). With this in mind, we created a new spreadsheet and organized it traditionally (CSV friendly) by copying the same data and adding “Year” and “Prediction Level” as columns. Before we started analysis, we filtered down to only the high prediction level to envision the worst case scenario. This seemed like the best choice because of the severity of climate change which commonly surpasses expectations.

  The Flooding & Economic Impact dataset does not include rows for any of the 35 counties (mostly inland or northern ones) which are not predicted to experience the effects of sea level rise by 2100, the upper time bound for our analysis. The site-specific “UAID” column was also dropped from both the Population by Tract and Housing Units by Tract datasets from usboundary.com. Finally, the “STATEFP,” “TRACTCE,” “AFFGEOID,” “GEOID,” and “LSAD” columns, representing the”[c]urrent state Federal Information Processing Series (FIPS) code,” “[c]urrent census tract code,” “American FactFinder summary level code + geovariant code + '00US' + GEOID,” “[c]ensus tract identifier; a concatenation of current state Federal Information Processing Series (FIPS) code, county FIPS code, and census tract code,” and “[c]urrent legal/statistical area description code for Census tract,” respectively, were dropped from the Cartographic Boundary Files - Shapefile dataset; all of this data was either unnecessary to our efforts, redundant, or a mix of concatenated and redundant, in a way that made it useful to drop these entire columns.

  From the National Elevation Dataset (USGS), we ran a query for data on all Civil features in Florida which resulted in 1,109 rows. Then we grouped all the features by county and calculated the median elevation to give us a single representative value. Extraneous columns such as “Feature Name”, “ID”, “Entry Date” were dropped from the final data frame.

  For State-to-State Migration data, manual extraction was used due to the inconsistency and formatting of data.

  For median house pricing used for pricing predictions, Pandas was used to identify target columns, as well as replacing missing values, casting columns to data type, stripping inconsistent string characters, and data aggregation (exporting clean data to new spreadsheet). The result was time-series data for each respective tract in Florida with its yearly median value estimate for 10 years (2010-2019).

- Datasets Enriching

  For the Redfin Metro Migration data, we first created a table with latitude and longitude for every county plus seven metros. Then, we used a K-Nearest Neighbors model to group counties by the nearest metro. This split Florida into 7 sections each labelled by the corresponding metro (Miami, Cape Coral, North Port, Tampa, Orlando, Jacksonville, Tallahassee). The first 6 were chosen due to their availability within Redfin data and Tallahassee was included for completeness and having a better representative of the western panhandle. For the Flooding & Economic Impact data, we changed "$##" strings to integers with (in USD) the header. In regards to the Population by Tract and Housing Units by Tract datasets, we renamed columns to clearly indicate "Tract" and either "Population" or "Housing Units," respectively, and removed alpha characters from Tract values, filling empty Population and Housing Units cells with '0,' grouping by Tract and summing, reindexing on Tract. Finally, for the Cartographic Boundary Files - Shapefile, we renamed columns to clearly indicate "Tract," "County," "Land_Area," and "Water_Area," reindexing on Tract.

- Datasets Joining

  We joined the "Cartographic Boundary Files - Shapefile," "Population by Tract," "Housing Units by Tract," and "County Codes" datasets. Mapping of inundation by county to the aggregated land/water area yielded values for land area at each year marker (2030, 2045, 2060, 2080, 2090, 2100) demonstrating a continuous reduction.

EXPLORATORY DATA ANALYSIS (EDA) & MODELING

Once our datasets were defined and our wrangling completed, next was to go into the house pricing data. One caveat of real estate data is that it is highly privatized, and so the US Census Bureau was the best data source we found. We wanted to see which features were most important. In order to do so, we reduced from 150 features to about 60 using manual data screening and domain knowledge. Next we performed generated correlations across the data, and quickly figured a lot of the columns were collinear. Using the SelectKBest function from sklearn.feature_selection module, we are able to pass a parameter that tells the function how many best features to reduce to. After trying multiple values, our best results came from keeping 10 features which we still found to have some collinearity, and using SelectKBest for further reduction was not proving useful. Despite this, we proceeded with our reduced set of features, understanding the nature of our data, and its limits.

The last step to fit our model was in standardizing our variables, and running a simple linear regression to fit our model; this resulted in a R^2 value of ~ 70% and an RMSE ~ $93,000.

- Note: despite this we achieved a roughly decent model, capable of showing that our model could fit to a function, which we could take further and run through a more advanced time-series model.

Our hopes were to predict house prices far into the future and better estimate the total loss in later years. To attempt this, we decided on using FB Prophet, a Python library that enables the user to predict values for time-series data. The two reasons for choosing this library specifically were 1) time constraints in learning complex regression models, and 2) the convenience of having various hyperparameters, including seasonality already baked in.

- Note: When we tested across all tracts, the data was proving difficult to test, as the run time would push to over 4 hours; we decided it was more effective to reduce the data set to a County level as opposed to more granular at the tract level.

Once we began testing (with our county level data), it was evident that with just 10 years of data (1 data point per year), it would become difficult to predict 80 years into the future, but by implementing a smoothing algorithm as a preliminary step, we figured it would improve. This algo works by replacing outliers by the mean value of the series, which, if not treated, would give the model a difficult time in making predictions.

Secondly, the algo also loops over a combinatorial matrix of hyperparameter values, considering a multitude of configurations for our model. This was done since the smoothing functionality of the algo alone was not reducing the sensitivity of the model and thus output trends that were unrealistic, like negative trends which would predict either very low or negative house prices.

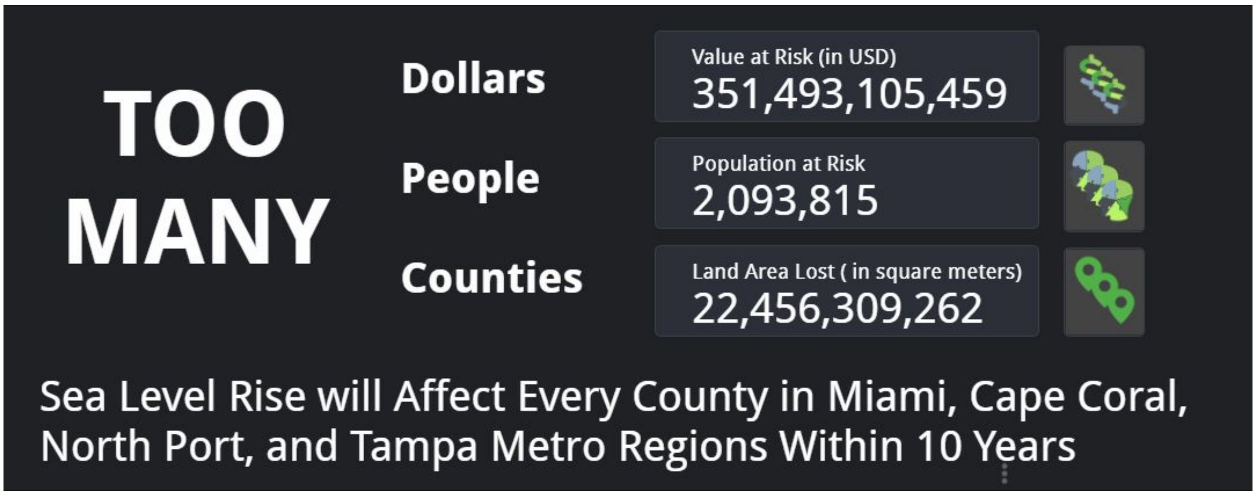
Thus, we were able to obtain values, but FB Prophet turned out to be unable to predict accurately for about 13 of our 68 counties. This is something that we were at least proud to be able to build, and includes future work to accurately predict house prices, so we can estimate properly our economic loss outward on our timeline from 2010-2100. Despite this we had other relative data we still were able to use to deduce our estimated loss.

FUNCTIONAL DASHBOARD

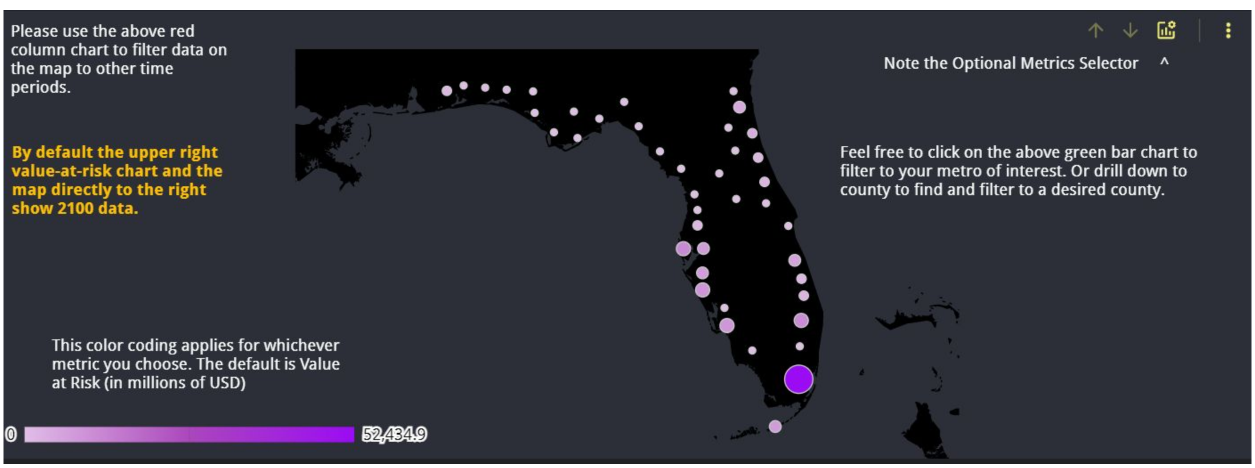
[Team 16 Dashboard › Top Insights]

The dashboard was created in Google Data Studio as a cloud-based informational tool aimed towards state and local government officials and developers. The included maps and charts are fully interactive, allowing users to easily filter on specific locations, years, and metrics. Sections are broken down as follows:

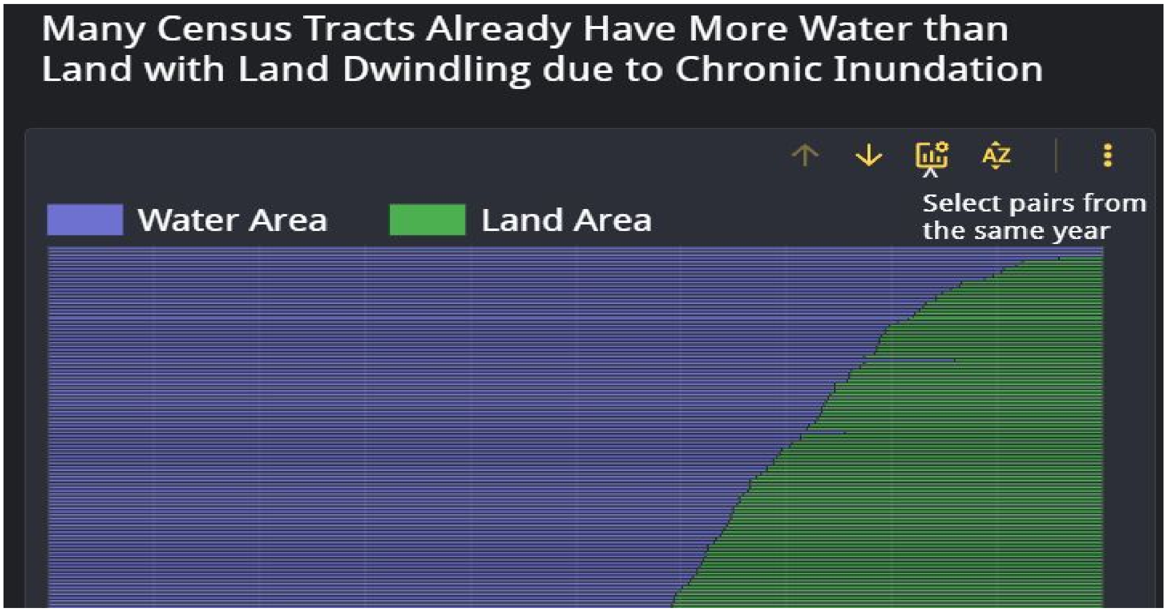
TOP INSIGHTS



FLOODING IMPACT MAPS



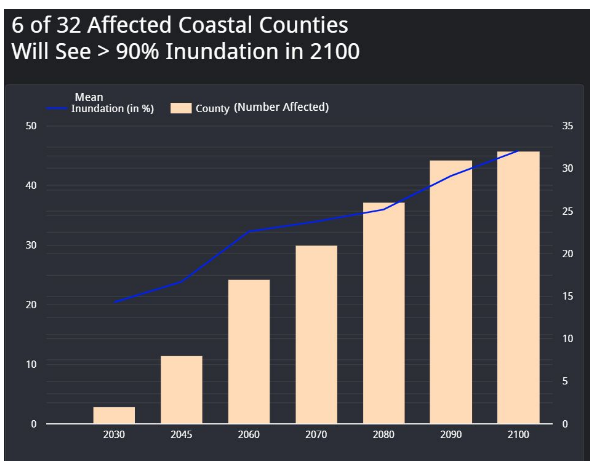
POPULATION DENSITY & PERCENT WATER AREA

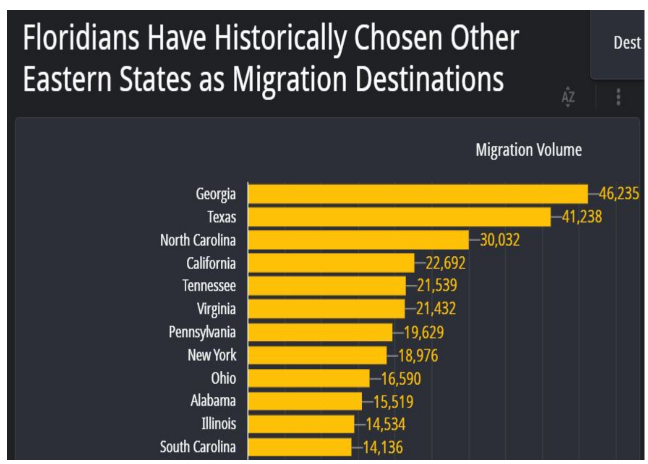


DATA SOURCES

Sources included the US Census Bureau, Union of Concerned Scientists, and Redfin Real Estate.

APPENDIX (ELEVATION & INUNDATION | STATE-TO-STATE MIGRATION)





[Relevant Image of SoFlo](#) + [Source Article](#)

CONCLUSIONS

Within 10 years, sea level rise will affect every county in the Miami, Cape Coral, North Port, and Tampa Metro regions. This is an area of 22,456,309,262 square meters at risk, represented by $351,493,105,459 value of property, housing a population of 2,093,815. And the situation will only worsen in the next century if no action is taken. Infrastructurally, this will also cause overwhelming migration to Georgia, Texas, California, Tennessee, and Virgina, among others, with direct impact on each of those states' economic resources. Additionally, by 2021, 6 of 32 affected coastal counties will chronically be more than 90% inundated with water. In others words, sizeable parts of Florida will cease to exist if government and developers do not work together to build policies and practices that actively combat sea level rise, including but not limited to considerations of climate change policies and sea wall infrastructure development.

Despite our ability to use our data's features to the best capacity, collinearity exists in the most robust dataset we could find; real estate as aforementioned is highly privatized, and finding the data we needed would be costly. However, we were able to reduce the number of features substantially, and successfully fit our data using linear regression with an R^2 of 70%. In an attempt to predict future house prices, we were partly successful, and able to accurately predict for only about 13 of 68 counties. Reducing to counties was necessary for the timeline of predictions we wanted, and the time it took to complete the model to run (approximately 4 hours!).

FUTURE WORK

The models underpinning this project should continue to be expanded and applied to other coastal states to better understand regional and nationwide potential impact that may begin to warrant consideration for federal recommendation, upon similarly trended new findings. Additional data, such as carbon emissions, should also be incorporated into this model to specifically build climate change response recommendations by measuring how effective various climate change responses could be in addressing sea level rise holistically. Additional reporting on societal issues in relation to sea level rise should be also be considered to better understand potential for full displacement of people, causing them to become unhoused, or other such harm. Gentrification must also be considered and reported on, in this model to get a better understanding of the potential for full displacement derived from sea level rise. Supporting the continued work of this research could benefit proposals to larger governments and development offices, seeking to build the aforementioned policies and practices.

REFERENCES

South Florida PBS Documentary

SeaLevelRise.org

Forbes 2050 Opinion

Forbes on At-Risk Property

Elsevier Historical Case-Study

Associated Press on Coastal Property

Sea Level Rise & Mortgage

THANK YOUS

Correlation One, Yash Desai, Spencer Stebbins, Kadija Yilla