Research Project - Results and recommendations (Team A)

APAN 5900 Solving Real World Problems with Analytics

Predicting the Financial Risks of Real Estates Owned by REITs Affected by Hurricanes

Team Members: Raven Feng, Zixin Chen,Yicheng Shi, Zilin Zeng, Zehua Rong, Jingshu Yang, Aansh Mehta, Maddie Tremblay, Romauli Butarbutar

12/07/2021

Results:

Background：

Standard & Poor Global Market Intelligence segment integrates research and analytical capabilities using data to answer clients' questions. Through a comprehensive environmental performance profile, the Global Trucost segment analyses risks related to climate change, natural resource restrictions, and other ESG concerns. Driven by these two business interests, our project is mainly to help our clients to find the relationship between the financial impacts and climate change by collecting raw data and developing forecasting models for our clients. According to the prediction from the National Oceanic and Atmospheric Administration's Office, when relative changes in hurricane activity are taken into account, average yearly losses are expected to rise by $7.3 billion, bringing the total cost of hurricanes and other coastal disasters to $35 billion per year. (NOAA) Real estate is particularly vulnerable to the effects of climate-related disasters since it is an illiquid, long-term investment. Therefore, it is important and constructive to begin keeping an eye on the effects of climate change on clients' existing and target real estate assets in order to assess the risks. Despite the present availability of tools and resources to assist clients in assessing their exposure to these risks, many still struggle to factor the risk into their investment decisions. On one hand, it is difficult to quantify climate activities. On the other hand, people cannot rely on historical climate and weather data to predict future events, because we know that these patterns are changing. To help our clients to make more informed decisions when investing in the real estate sectors, our project aims to build a clear and actionable climate risk assessment framework that outlines the potential hurricanes risks and the real estate market. Before constructing the predictive model, we plan to use the historical data to find the relationship between climate activities and the real estate market.

Research Questions:

In this study, our major research question is "What is the relationship between hurricane intensity and the financial loss of impacted companies?" Below are three points that explain where was our data collected and how we built the data of hurricane and real estate assets. The first one is that hurricane intensity is measured by maximum wind speed per year and the number of hurricane encounters. The second one is that impacted companies are defined to be any companies within a 50-mile radius of hurricane landfall. The third one is that financial loss is defined to be asset write-downs.

Methodology:

We have four steps of the methodology for our analytical research. First, we explored and understood the structure of the data including what financial data is available over which years, how hurricane data is presented, features to join the financial data with hurricane data, and specific attention to geolocation (longitude and latitude of hurricane or property). For each storm, we have extracted its latitudes, longitudes, wind speeds, category, and year. We have then enriched the dataset in several ways. We've calculated the average latitude and longitude and extracted the maximum wind speed for each storm. Then, using the latitude and longitude coordinates, we've determined whether or not the storm fell within the US. We only want to consider storms within the US because we will be considering macroeconomic data that will be different across different countries - and we've specifically pulled the data for the US. Secondly, we have determined the scope of our research. We choose to use the existing data from 1995-2020. The data is divided into two sections: 1. The data set of the real estate listed the public trade company. 2. The data set of the hurricane in the United States. Third, we have different ways to analyze each research problem. Our goal is to know the financial losses caused by the

hurricane in the real estate industry. We will use macroeconomic data as control variables in our model, especially the GDP and the inflation rate -- to integrate external and non-hurricane market factors into our model. For each company, we pulled their institution name, the properties they own, their net book value per year, and the state where the property is located. This data can help us understand the change of the write-down value of the company's assets that have been impacted the most by hurricanes over time. We will establish a model of the key variables to see how the companies' financials are changing in correlation with hurricane hits (and the strength of those hits). Finally, we will evaluate and improve the model iteratively based on model prediction ability and interpretability of the model.

Preliminary Findings:

a. Data Preparation:

First, for hurricane data, there are two available datasets having detailed records of each hurricane for different time periods. The first dataset is called IBTrACS, which is downloaded from the NOAA website. The time period for this data set is from 1842-2021, with an indication of whether a hurricane has made landfall. However, the surface wind speed data is not very reliable due to the measuring method. Thus, our team decided to use the other dataset called ADT-HURSAT for a more accurate record of hurricane surface wind speed. The reason why this dataset has a more accurate record is because of the improved measuring method, which applied the advanced Dvorak Technique to a globally homogenized record of geostationary satellite imagery. We've included a visual comparison between the two datasets in the appendix.

For the asset data, we used the S&P Capital IQ database. Specifically, we relied on two types of asset data. First, we looked at company-level data for publicly-traded REITs in the United States which firms are required to report on a quarterly and annual basis. Relevant variables used include total assets and any asset impairment charges related to the company for that period. For periodicity, we used 1995-2020 which is the limit of the S&P data. Second, we used property level data also found within a different search tool within Capital IQ. This tool shows the address and parent company for each asset. Assets include commercial and residential properties within the United States. Similar to the company level data this was available on a quarterly and annual basis. Fields used include total assets and asset writedown. The database yielded over 40,000 properties across the continental United States.

Then, we merged these two datasets by storm id. Thus, the final dataset contains wind speed data from ADT-HURSAT until 2017, and wind speed data from IBTrACS from 2017 to 2020. It also includes the indicator of whether the hurricanes have made landfall.

b. Data Exploration

To combine hurricane data and real estate data, we filter to include only the hurricanes that made landfall in the US. Then, we use the latitude and longitude to find out which real estate is within the affected area, which we defined as a 50 miles radius around the hurricane.

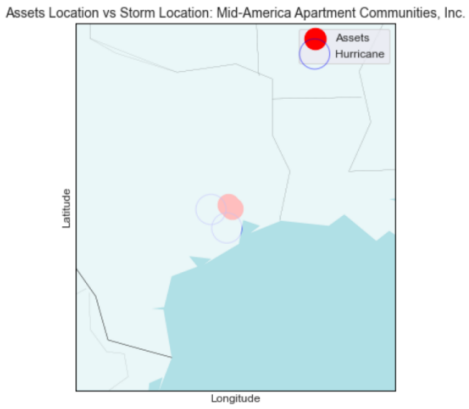


Figure 1: Example of how to define the affected real estate

In this way, we plotted all the hurricane landfalls from 1995-2020 represented by the blue circle, and the location of each property represented by red dots. The most affected area is the southeast coast of the US.

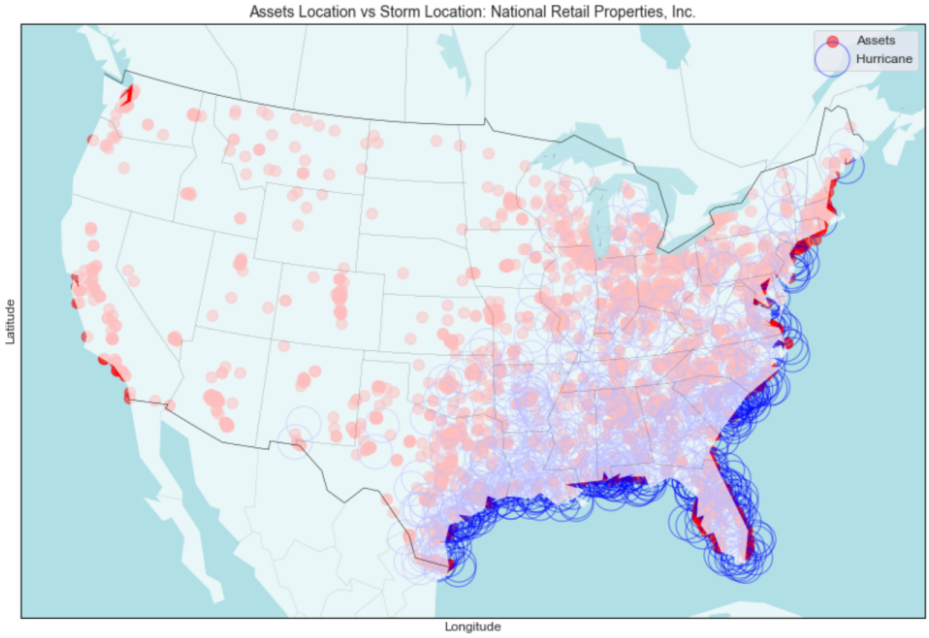


Figure 2: Asset location VS Storm location: National Retail Properties, Inc.

To have a very brief view of the affected companies, we took a look at the top 30 affected companies from 1995-2020.

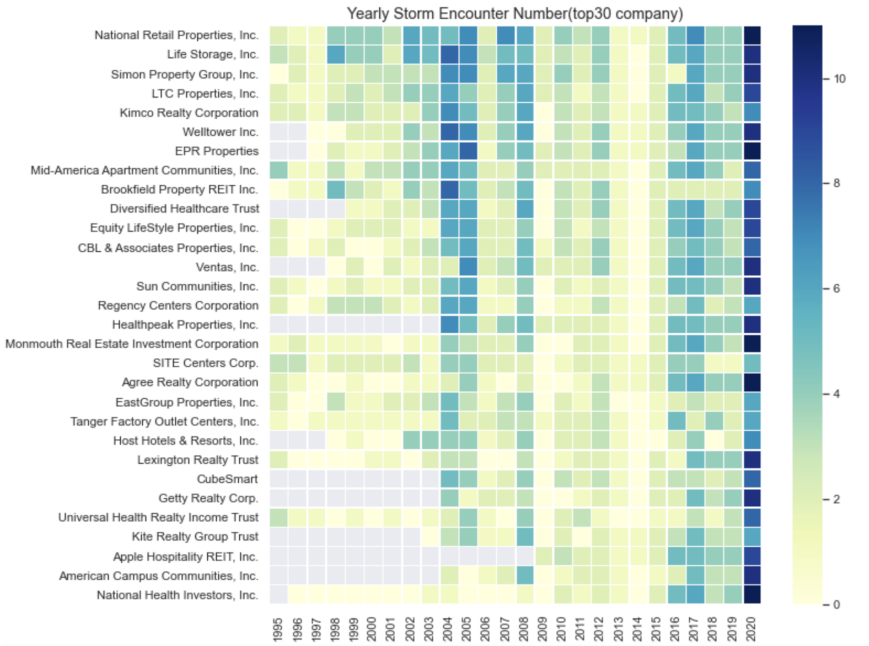


Figure 3: Top 30 Affected Companies

Since the loss of affected companies is measured by the write-down value, we select one company to show the trend of the write-down values and the number of hurricanes encountered over years. Below is an example of the company Macerich. The trends of the two lines become closer after 2006, which means that with more hurricanes encountered, there are higher writedown values.

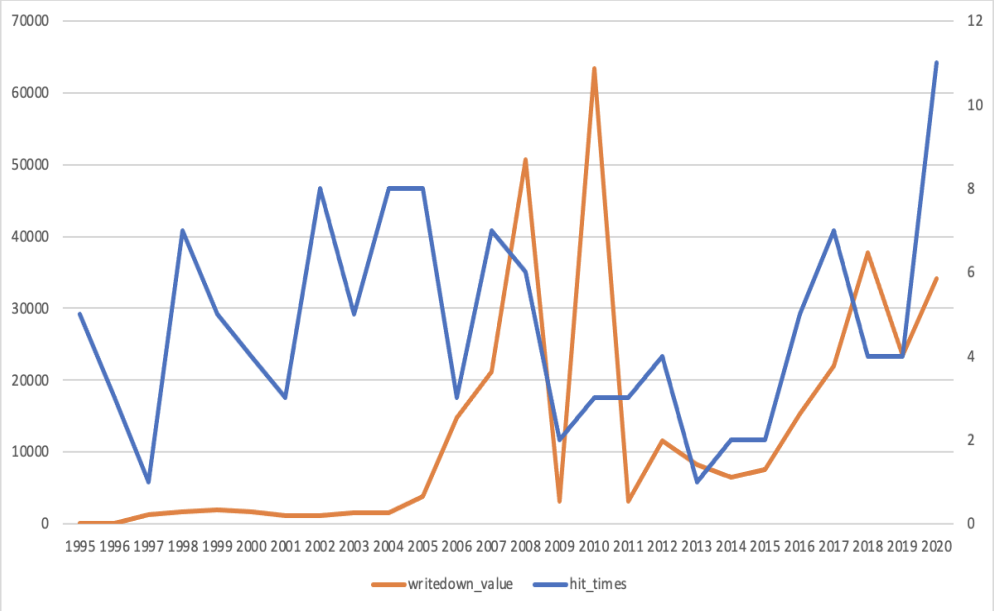


Figure 4: Total Asset Write-down Value & Total Hurricane Hit times by Year: Macerich Company

Modeling Results:

First, we finished a data exploration and found that our independent variables and dependent variables were likely to have a positive relationship. Then we tried linear regression to predict the values of the dependent variables based on the independent variables. The p-value of max_storm_speed is lower than 0.05, which means it is significant. And the coefficient of it is 0.0086, which is positive. So it means that the larger the max storm speed is, the higher the write-down value of the companies will be. However, the R square was 0.027, which represents that the model was not good at predicting the results. It may not be accurate enough. The R square is too low, but it can still prove that it has a certain relationship with max_storm_speed. The accuracy of linear regression is not very good, so we kept trying random forest instead of linear regression. The MAPE (Mean Absolute Percentage Error), which can somehow show the

accuracy of the regression model was 41%. It was still not high enough but at least much higher than the R-square of linear regression.

Then, due to the low accuracy of the regression models we tried, we decided to use classification modeling. The classification model used write_down, which was whether the company recorded write-down value, as a final dependent variable. We used all other supported data to predict if the company has a write-down value or not.

Since it is a classification model, we plan to use logistic regression, and the accuracy of the logistic regression model is around 57%. Then we tried the random forest classifier, which has a 60% accuracy. There is still a slight difference between the logistic regression model and random forest classifier because the precision (the fraction of relevant instances among the retrieved instances) and recall (the fraction of relevant instances that were retrieved) of them are different. We could choose different models for different purposes.

Recommendation:

Our project examined the impact of hurricanes' intensity on the financial value REITs properties in the US. After selecting the properties in the hurricane area from our database, we find that the intensity of hurricanes appears to have a significant impact on asset write-down values.

The preliminary findings of the datasets and the results of this study are important for REIT investors deciding whether to allocate or construct additional capital, especially if the risk of additional hurricanes in an area is increasing due to climate change.

In the future, once the REIT investors have climate prediction data, like the number of storms and the maximum wind speed, they can have some preliminary concepts about the write-down value of their assets by using our model. Moreover, if the investors are considering finding a new investment opportunity, S&P may be able to advise them on which locations are not being affected by hurricanes very frequently. For instance, from our database, most properties being impacted by hurricanes are located on the east coast.

Moreover, we believe our findings may not only help REIT investors but also insurance companies. With the model we provide, the insurance company may calculate the expected damage brought by the hurricanes to each property with the latest weather forecast information. As a result, they can change the premium rate according to the likelihood that a property may be impacted by the hurricanes. Additionally, they can target home insurance marketing on people who have properties in the area which is affected by the hurricane the most. If S&P is interested in this topic, the team can continue to retrieve the data about the change rate of premium of properties we examined in our project to see if the hurricanes affect the premium rate of those properties.

In a future follow-up study, we would recommend the team to have more structural and amenity information about the real estate properties for the consideration of the impact of hurricane intensity. For the structural variables, the team may collect the data like the number of bedrooms and bathrooms, the floor of the properties. Therefore, we may know if the size and the number of floors would affect the write-down value when the hurricane hits. For the amenity and location variables, the team may collect the data like the distance to the beach and the type of view of the property. For example, we may expect that an apartment with an ocean view may be

affected more than one with a city view. An updated model considering these more specific structural and amenity variables can help investors have a more thorough understanding of the potential financial loss of different types of REITs investment targets.

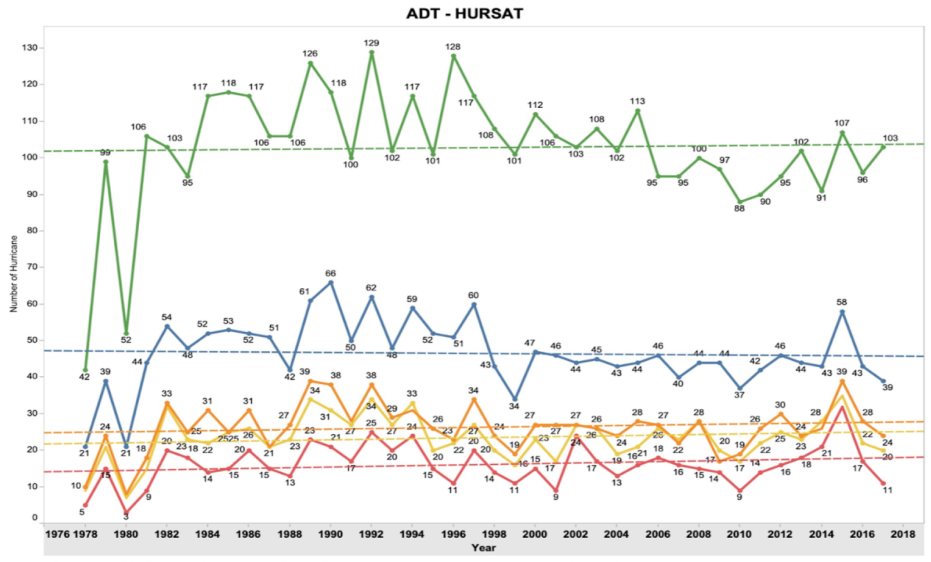
Furthermore, in the time scope from 1995-2020, there were several financial crises causing serious economic recessions in the US. Even though we included the inflation rate as one of our variables, we did not more deeply consider the impact of financial crises on the real estate market, especially the subprime mortgage crisis in 2008. We would advise a team with a more advanced financial background to help improve the present model and add more financial market variables.

Finally, natural disasters like hurricanes and wildfires cannot be eliminated in different regions, where many real-estate properties have already been constructed. But, a more accurate understanding of the impact of natural disasters and the side effects of risk exposures on housing markets can help improve clients' readiness to handle these issues. Future research on insurance policies for natural disasters and more detailed climate change adaptation methods are also worthwhile ways to build on the results of this project.

Appendix:

Comparison between ADT-HURSAT and IBTrACTS

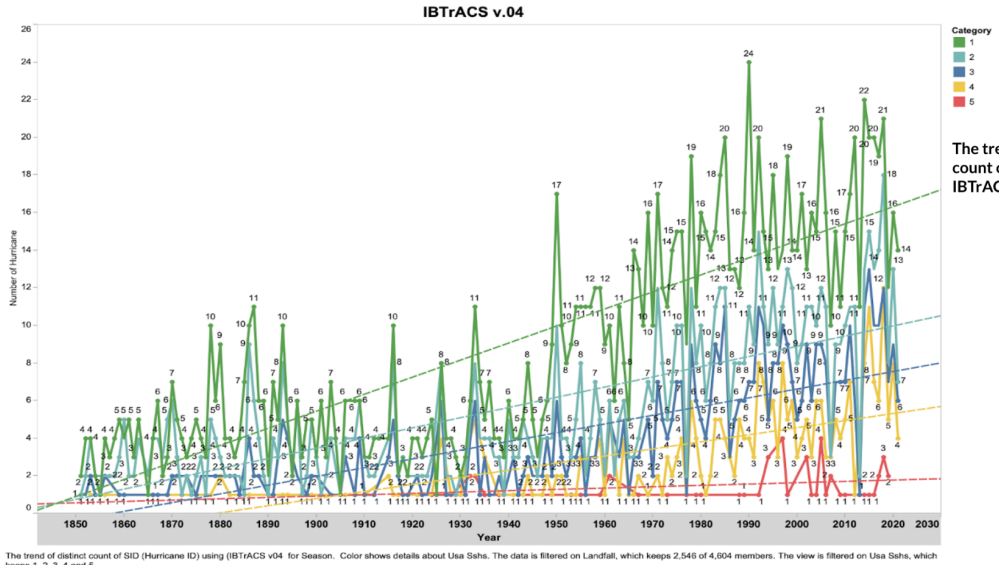
ADT-HURSAT vs IBTrACS



The trend based on distinct count of Hurricane ID using ADT-HURSAT data 1978-2017

The trend of distinct count of SID (Hurricane ID) using ADT- HURSAT data for Year. Color shows details about Category. The view is filtered on Category, which keeps 1, 2, 3, 4 and 5.

ADT-HURSAT vs IBTrACS



The trend based on distinct count of Hurricane ID using IBTrACS v04 data of 1842-2021

The trend of distinct count of SID (Hurricane ID) using (IBTrACS v04 for Season. Color shows details about Usa Sshs. The data is filtered on Landfall, which keeps 2,546 of 4,604 members. The view is filtered on Usa Sshs, which keeps 1, 2, 3, 4 and 5.