Revealing How Real Estate and Insurance Markets React to Natural Disasters

Augustus Cantella, Emma Sainovic, Jessica Hope, Haedam Cho

Digital & Data Studies, Harpur College SUNY Binghamton

DIDA 425: Digital Inquiry & Practice

Instructor Jacopo Mazzoni

May 9, 2025

Abstract

Studies suggest that climate change is worsening as the years progress and is causing natural disasters to occur more frequently in the United States, and thus causing economic impact. This paper will study the economic ramifications of natural disasters in the USA, specifically in terms of the impacts they have on insured damages and the housing market. We will utilize two main datasets, one which we obtained from the EM-DAT website, and the other is the Zillow Home Value Index (ZHVI). We will have a focus on the way storms, mainly in the form of a hurricane, will affect insurance rates in the US, the way a campfire in California affects Butte County's houses ZHVI, and the way droughts in Florida impact home ZHVI in the state. This may be useful in order to predict future housing values and insurance-related changes post-natural disaster. According to our initial statistics from the EM-DAT dataset, storms, droughts, and campfires, were the natural disasters that had the most economic and insurance-related impacts. We found that only storms, mainly in hurricane-form, and campfires, can help predict the real estate market and insurance trends in the USA.

Introduction

Natural disasters can occur anywhere at any time, and their unpredictability is what makes them so risky and oftentimes destructive. In this research, we will focus on natural disasters that occur in North America between 2000 and 2024, and use this data to analyze insurance and home market trends. The risk of economic loss that comes with natural disasters is what motivates people to enroll in insurance plans that would allow them to recover financially in the case of economic loss due to damages caused by a natural disaster. This entails the customer or client paying a monthly premium to an insurance company, otherwise known as a rate or subscription, which is set depending on factors that will be discussed later on. Not every country offers insurance, although the United States has a plethora of insurance plans available which someone could choose based on their needs, from property insurance, crop insurance, or many other types. We want to analyze how the economic loss that follows a natural disaster affects the values in the home market, as well as the way insurance companies react to the disasters.

We are doing this research because natural disasters are starting to affect real estate markets and insurance trends across the United States. As climate change continues to increase the frequency and intensity of events like storms, floods and wildfires, these disasters are beginning to reshape where people live, home values, and how accessible property insurance is. This topic is important not just for homeowners and investors, but also for local governments, planners and policymakers who are trying to make communities more resilient and financially stable in the face of ongoing environmental risks.

To reveal the effects of natural disasters on real estate and insurance markets, this paper studies different disaster types using a variety of techniques such as creating and interpreting data visualizations and maps to understand real estate trends in relation to other variables and geographies. By taking a look at total insured damages and total damages of storms in the US we found that the strength of a storm based on the Saffir-Simpson scale is what impacts adjusted total damage and adjusted total insured damage. We also examine the statistical significance of specific disaster case studies such as Camp Fire by conducting difference-in-difference analyses and find that the disaster event did significantly impact Butte County's housing values negatively long-term, but had a weaker immediate impact. Finally, we attempt to predict housing values based on drought statistics using linear and logistic regressions, which both performed poorly, suggesting that there is a weak relationship between our drought data and housing values or that better predictors are required.

Throughout this paper we will review prior literature and reflect on what they mean for our research. Then we will explain in detail our methods, where our datasets come from and how we interacted with them in order to get our results. We then explain our results and give some insight into what could be improved upon. We will finally tie everything together with our research question and all previous research in our conclusion

Literature Review

Climate change is changing how we think about housing risk and long-term property values, and insurance is one of the main tools people use to deal with the financial side of that risk. The article *Insurance as an Alternative for Sustainable Economic Recovery after Natural Disasters* (2022) pulls together over 260 sources and argues that disaster risk insurance is one of the best ways to recover financially after a natural disaster. It breaks down how things like climate change, location, and whether the insurance is public or private affect premium pricing. One part that stood out was how tornadoes, even though they're common and destructive in the U.S., still don't have their own standalone insurance policies so people have to rely on general property insurance instead. This sets the stage for understanding how insurance access and structure affect recovery and stability after disasters.

Building on that, Zac Taylor's article *The Real Estate Risk Fix* looks at how Insurance-Linked Securities (ILS) take this idea one step further, by providing insurance for the insurers themselves. His case study of Florida shows how investors can absorb financial risks from hurricanes through ILS, which can help stabilize housing markets. But he also points out the risks of over-relying on financial tools instead of addressing structural problems like zoning or overdevelopment. This complements the first article by showing what happens when insurance becomes financialized, and what that means for long-term resilience.

Michele Schroeder's article *Environmental Insurance* shifts the focus to how individual transactions are affected. Instead of broad markets or insurance companies, she looks at how buyers, sellers, and developers can use environmental insurance to transfer risk. Her comparison of traditional tools versus insurance shows how coverage (or lack of it) changes not just liability, but how properties are valued. This builds on the idea from the first two articles that access to effective insurance is deeply connected to how markets function after a disaster.

Taking a broader view, Contat et al. (2024) *When Climate Meets Real Estate* surveys over 140 papers and shows how all of these issues like physical damage, transition risk, policy, insurance, and buyer perception fit into the national conversation on climate and housing. They show that markets haven't fully priced in long-term risks like sea level rise or drought, and that responses vary depending on local beliefs, policies, and data quality. Their work connects the previous articles by showing how all these individual responses and tools interact on a larger scale.

That complexity is also reflected in another paper by Contat, Doerner, Renner, and Rogers (2024), who study Hurricane Ian's impact on Florida housing prices. Surprisingly, they found prices went up 5–11% right after the storm. This challenges the assumption that disasters always lead to a downturn and shows how outcomes can vary based on local context. They also talk about the limitations of public data, which links back to the spatial and measurement issues raised by Contat et al. (2024). Both of these papers help explain why disaster responses aren't one size fits all and support the need for mixed method approaches like ours.

Andrea Martino's thesis *Climate Risk and Real Estate* digs deeper into how different disaster types affect property values, using models like hedonic pricing and difference in differences. She emphasizes how spatial data and risk visibility shape buyer behavior, which reinforces the themes we saw in Schroeder and Contat et al. that perception and information gaps are just as important as the physical event itself when it comes to price shifts.

Barthel and Neumayer (2012) step back and ask whether the rising losses we see in insurance data are due to worsening disasters or just changes in what's insured. By normalizing data over time, they find that in places like the U.S., storm and flood losses are going up. Their work adds historical context to the more recent papers and gives us a way to think more critically about trends in insurance and damage data.

Finally, Leeper et al. (2022) look at drought specifically, mapping how drought patterns differ between the West and East over the past 20 years. Their findings help explain some of the patterns we noticed in our own data, like why the Southeast, especially Mississippi, has been hit hard by D4 droughts recently. It also expands the scope beyond fast disasters like hurricanes to slower-building risks like droughts, which is something Contat et al. (2024) also call for in their broader survey.

Methods

In order to properly assess how natural disasters and climate related-risks affect housing valuations, this paper will feature Zillow Home Value Index (ZHVI) data which focuses on the middle-third of homes, reflecting the typical home value for a given area, rather than the median. To calculate the ZHVI ($), the index uses a Zestimate model which uses a combination of public, multiple listing service, and user-submitted data to estimate a home's market value. In addition to monthly ZHVI data from the year 2000 to the present, it consists of various geographic ID data such as SizeRank, State, and Metro. Since these datasets are official public releases from Zillow, the files come fully cleaned and organized. Zillow also allows for different geographic level and housing type versions of ZHVI, so for a more granular approach we chose to work with a county-level version of the dataset. It is important to note that the original dataset only contains 3,073 counties, 71 short of the US's 3,144 total counties. Our merging and analysis processes maintain the 3,073 counties, which will be apparent in mapping visualizations. The ZHVI dataset was merged with a master dataset from note that contains the all US states and standardized county names with fips codes.

EM-DAT is the International Disaster Database which compiles data from various sources into one about worldwide disasters since the year 1900. We downloaded data about natural disasters in Northern America between 2000 and 2025. The dataset includes a column with a unique identification number to identify natural disasters and map the disaster impacts at the country level. The impacts of a single disaster can be different between countries, so EM-DAT organizes their data by country, and could have different information on impacts from the same disaster based on that country. Some of the human impact variables used on EM-DAT include total deaths, number of people injured, number of people affected, number of homeless, and total affected people, which is the combination of the last three columns. They also note that there is some uncertainty with these numbers because it is all based on what is reported and they may have incorrect data because of things that are not or are falsely reported. As for economic impact variables, EM-DAT includes columns with information about insured damage, adjusted insured damage, total damage, and adjusted total damage. Our dataset has values given as $USD written in thousands of dollars (‘000 US$). They also note on the site that economic loss after natural disasters is typically largely underreported and they usually only have this information available for countries that offer insurance, and only after high-impact disasters. This caused some limitations in the data because there were many rows that had empty cells in the economic variables which made it difficult to analyze properly. Finally, we included a statistics table in the results section in order to show percentage increases in number of storms, adjusted storm damages, and adjusted insurance costs between the years.

The U.S Drought Monitor provides a very comprehensive dataset of drought levels throughout the continental United States from 2000 to 2025. The data can be downloaded differently by area type and statistics category. Area type can be selected to be national, state, county, etc. The statistics category can be selected to be percent area, total area, percent population, and total population. We decided to use the county level data and use percent population statistics. The columns consist of MapDate, FIPS, County, State, None, D0, D1, D2, D3, D4, ValidStart, ValidEnd, StatisticFormatID. FIPS, County, State are the geographical locations of drought locations. None, D0, D1, D2, D3, D4 are the level of drought that location is experiencing, U.S Drought Monitor describes None as normal or wet conditions, D0 as abnormally dry, D1 as moderate drought, D2 as severe drought, D3 as extreme drought and D4 as exceptional drought. These have a percent value of the total population of the area affected by a given level of drought. They determine these drought levels by usings observations “To determine drought intensity, USDM authors use a convergence of evidence approach, blending objective physical indicators with insight from local experts, condition observations and reports of drought impacts." (U.S Drought Monitor). ValidStart, ValidEnd is the time the data was observed, USDM observes data every week on Tuesday which is the ValidStart, the ValidEnd is the following Monday. We decided to do 2022 to the start of 2025.

Using an R-script provided to us by our instructor Professor Mazzoni, we were able to merge our ZHVI dataset with a master fips dataset (provided by Professor Mazzoni) to match standardized US county names with the fips codes in our Zillow data. This left us with 21 unmatched counties, which we manually entered into the dataset, allowing us to then merge it with our original Zillow dataset so that each county is matched with the monthly ZHVI values. To get an idea of what housing valuations looks like over time, we created a time series with years ranging from 2020 to 2024.

We manipulated the Zillow ZHVI dataset using python pandas pct change to transform the columns into percent change from the previous month. We then combined the Zillow ZHVI data set with the Drought Monitor set, joining them on the county column using pandas. We wanted to test out a couple different response variables for when we did machine learning so we created another column where the percent change in ZHVI was not the month that the drought was observed but the month after, this we called Month After Per Change. Another predictor we wanted to test was the percent change in ZHVI in a given county versus the state average ZHVI percent change. For this response we created two columns, one for the month the drought was observed and one for the month after the drought was observed. These are labeled State Dif and Month After State Dif. For this dataframe with these created columns we created another dataframe where each of these columns were transformed into a binary classification, where if the percent value was negative it became a 0 and if the value was positive it became a 1.

To tackle our research question, we decided to categorize natural disasters into types: wildfire, drought, and storm. By analyzing case studies of each of these types, we aim to better understand the effect natural disasters have on various real estate markets and insurance trends. For wildfires, we found that the most economically devastating incident was 2018's Camp Fire in Butte County, California. Understanding that we have time series data and potential control groups in Butte's neighboring counties, we decided to go with a difference-in-difference analysis with a model that compares ZHVI trends in Butte County to those in neighboring counties before and after Camp Fire with the following model equation:

$$ZHVI_{it} = \beta_0 + \beta_1 \cdot PostFire_t + \beta_2 \cdot Treatment_i + \beta_3 \cdot (PostFire_t \times Treatment_i) + \varepsilon_{it}$$

The PostFire variable controls for changes in the housing market that occur after Camp Fire across all counties in the model, including things like economic shifts and external shocks. Treatment variable controls for any baseline differences between Butte County and its neighbors in average ZHVI. Finally, PostFire x Treatment is the key interaction term that estimates the additional change in Butte County's ZHVI relative to its neighbors post Camp Fire. In other words, this key variable captures the local impact Camp Fire had on Butte County. The subscript *i* represents county and *t* for time. After conducting this analysis, we realize that the COVID-19 housing value spike could be a confounder in the model. So, we also ran a refined diff-in-diff analysis with a shorter post-treatment window to better isolate the fire's impact on Butte County. The model equation stays identical in this refined version. To ensure data quality and suitability, proper robustness checks are conducted for the diff-in-diff model as well.

We cleaned the Em-DAT dataset by deleting columns we found unnecessary to our research, as well as the rows that held information about natural disasters occuring in 2025 because there is not enough data recorded yet in 2025 which may skew our analysis results. Much of the analysis on storms was done so by using Python's pandas and matplotlib libraries in order to make charts and data visualizations to find patterns or trends in the data. We started off by analyzing how many times the different types of storms occur in the entire dataset followed by the number of storms occurring per year, and we made a line graph for the latter entitled "Years with the Min & Max Number of Storms". The dataset includes data for each month of every year, although we used the "Year" column as the independent variable for the purpose of finding a pattern as the years progress, and because hurricane "season" is an annual time period. Later on we will compare this line graph to line graphs containing information about insurance rates so that we can see how the number of storms affected the amount of adjusted insured damage. We created another line graph that has the year as the independent variable, with the adjusted insured damage value as the independent variable, entitled "Outliers of Storm Damages Covered by Insurance". We used the adjusted values rather than the unadjusted value because the unadjusted value is the general value of what insurance should cover, while the adjusted values are what the insurance actually covered, which is more accurate, thus providing better data for analysis. Finally, we made a line graph displaying the adjusted total damage per year, with the year as the independent variable and the adjusted total damage as the dependent variable, entitled "Outliers of Total Damages Caused by Storms".

Using Tableau we were able to create key visualizations to gain insights into our data's distribution across america and how that changes over time. Two important maps we wanted to make were distribution of drought levels across the United States and a timelapse of ZHVI values over the years. We also did machine learning methods to try to predict the different ZHVI percent change given the drought level information. These methods include linear regression and logistics regression. For linear regression we did simple linear regression and multivariate linear regression. For the simple linear regression we used just D4 as the predictor variable. For multivariate we used all drought levels besides for None as the predictor variables. When doing the logistic regression we used the same variables against the binary classification columns.

Results

Using the Em-DAT dataset, we identified the three most impactful disaster types by examining some descriptive statistics in the following table:

	Insured Damage ('000 US$)		Insured Damage, Adjusted ('000 US$)		Total Damage ('000 US$)		Total Damage, Adjusted ('000 US$)	
	mean	sum	mean	sum	mean	sum	mean	sum
Disaster Type								
Drought	6,300,000.00	18,900,000.00	8,316,855.67	24,950,567.00	7,164,285.71	100,300,000.00	8,734,256.69	113,545,337.00
Earthquake	136,250.00	545,000.00	209,492.00	837,968.00	398,055.56	3,582,500.00	614,512.33	5,530,611.00
Epidemic	NaN	0.00	NaN	0.00	NaN	0.00	NaN	0.00
Extreme temperature	871,666.67	2,615,000.00	1,195,911.33	3,587,734.00	758,333.33	4,550,000.00	975,538.00	5,853,228.00
Flood	506,870.97	15,713,000.00	666,235.13	20,653,289.00	950,856.02	83,675,330.00	1,190,649.50	102,395,857.00
Mass movement (wet)	NaN	0.00	NaN	0.00	460,000.00	920,000.00	558,917.00	1,117,834.00
Storm	2,733,955.15	546,791,030.00	3,328,719.88	652,429,097.00	3,769,941.81	1,229,001,030.00	4,538,184.09	1,402,298,885.00
Volcanic activity	NaN	0.00	NaN	0.00	500,000.00	500,000.00	606,718.00	606,718.00
Wildfire	2,163,420.00	54,085,500.00	2,851,719.26	65,589,543.00	1,503,449.38	91,710,412.00	1,953,532.54	111,351,355.00

Figure 1

Here, we can see that Drought, Storm, and Wildfire are economically (measured by damages) and insurance-wise (measured by insured damages) the most impactful disaster types, giving us reason to study these in particular. When further exploring the data, we saw that the most economically impactful wildfire was Butte County, California's Camp Fire in 2018, with an adjusted total damage of $20,021,680. Accordingly, we raise the question: did Camp Fire significantly affect Butte County's real estate market? And if so, how? To evaluate this, we compare Butte County's ZHVI over the years to that of its neighboring counties by log-transforming the data and plotting it as a time series. For our regression analyses, the ZHVI is log-transformed, but for better visualization purposes, we kept the ZHVI as a dollar amount. The real estate effects of the 2008 Financial Crisis and COVID-19 housing bubble are clearly shown in the graph, and taking note of the presence of such events will be crucial.

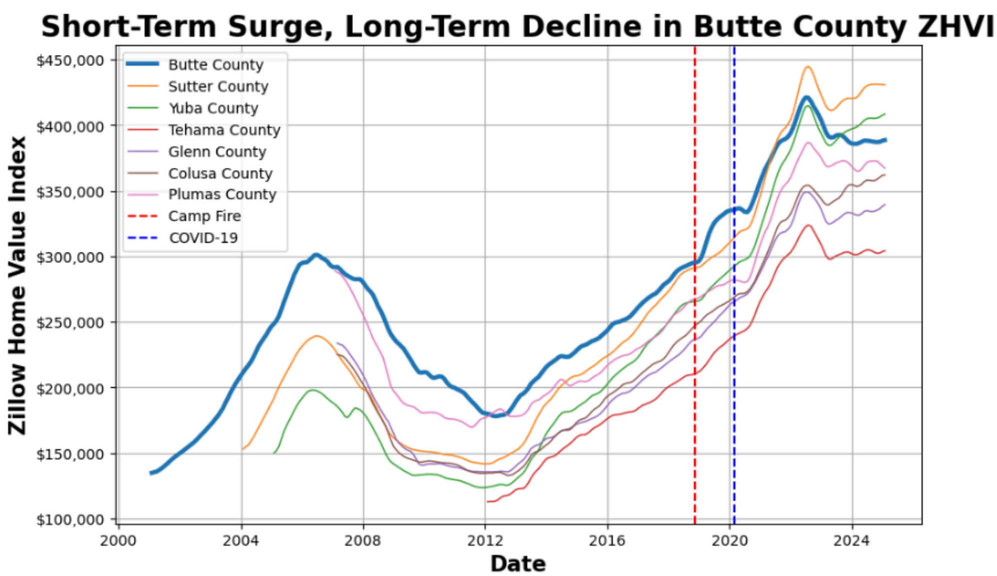

Figure 2

Visually, there is a clear uptick in ZHVI post-fire in Butte County (in blue) In order to examine Camp Fire's statistical significance on Butte County, we run a difference-in-difference analysis with the neighboring counties as control comparisons. Before proceeding, however, we make sure to validate this methodology by checking that our data exhibits the parallel trends assumption, which means that in the absence of the treatment (Camp Fire), the trends in both the treatment and control groups are generally the same. We do this by running an OLS regression with ZHVI as the dependent variable, and Trend_Interaction as the key interaction term that measures the difference in trend slopes between Butte and neighboring counties. The results are found in the following Figure 3.

Variables	OLS Regression Results
Intercept	< 0.001
TimeIndex	< 0.001
Treatment	< 0.001
Trend_Interaction	0.922
N	1064
R^2	0.231

Figure 3

Since *Trend_Interaction*'s p-value demonstrates statistical insignificance, it can be said that the parallel trends assumption is indeed observed in the data. This gives us reason to move onto the diff-in-diff analysis using the regression equation in Figure 3. The outcome of this test tells us whether Camp Fire had a significant effect on Butte County when compared to similar (neighboring) groups that did not receive the treatment by focusing on data pre- and post-fire. We also add time-fixed events by adding dummy variables for each time period (monthly) to account for potential variations in ZHVI caused by time-related effects. With the following regression results, we see that with a negative t-statistic, the interaction term is highly significant, meaning Camp Fire's effect on Butte County is significant compared to its neighboring counties with an overall negative effect on ZHVI.

Variables	t-statistic	p-value	other
Intercept	106.480	< 0.001	-
Post_Fire	10.597	< 0.001	
Treatment	30.041	< 0.001	
Interaction	-8.623	< 0.001	
N	-	-	1589

R^2			0.917

Figure 4

However, after reassessing the model, we realized that the COVID-19 housing bubble could be a confounder, as it is included within the full-period time window (November 2018 onward) To address this issue, we shortened the post-treatment time window to stop before March 2020, when COVID-related real estate effects began to take place. This refined diff-in-diff analysis resulted in a lesser, mildly significant result with a weaker negative impact of Camp Fire on Butte County as shown in the results in Figure 5.

Variables	**t-statistic**	**p-value**	**other**
Intercept	105.215	< 0.001	-
Post_Fire	8.586	< 0.001	
Treatment	29.684	< 0.001	
Interaction	-2.388	< 0.017	
N	-	-	1176
R^2			0.842

Figure 5

After examining both results, we can see that the full-period DiD model shows a statistically significant interaction effect, suggesting that Butte County's ZHVI did negatively diverge from its neighboring counties after Camp Fire. However, since the full-period time window overlaps with the pandemic's housing disruptions, it is difficult to fully attribute the observed effect to the fire alone. The refined model, which isolates the immediate post-fire period before COVID, finds mild to no statistically significant short-term impact. Although Camp Fire resulted in an *immediate* surge in home values due to a decreased supply and increased demand from displaced residents, the overall short term effect was a modest decline compared to its neighboring counties. However, looking at the years after COVID, the gap widened significantly, meaning that Butte County's housing market didn't recover like its neighbors did. This means that destructive natural disasters such as Camp Fire can have lasting effects on home values of the affected region, especially when compounded by broader macroeconomic events such as COVID. Possible driving forces of these lasting effects could include permanent housing loss, infrastructure disruption, and changes in risk perception. Hence, policy and planning should account for both immediate and long-term effects that address the sustained aftermaths of such disasters.

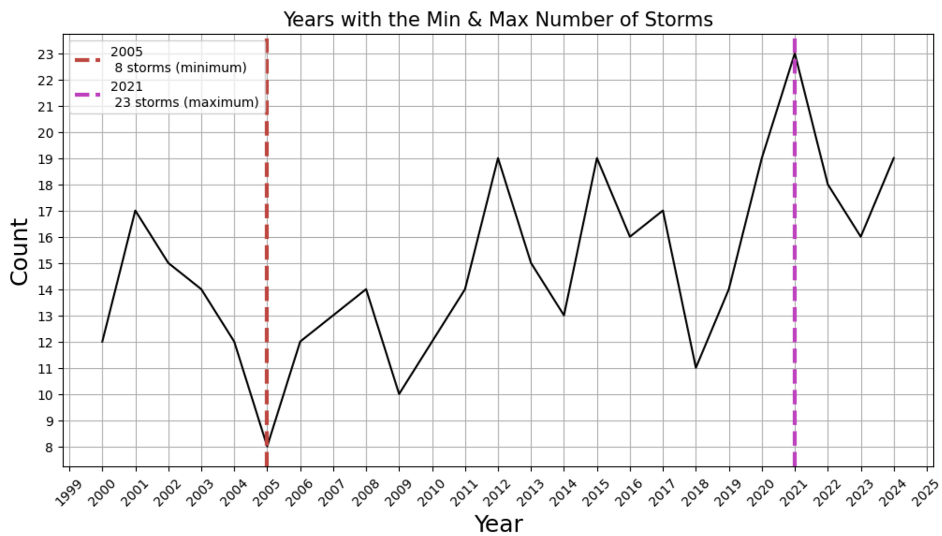

Figure 6

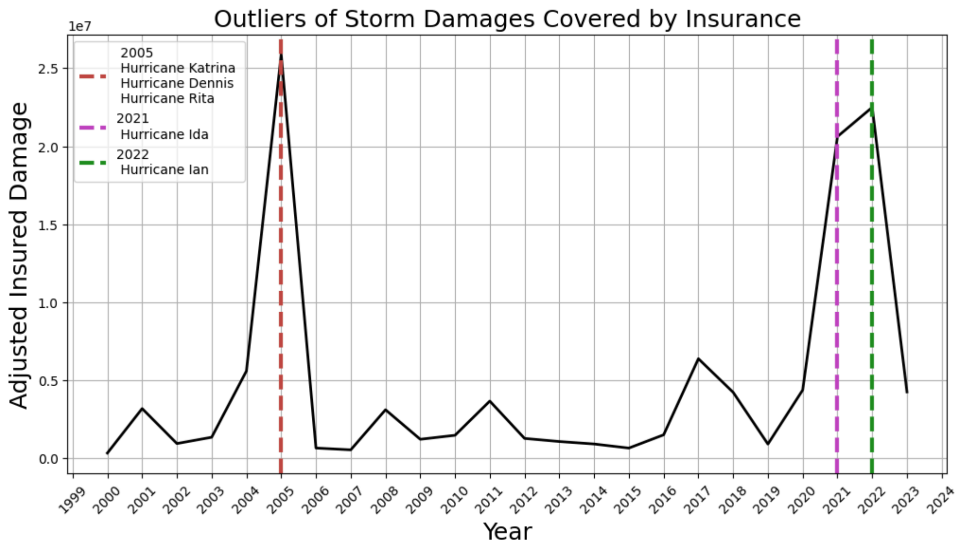

Figure 7

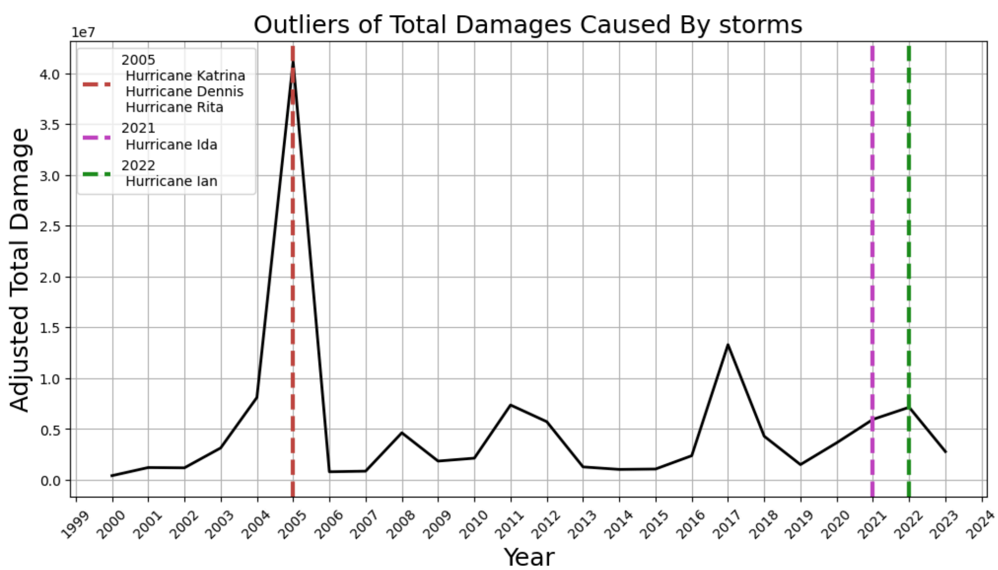

Figure 8

The literature review about sustainability tells us that the number of natural disasters occuring every year is increasing due to climate change, and Barthel and Neumayer article tells us that there is a trend of increasing insurance rates, especially for storms. Because of this, we found it important to compare these graphs. We want to focus on the years 2005, 2021, and 2022, as they are outliers in figures 7 and 8. We will refer to "Adjusted Insured Damage" from figure 7 as Adj.ID and "Adjusted Total Damage" from figure 8 as ATD.

Year	Storms #	% Increase	$ Adj.ID/year	% Increase	$ ATD/year	% Increase
2004	12	-	5,594,578.00	-	8,098,834.00	-
2005	8	-33.33	25,889,050.00	362.75	41,146,110.00	408.05
2006	12	50.00	654,722.70	-97.47	794,007.30	-98.07
2020	19	58.33	4,372,876.00	-	3,645,022.00	-
2021	23	21.05	20,615,580.00	371.44	5,944,133.00	63.08
2022	18	-21.74	22,489,140.00	9.09	7,136,601.00	20.06
2023	16	-11.11	4,250,000.00	-81.10	2,779,688.00	-61.05

Figure 9: Statistics Table

According to "*Science and the Storms: the USGS Response to the Hurricanes of 2005*", storms are rated from 1 to 5 on their intensity based on the Saffir-Simpson scale. Several hurricanes in 2005 reached a 4 or 5 on the scale including Hurricane Katrina in New Orleans, Hurricane Dennis in Florida, and Hurricane Rita by the Gulf of Mexico. These were some of the most destructive hurricanes of 2005 and would explain why the ATD and the Adj.ID per year increased so much, without the number of storms increasing at a proportional amount. If you look at the statistics table above, you can see that from 2004 to 2005, the number of storms actually decreased by approximately 33% (4 less), yet the ATD increased by over 408% ($41 million more), and the Adj.ID increased by over 362% ($26 million more).

According to "*Active 2021 Atlantic hurricane season officially ends*", 2021 is in the top three years most active storm years, specifically hurricanes. Scientists say it is because of the warm phase of Atlantic Multidecadal Oscillation (AMO) that occurred in 1995. The AMO caused the storms after 1995 to become stronger and last longer than before. The article also mentioned that May and August of 2021 had above-average levels of destruction from hurricanes, reaching levels 4-5, which were before the official start of storm season that usually occurs June first and on. The most destructive hurricane of 2021 was Hurricane Ida which was a category 4 hurricane during August in Louisiana. The statistics chart shows that from 2020 to 2021, the number of storms increased by over 21% (4 more) while the Adj.ID increased by over 372% ($21 million) and the ATD increased by almost 64% ($6 million). So similar to 2005, 2021 had such a large increase in Adj.ID and ATD percentages because of the extreme levels of destruction caused by the hurricane, rather than the number of storms occurring. Unlike 2005, the number of storms actually increased, although the amount of insurance and damage increase is extremely unproportional to the number of storm increases, so the storms must have been much more destructive than previous years.

According to “2022 U.S. billion-dollar weather and climate disasters in historical context”, 2022 was one of the top three most expensive storms in the US, while one of the other most costly was 2005. This is similar to the results we found and displayed in figure 2 where 2005 and 2021 are outliers. It is important to note, however, that the article refers to several types of natural disasters, not only storms. They specifically point out Hurricane Ian, which occurred heavily in Florida, was a category 4 storm in 2022 that caused around $112.9 billion in damages across all areas affected. It was also the first storm to occur in Florida that exceeded the typical approximate value of $50 billion dollars that is usually spent on uninsured and insured damages, reaching over $100 billion dollars. Hurricane Ian clearly had a large impact on insurance rates and coverage in the year 2022. The statistics chart shows us that there was a decrease in storms by almost 22% (5 less) from 2021 to 2022, but the Adj.ID increased by about 9% ($22.5 billion) and ATD costs increased by about 20% ($7 billion). note

So 2021 and 2022 made 2 consecutive years of having extremely high insured damage and total damage after relatively low numbers from 2006 and on. The articles showed us that the statistics in our chart, specifically the adjusted insured damage and adjusted total damage were caused by the category level of the storms occurring those years because they caused significant damage to the locations they affected. The number of storms did not matter as much because as we saw, some years had an increase while others had a decrease in the number of storms.

We have created interactive visualizations with Tableau, the first one we have created is a timelapse of standardized ZHVI values from 2000 to 2025 across America. This visualization provides us with major hot spots of high ZHVI which we can use in comparison with the second visualization which is the map of the continental U.S affected by a given drought level which the user can select. There are two different dashboards, one is a map of the whole U.S and a coloring of it based off of percent area affected by selected drought level. The second dashboard is split into 2, a histogram that displays the top ten states affected by D4 Drought during a selected year and a map of a selected state and a coloring based off of percent area affected by selected drought level. Users can click on a state's bar in the histogram to change the map to that state. On both of these dashboards the user can select a year from the dropdown, the default is average over the years of the data.

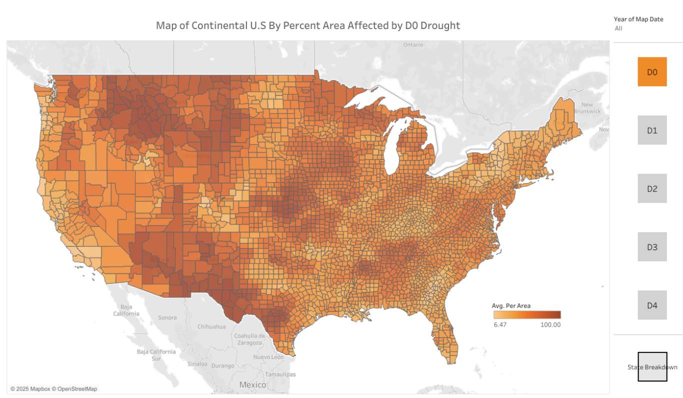

Figure 10: Drought Map Dashboard 1

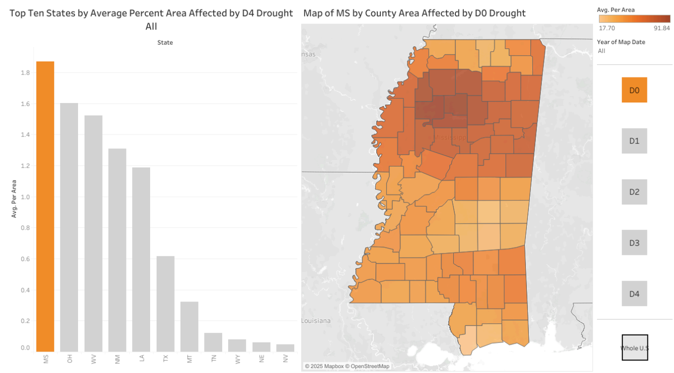

Figure 11: Drought Dashboard 2

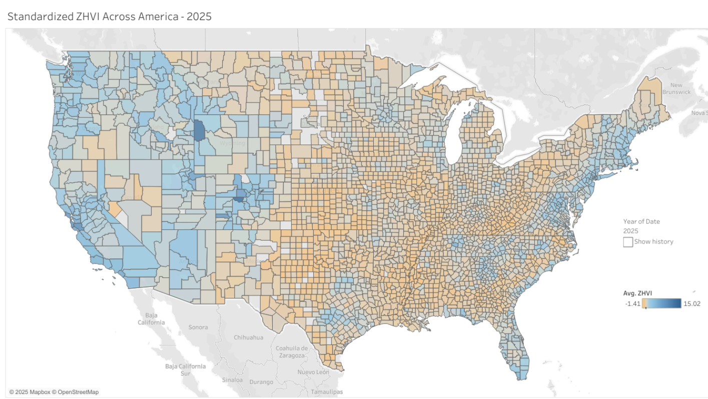

Figure 12: Standardized ZHVI Timelapse

We decided to try and do some machine learning methods to try and predict ZHVI. When first doing exploratory analysis we plotted what we expected to be our strongest predictor D4 against our response ZHVI. The results are symmetrically and randomly distributed around zero and no clear trends which does not bode well for linear regression but we decided to try it anyway.

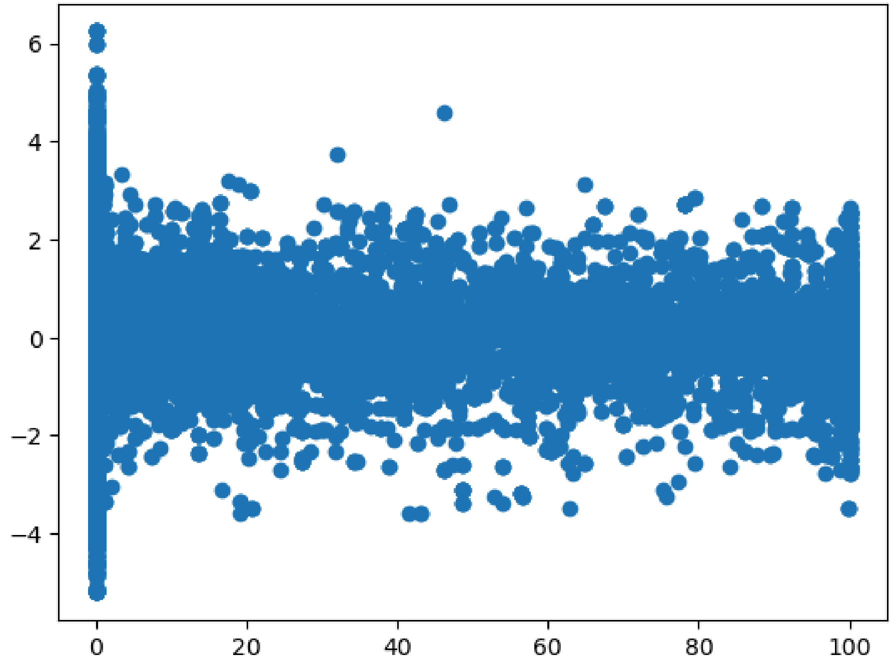

Figure 13: Scatter plot D4 against Percent Home Value

The results of our linear regression analysis were quite poor, indicating a weak relationship between the predictor variables and the response. The highest R squared value we obtained was only 0.0166, using Month After Per Change as the response variable. This model included all drought level indicators as predictors, excluding the None category to avoid multicollinearity. An R-squared value of 0.0166 suggests that the model explains only 1.66% of the variability in the response, meaning that drought levels do not appear to be strong predictors of changes of ZHVI over time in this context. This result suggests either a very weak underlying relationship or the need to incorporate additional variables or a different modeling approach.

With the poor results from linear regression, we decided to test logistic regression and see if that model would perform better. Again our best model included the response variable as Month After Per Change but also performed poorly. Our model predicted a vast number to be 1 (Positive Percent Increase) and misclassified a lot of 0 (Negative Percent Decrease). Furthermore the model's ROC curve yielded an AUC of 0.56, indicating very weak discriminatory power. Since an AUC of 0.5 corresponds to random guessing, a score of 0.56 suggests that the model is only marginally better than chance and reinforces the conclusion that there is a weak relationship or perhaps more and better predictor variables are needed. Some predictors that might be a good idea is median household income and population,

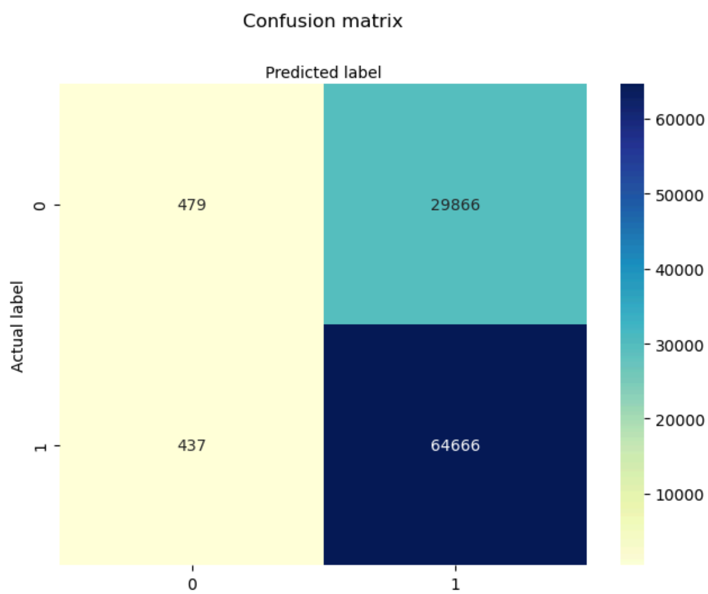

Figure 14: Confusion Matrix

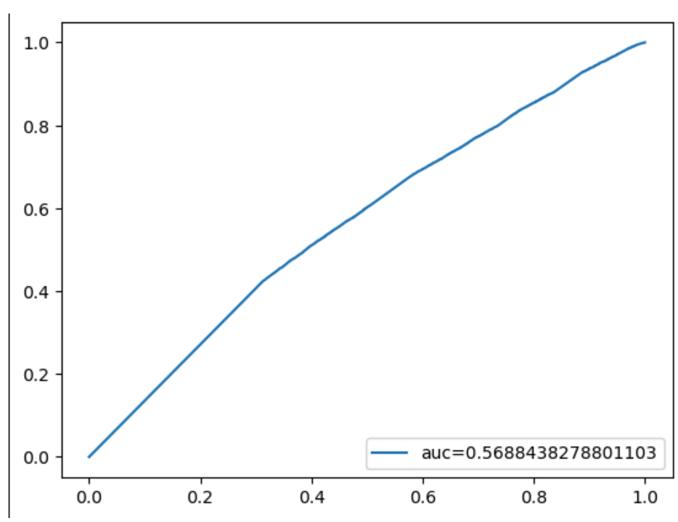

Figure 15: ROC Curve of logistic regression

Conclusion

Our research highlights the growing influence of climate related disasters on real estate markets and insurance systems in the U.S.. The clearest finding is that natural disasters, especially wildfires, droughts, and storms, have measurable but often complex effects on home values, depending on the type of disaster, geographic location, and timing. From our case study on the Camp Fire in Butte County, we found that even severe disasters don't always lead to a long term drop in home prices, possibly due to broader market forces like COVID-era housing spikes or federal aid. Our storm analysis showed that while the number of events has increased, the financial impact in terms of insured and total damages doesn't follow a simple pattern, reinforcing findings from Barthel and Neumayer that rising losses often reflect development trends more than disaster intensity.

Throughout our analysis, we found strong support for the arguments raised in the literature. For instance, Contat et al. and Martino both point to the limits of public data and the role of risk perception in driving housing prices, challenges we also faced when merging datasets and creating models. Taylor and Schroeder's work on insurance instruments aligns with our storm analysis, where the gap between adjusted insured losses and total damages became clear. This suggests there's still room to improve how the insurance industry protects against climate related housing loss.

Looking forward, one way to strengthen our research would be to incorporate more detailed data, like single-family home data or more local insurance claim records. That would help us see how different neighborhoods or different types of homes are affected by disasters. Running more advanced models like random forests could also improve how well we predict housing trends in areas at risk. Adding in things like climate risk maps or information on local building codes might also help account for how some places are already trying to reduce risk.

From a policy perspective, our findings suggest a need for more transparent risk communication and expanded federal support for hazard insurance. Subsidies that hide real risk, like those Martino critiques, may offer short term relief, but distort the long term pricing and investment. As climate change continues to stress housing markets, better insurance mechanisms, accurate risk mapping, and data shaping between public and private stakeholders will be essential to support both property owners and market sustainability.

References

Barthel, F., & Neumayer, E. (2011, December 10). *A trend analysis of normalized insured damage from natural disasters*. Climatic Change. https://link.springer.com/article/10.1007/s10584-011-0331-2

Blackwell, J. (2021, November 30). *Active 2021 Atlantic hurricane season officially ends*. National Oceanic and Atmospheric Administrationnoaa.gov/news-release/active-2021-atlantic-hurricane-season-officially-ends. https://www.noaa.gov/news-release/active-2021-atlantic-hurricane-season-officially-ends

Canario, Y. (n.d.). *Home Insurance*. Kaggle. Retrieved May 6, 2025, from https://www.kaggle.com/datasets/ycanario/home-insurance

Contat, J., Doerner, W., Renner, R., & Rogers, M. (2024, May). *Measuring Price Effects from Disasters Using Public Data: A Case Study of Hurricane Ian*. FEDERAL HOUSING FINANCE AGENCY. https://www.fhfa.gov/document/wp2404.pdf

Contat, J., Hopkins, C., Mejia, L., & Suandi, M. (2024, April 15). *When climate meets real estate: A survey of the literature*. Real Estate Economics. https://onlinelibrary-wiley-com.proxy.binghamton.edu/doi/full/10.1111/1540-6229.12489

Cui, J. (2023, September 20). *Map: New data shows where natural disasters could affect homeowners' insurance*. NBC News. Retrieved April 29, 2025, from https://www.nbcnews.com/data-graphics/map-new-data-climate-change-homeowners-insurance-rcna105632

Dantewadia, P. (2025, January 27). *Map Shows Rising Home Insurance Costs by State as Climate Risks Reshape the Market*. Realtor.com. Retrieved May 6, 2025, from https://www.realtor.com/news/trends/maps-show-insurance-costs-rising/

Farris, G. S. (n.d.). *The Major Hurricanes of 2005: a Few Facts*. The Major Hurricanes of 2005: a Few Facts. Retrieved May 6, 2025, from https://pubs.usgs.gov/circ/1306/pdf/c1306_ch2_b.pdf

Federal interagency wildland fire maps and data for all. (2018). National Interagency Fire Center. Retrieved April 29, 2025, from https://data-nifc.opendata.arcgis.com/

Flavelle, C., & Rojanasakul, M. (2024, December 18). *Insurers Are Dropping Homeowners as Climate Shocks Worsen*. The New York Times. Retrieved April 29, 2025, from https://www.nytimes.com/interactive/2024/12/18/climate/insurance-non-renewal-climate-crisis.html

Hazard and Disaster Magnitude Units. (n.d.). EM-DAT Documentation. Retrieved May 6, 2025, from https://doc.emdat.be/docs/data-structure-and-content/hazard-and-disaster-magnitude-units/

Historical Hurricanes and Tropical Systems Impacting the NWS Mobile CWA. (n.d.). National Weather Service. Retrieved April 29, 2025, from https://www.weather.gov/mob/tropical_events

Housing Assistance Program Data - Owners - v2. (2025, April 13). FEMA. Retrieved April 29, 2025, from https://www.fema.gov/openfema-data-page/housing-assistance-program-data-owners-v

Housing Data. (n.d.). Zillow. Retrieved April 29, 2025, from https://www.zillow.com/research/data/

Insurance as an Alternative for Sustainable Economic Recovery after Natural Disasters: A Systematic Literature Review. (n.d.). MDPI. Retrieved April 29, 2025, from https://www.mdpi.com/2071-1050/14/7/4349

Leeper, R., Bilotta, R., Petersen, B., Stiles, C., Heim, R., Fuchs, B., Prat, O., Palecki, M., & Ansari, S. (2022, April 15). *Characterizing U.S. drought over the past 20 years using the U.S. drought monitor*. International Journal of Climatology. https://rmets.onlinelibrary.wiley.com/doi/full/10.1002/joc.7653

Martino, A. (2023, March). *Climate risk and real estate: influence of weather-releated events on real estate and market response*. Politecnico di Torino. https://webthesis.biblio.polito.it/secure/26430/1/tesi.pdf

Monitoring hazards & disasters worldwide since 1988. (n.d.). EM-DAT - The international disaster database. Retrieved April 29, 2025, from https://www.emdat.be/

Posey, M., & Todoroff, N. (2024, June 28). *The Impact of Natural Disasters on Insurance Rates in 2025*. Bankrate. Retrieved May 6, 2025, from https://www.bankrate.com/insurance/homeowners-insurance/natural-disaster-costs/

Schroeder, M. (2013, May 16). *Environmental Insurance: A Risk Management Tool for Real Estate*. Environmental Claims Journal. https://www.tandfonline.com/doi/full/10.1080/10406026.2013.781471

Smith, A. B. (2023, January 10). *2022 U.S. billion-dollar weather and climate disasters in historical context*. Climate.gov. Retrieved May 6, 2025, from https://www.climate.gov/news-features/blogs/beyond-data/2022-us-billion-dollar-weather-and-climate-disasters-historical

Source Cooperative. (2023, November 9). Source Cooperative - Source Cooperative. Retrieved April 29, 2025, from https://source.coop/repositories/giswqs/nwi/description

Spotlight on: Catastrophes - Insurance issues | III. (2025, February 28). Insurance Information Institute (III). Retrieved April 29, 2025, from https://www.iii.org/article/spotlight-on-catastrophes-insurance-issues

Taylor, Z. J. (2020, January 5). *The real estate risk fix: Residential insurance-linked securitization in the Florida metropolis*. Sage Journals Home. https://journals-sagepub-com.proxy.binghamton.edu/doi/full/10.1177/0308518X19896579

Tinker, R., & Johnson, L. (n.d.). *This Week's Drought Summary*. Current Map | U.S. Drought Monitor. Retrieved 2025, from https://droughtmonitor.unl.edu/CurrentMap.aspx

Tinker, R., & Johnson, L. (n.d.). *U.S. Drought Monitor*. Current Map | U.S. Drought Monitor. Retrieved May 6, 2025, from https://droughtmonitor.unl.edu/CurrentMap.aspx

Welcome to the EM-DAT Documentation. (n.d.). EM-DAT Documentation. Retrieved April 29, 2025, from https://doc.emdat.be/docs/

Young, R. (n.d.). *Business of Disaster | FRONTLINE*. PBS. Retrieved May 6, 2025, from https://www.pbs.org/wgbh/frontline/documentary/business-of-disaster/