Navigating Climate Changes in the Insurance Industry

Attributes of an Effective Data Communication System

by Erica Laidler

December 2023

1 Abstract

In recent years, the world has experienced a dramatic surge in unusual weather and extreme weather events, including catastrophes such as tornadoes, floods, and wildfires. This trend has a profound impact on insurance companies, because natural disasters often lead to property damage and massive losses in claims. Thus, climate change necessitates that companies perform a complete reevaluation of their strategy, with a particular focus on enhancing their data analysis and communication systems.

The objective of this report is to present the key attributes of a universally effective data communication system for property insurance companies. After conducting research and exploration, three main findings emerged. First, insurance companies should establish clear systems to communicate past weather trends and past claims history. Additionally, they need well-constructed predictive models and clear communication of future weather and loss projections by region. Lastly, an insurance company's data system must cater to the diverse needs of employees, providing clear, dynamic interfaces for employees at all levels of the company. To demonstrate the attributes of an effective data system, this report showcases visualizations tailored for specific contexts, along with the outcomes of a multilinear model predicting company losses based on historical climate conditions.

Overall, it is the companies with the most accurate predictive models and robust data communication systems which will excel beyond competitors in the coming years.

2 Introduction

With increasing temperatures, rising sea levels, and a heightened frequency of extreme weather events, the global risk of catastrophe is steadily rising. In 2017, an astonishing 3,433,434 homes were destroyed in Peru, leading to over a billion US dollars in damage. Intense downpours led to river floods and landslides, claiming the lives of 114 citizens and displacing 184,000 more (Son et. al).

Wildfires have also been particularly damaging in the last twenty years. Wildfires have the capacity to wipe out massive amounts of property and nature at once, and are difficult to predict and prepare for. This presents a new and challenging risk to insurers.

Rainfall, hurricanes, wildfires, droughts, floods - all of these climate concerns introduce risk to the insurance industry. Overall, insured property losses in the United States from natural catastrophes have soared in recent years, increasing from 34.1 billion in 2016 to 137.4 billion in 2017 and 62 billion in 2018 (Albin). In his 2012 report, Laurens Bouwer points out that the "increasing exposure and value of capital at risk to future losses" is also a factor, as more properties, especially expensive properties, are built in catastrophically risky zones.

However, despite being unfortunate for the average homeowner, property risk is not inherently unprofitable for insurance companies.

In a property insurance company, profit derives from taking on properties and setting premiums which are high enough to offset the cost of claims and operating costs. According to Dustin Lemick, an author at BriteCo, "A policyholder's premiums are calculated based on coverage type, risk assessment, and the probability of claims. By effectively pricing premiums, insurance companies can cover potential losses and operating expenses while maintaining profitability."

Actuaries and data scientists work together to establish pricing models, which underwriters then apply to specific claims and make pricing decisions on a case-by-case basis according to these guidelines. In its pricing model, a company must consider a wide range of potential risks. The policyholder's history, the protective measures of the property, and the property's risk of destruction by catastrophic event all typically help inform the premium pricing. Policyholders with a lower risk assessment can expect a lower monthly premium.

In recent years, data science has exploded, assuming a much more prominent role in risk assessment and pricing. In the 1980's, Catastrophe (CAT) models were developed, and quickly grew to become the standard in the industry for assessing climate-related risk of destruction. From then on, insurance models have become increasingly sophisticated. An article in Forbes in December 2015 by Bernard Marr announced that "big data is changing insurance forever", underscoring the industry-wide shift towards the extensive use of data on a massive scale.

In the face of new climate changes, insurers must place a newly increased emphasis on the precision of catastrophe modeling and the implementation of thoughtful systems for effective decision making, particularly in highly catastrophic zones. According to Jordan and Philips, there is "an expectation that the insurance sector should be leading business efforts in responding to climate change due to their inherent exposure to increasing risks." Failing to address these risks could lead to substantial losses in claims.

In his article about risk, Pedro Diaz de Leon points out that climate change will affect the industry in a large number of ways, including the calculation of reserves and the premium pricing model. "Insurers may face higher claims payouts, increased reinsurance costs and difficulties in assessing and pricing climate-related risks. There are significant regulatory requirements currently in place related to climate change, and these requirements will continue to intensify as regulators continue to focus on climate change," writes Diaz de Leon.

In 2007, the ClimateWise principles, a voluntary initiative within the insurance sector, was launched as a set of guidelines for insurance companies to address climate-related risk. While this offered some support, it was not widely implemented (Jones et al). However, around this time, insurance companies across the world did begin to take a serious look at their models and strategies.

In India, where variable rainfall, coastal flooding, and intense droughts and cyclones are frequent, Mamata Swain writes about the huge risk to agriculture and the subsequent rise of crop insurance schemes.

Later, in 2022, Gatzert and Reichel discuss the increasing adoption of climate-related awareness and initiatives among insurance companies in the United States and Europe. Using logistic regression analysis and a linear fixed effects model on data from the Refinitiv Eikon database, they found that from 2009 to 2018, property and casualty insurance companies were most likely, out of all insurers, to prepare for climate change. This said, European insurance companies were significantly more likely than American companies to exhibit awareness of climate change-related risks. This demonstrates a real opportunity for American insurance companies to bridge the gap and take charge on reevaluating their systems (Saffron and Smith).

There are many possible approaches for grappling with climate changes. One option is for insurance companies to pull out of particularly catastrophically risky zones altogether. California, for example, is known for its extensive problems with wildfires, earthquakes, floods, and droughts. In response, seven of the top 12 insurance companies in the state, including Allstate, State Farm, Farmers Insurance and American International Group (AIG), have withdrawn from California or scaled back their provision of new policies in the last year (Brooks).

Another option is for insurance companies to heavily encourage or require policyholders to implement cautionary measures in their properties to help prevent potential damage. This might include installing a basement waterproofing system, spark arresters on chimneys, or storm shutters.

Enhancements in technology are also important. Diaz de Leon states that the companies which effectively automate tasks at scale will set successful companies apart from their competitors. "Rapid advancements in technology can disrupt traditional insurance models. Insurtech companies leveraging AI, big data and blockchain can provide innovative insurance products and services, challenging established insurers."

Specifically, robust data analysis and communication systems play a pivotal role in an insurance company's response to climate change. Developing effective CAT models and pricing systems is important. Utilizing a system of communication, including comprehensive visuals which efficiently synthesizes trends across various regions and catastrophic events, is also essential.

This communication of data trends is so crucial that some researchers have undertaken assessments to determine the characteristics of a visual representation which are most effective and resonate most strongly with viewers. In 2013, O'Neill and Smith conducted a critical analysis of climate change imagery across various societal realms, including in news media, advertising, climate science, art, and marketing. They observed trends to suggest that certain types of imagery have gained popularity, shaping " the cultural politics of climate change in important ways". Thus, it is evident that the way climate change is communicated, whether in insurance companies or society as a whole, greatly influences interpretation.

In 2023, Li et. Al found, for instance, that in a study of over a thousand U.S. adults, artistic visualizations elicited stronger positive emotions than neutral data graphs. It is likely that designing visuals that optimize clarity and viewer engagement ensures continuity throughout a company, enhances employees' understanding of climate-related risks, supports better-informed underwriting decision-making, engages stakeholders, and overall prepares employees to respond adequately in the face of climate change.

To enable effective data communication, it is important to have a system of functional underlying models, which predict property loss in various regions based on expected weather conditions.

In 2014, Odening and Shen wrote about the challenges inherent to this prediction task. "Weather risks show characteristics that often violate classical requirements for insurability," they wrote. Climate change has produced an increase in the "volatility of weather variables" and brought about "non-stationary loss distributions." In other words, weather events have become more variable, both in their occurrence and the likelihood of causing destruction. This increased variability makes them significantly more challenging to predict and accurately price. For instance, regions like the Northeastern United States, traditionally characterized by relatively predictable cold winters and warm summers, is now experiencing much warmer temperatures in the winter and more rain. However, it will occasionally face extremely heavy snowfall capable of causing substantial damage to roofs and cars. Odening and Shen introduce techniques like time diversification and other statistical tools to address this challenge.

Another challenge is that different weather conditions exhibit distinct characteristics, encompassing variations in their likelihood of occurrence across different locations and their potential for causing damage. Some weather conditions, like drought, are slower to develop and fairly predictable, affecting certain regions for an extended period. Droughts do not typically cause sudden immense destruction, but can cause problems over time. On the other hand, wildfires are savagely destructive, may progress rapidly, and are very difficult to anticipate. Heavy rain, with a relatively high likelihood to occur across most regions, may cause some destruction. In contrast, tsunamis are very rare and only affect coastal areas, but are far more likely to cause sudden massive destruction when they occur. This diversity demonstrates the need for tailored model development, risk assessment, and preparation for each type of weather condition.

One approach to address the various characteristics of weather conditions in the prediction process is to focus on one catastrophic event at a time. In 2017, Lin and Shullman focused on flooding, introducing an integrated dynamic risk analysis for assessing the risk of coastal flooding in different regions "at regional scales, considering integrated dynamic effects of storm climatology change, sea-level rise, and coastal development." Their framework is composed of two components: a modeling scheme to collect and combine information, and a formal Poisson-based model to derive various risk measures of interest.

In 2014, Royse et. al also developed a method for predicting occurrences of flooding and simultaneously estimated the distribution of losses associated with potential flooding. "In this paper we show that integrated environmental modeling (IEM) techniques can be used to generate a catastrophe model for groundwater flooding. Catastrophe models are probabilistic models based upon sets of events representing the hazard and weights their likelihood with the impact of such an event happening which is then used to estimate future financial losses." Thus, this approach considers the probability of a flood occurring, the likelihood of its severity, and the potential levels of damage associated with this event.

This methodology proves particularly useful in the insurance industry, where merely knowing the number of floods in a given year is insufficient. The model is creative because it accounts for both the potential flooding risks and the loss associated with these risks, all weighted by the probability of their occurrence. This type of model might be expanded to a more intricate system or set of models which take into account many catastrophes at once, each with their associated risk of damage and loss.

In summary, climate changes are prompting adaptive measures in the insurance industry, such as ceasing policies in highly catastrophic regions, prioritizing the implementation of proactive measures in homes, optimizing the use of technology, and establishing functional systems of data analysis and data communication to all levels of the company.

This report will primarily focus on this latter point, presenting a standardized data communication system for insurance companies to use at every level of operation. In insurance companies, it is not enough to simply own data; the utility of data analysis is dependent on its clear and comprehensible communication. It is crucial for insurance companies to develop a system that effectively presents information on both a large and specific scale.

In this project, I plan to outline the key attributes of an effective system of data communication within a property insurance company. I will demonstrate how past trends and future projections can be communicated to employees via tables and visualizations. I will also develop a linear model to predict future claims losses by state, based on previous climate conditions. Overall, the goal of an insurance company is not to accurately predict every single catastrophic event and the ensuing damage down to the location and time. Instead, the goal is to maintain a system of informed models and visualizations which enable employees to holistically prepare for climate change, remain up to date, communicate well, and charge premiums in order to abundantly cover all losses and make profit.

3 Methods

The development of an effective data system at a property insurance company requires a nuanced understanding of the insurance industry, the data involved in risk assessment and pricing models, and an analysis of the methods which most effectively communicate important data insights to employees. In this project, I thoughtfully considered the most important attributes of an effective data communication system. I drew upon my experience from working briefly at an international insurance company and delved into further research on how insurance companies are navigating the implementation of expanding systems of data. I also explored many different techniques to assess their effectiveness in different contexts.

One important attribute of an effective system includes visuals. I experimented with many data visualizations, comparing their functions, unique traits, and parameters. To assess various data visualizations and decide on some of the most effective approaches, I experimented with precipitation data from the United States. I used a data set from The National Centers for Environmental Information, which contains extremely detailed climate information over the past few centuries. This particular data set I used contains comprehensive data on precipitation in the United States. My focus was on determining the best way to document the progression of precipitation trends in various regions, with a specific interest in documenting increases in the frequency of severity of precipitation, and comparing trends between regions.

I found that some visualizations, like bar or line plots, are well suited to depict the progression of precipitation over the years. Other plots, like overlaid bar plots, are useful for comparing catastrophic events over the years between two regions. In contrast, a map of the United States can effectively compare climate events based on relative severity, with the use of an accompanying legend. Overall, each visualization has particular features that may make it better suited for specific tasks, whether on a large scale, such as presenting a substantial amount of information to a stakeholder at once, or on a small scale, such as isolating the trend of a particular catastrophic event in a specific region. Certain visualizations are also more suitable for specific audiences, such as executives or underwriters. A well-chosen system of visualizations helps optimize the effectiveness of data communication between employees at all levels of the company.

In addition to brainstorming important attributes of a data communication system and determining some of the best visualizations for communicating past trends, I also analyzed the best way to communicate about future predictions. To begin, I built a linear model to predict loss based on climate variables. The intention was to create a model to predict future property loss in the US based on the catastrophic weather event data from previous years.

In this model, the response variable, incurred losses, represents an insurance company's incurred losses as a result of property claims, by state, in dollars, in a given year. This is an important value because insurance company profit derives from the net value of premiums subtracted by losses and operational costs. Though loss alone is not an adequate assessment of company profit, it does provide insight on the amount of property destruction which occurred in a given year. The data for the loss variable came from The Insurance Institute, which represents data from a large body of insurance companies in the United States. Specifically, I import the tables containing Incurred Losses By State, Property/Casualty Insurance, for 2019, 2020, and 2021. Note that the loss data comes from property and casualty insurance, due to a lack of tables which exclusively represented property data. This is a limitation of the model, but the results are still interesting.

The predictors consist of climate data. Specifically, the predictors are the measurements, by state, of heavy rain, winter weather, ice storm, heavy snow, hail, tornadoes, windy weather, floods, wildfires, and dust storms which have occurred in the past few years. The data comes from The National Centers for Environmental Information website, specifically from their Storm Event Database from 2019, 2020, and 2021. The units for these measurements vary depending on the weather condition, however their relative values retain the important information.

The model is a multiple linear model. This was appropriate for the task because the objective was to predict loss based on ten continuous variables. Before implementation, it was important to test the linear assumptions, including linearity, independence, constant variance, normality of residuals, and no perfect multicollinearity, which hold.

It was not evident at first exactly how to structure the linear model. For instance, the model might represent loss in one year by training on data from the previous year. However, I decided that it is better for the model to train on three years of data, and to take the average between the annual total measurements from each of the three years. This helps control for variability from year to year. Also, this means that the model will be able to take climate data from the past three years and make predictions regarding the loss expected in the following year.

Some specifications were made to the model. A few outliers were removed from the data, because they were very extreme outliers in certain measurements and interfered with the linearity assumption. A log transform was also performed on the response variable, to further mitigate the impact of outliers. Forward selection was used for variable selection, and the best combination of predictors was retained.

4 Results

The goal of this project was to propose an effective, dynamic system for an insurance company to systematically communicate data insights and meet the needs of employees at all levels of the company. An optimal set of systems will prepare an insurance company to confront the challenges of climate change in the coming years.

Overall, a universally effective system for data communication at an insurance company must meet several important criteria. This includes having clear and efficient systems for analyzing and visualizing past trends in climate conditions and claims. Additionally, there should be a well-structured framework for predicting and communicating projections of future weather conditions and expected loss or claims, categorized by region. Lastly, the system should encompass various forms tailored to meet the diverse needs of employees at all levels of the company. These systems should be dynamic, regularly updated, and provide efficient access to information by region, year, and catastrophic event, both for historical and future data.

Some of the most important positions in an insurance company include executives, actuaries, data scientists, and underwriters. Executives make decisions about the company on a wide scale. Data scientists develop predictive models to predict future trends in weather conditions and assess risk. Actuaries work with data scientists to set pricing guidelines. Finally, underwriters work on a case-by-case basis to determine what policies to take on and what premiums to charge, based on pricing guidelines and an assessment of policy risk. Ultimately, underwriters must make balanced decisions, taking on moderately safe policies and charging enough premium in order to offset future losses and the operational costs of a company.

The first requirement of an effective data system is that it must provide employees with access to historical data, including data on past climate events and property claims. This is important because it helps explain why there may have been a lot of claims in particular regions in the past years. Knowing about past weather conditions also helps the employee understand whether current weather events are an unusual circumstance. For instance, if this is the first year that Oregon is experiencing wildfires, then it suggests that underwriters may have to make changes in their policy decisions.

As mentioned, the optimal presentation of past climate data may vary based on the needs of the employee. For instance, stakeholders and executives tend to benefit from reports and presentations which present data on the macro scale. They are interested in large-scale trends which are likely to affect the company as a whole.

For instance, in a one-time presentation to stakeholders, the presentation should be centered on macro data, simplified to the trends which best support the presenter's argument. Forming the presentation based on the needs of stakeholders by keeping data presentation as comprehensible and effective as possible, while still encompassing a huge amount of information, is key. Considering the number of variables at play, including different regions, climate conditions, and years, this can be a challenge.

Imagine that a speaker is passionately advocating for the adoption of new procedures to address the escalating challenges posed by climate extremes. It would useful for this presenter to provide plots which depict the trajectory of intensifying extremes in precipitation across the country over the past century.

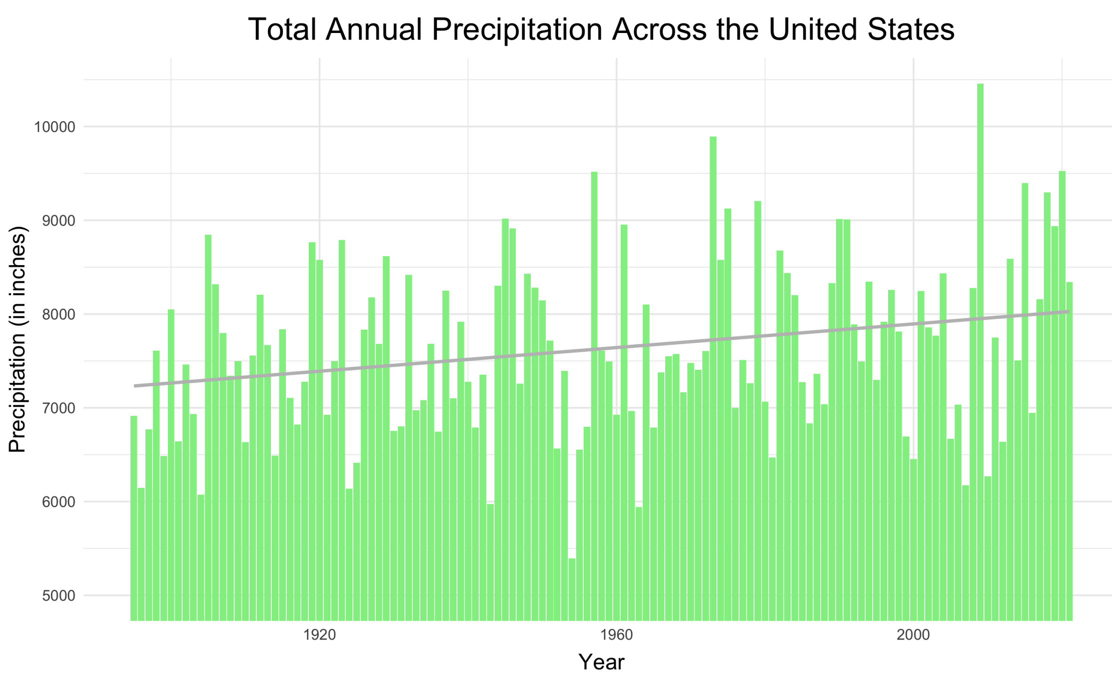

Figure 1.

The plot above depicts the total annual precipitation, measured in inches, for each year across the United States from 1890 to 2021. In general, the amount of precipitation across the country has been climbing, as depicted by the gray line, which represents the fitted linear trend in the data. In 1890, there was less than 7,000 inches of total precipitation observed across the United States, which climbed to over 9,000 inches in 2020.

A presentation to stakeholders might also present a plot depicting the trends over time for a particular state or region.

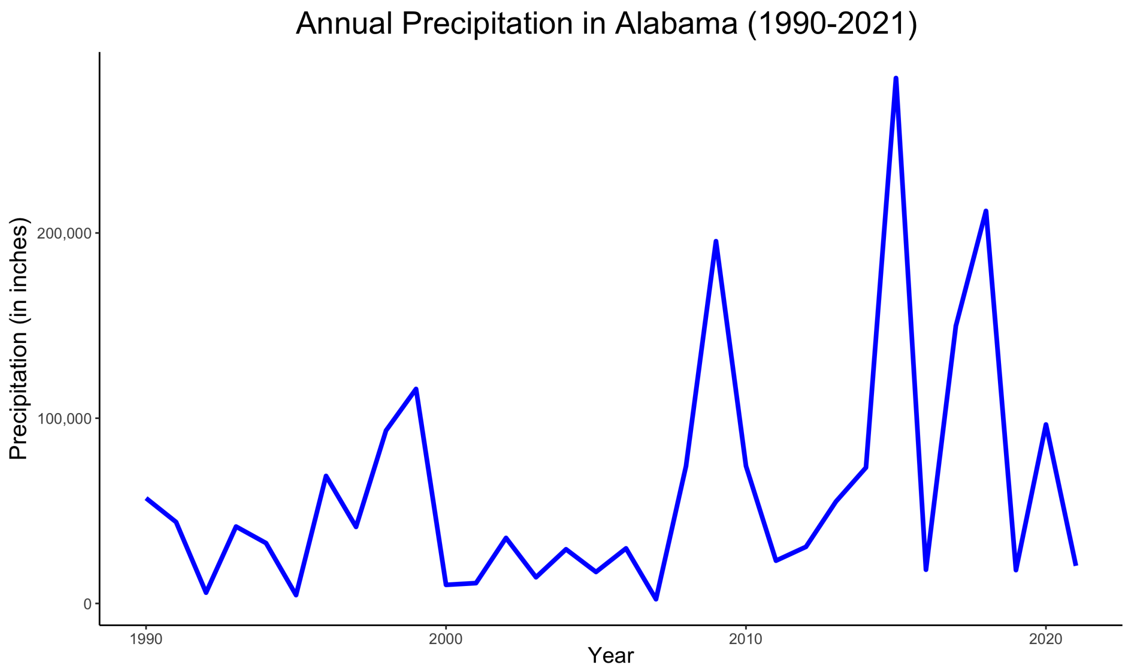

Figure 2.

The plot above illustrates the annual precipitation in Alabama from 1990 to 2021. Alabama, renowned as one of the rainiest states in the United States, exhibits a particularly dramatic surge in precipitation over this period. In 1990, the recorded precipitation was approximately 50,000 inches, but around 2015, there was a massive spike, reaching over 300,000 inches.

The plot also depicts heightened and intense fluctuation after 2010, indicating not only a trend toward increased precipitation over the years but also a general increase in weather variability. This implies that certain years may experience considerably lower precipitation levels than others, making prediction more difficult. This may make it more likely for insurers to wildly under- or overestimate potential damages in particular regions. For this reason, this plot highlights to stakeholders the importance of prioritizing the development of accurate models and maintaining them consistently.

Alternatively, a presentation addressed to stakeholders might be slightly more niche, making the argument that a certain region is too risky to be profitable based on trends in recent years. For instance, many insurance companies made the executive decision to cease all policies in California after determining that the wildfire situation and other risks made the region too unpredictable.

The plot below compares the precipitation in inches between Georgia and Kansas for each year from

1990 to 2021.

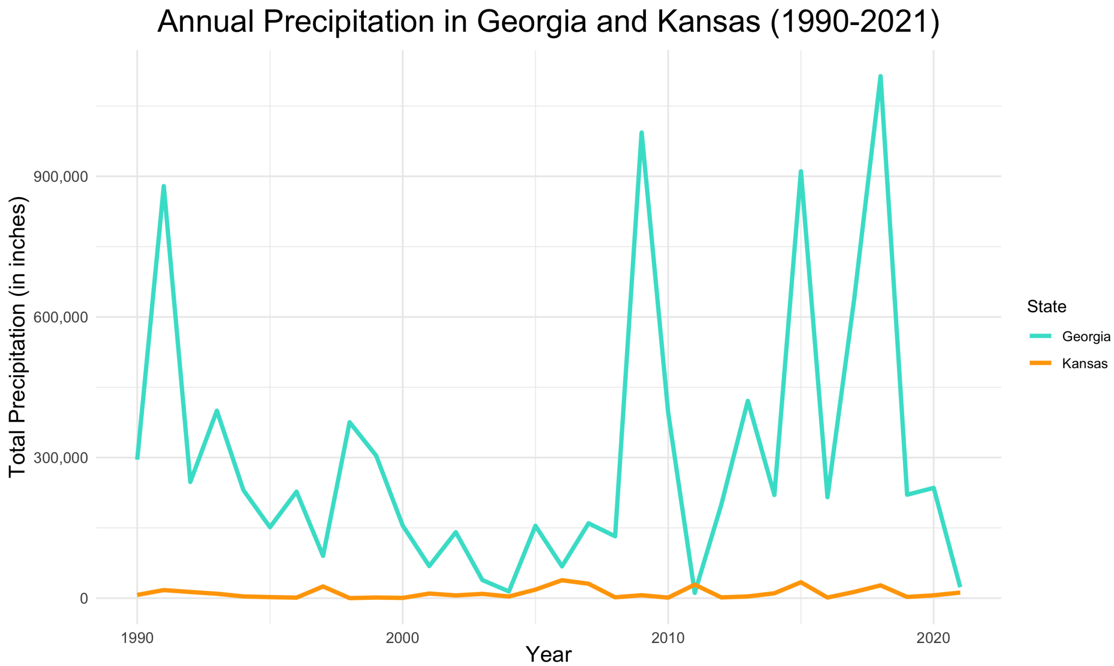

Figure 3.

Georgia, depicted as the blue line in the plot above, experiences large fluctuations in data, experiencing some years with an incredibly high amount of precipitation, and years with much less. Overall, however, it tends to experience more precipitation than Kansas. For example, there was an observed total of over 100,000 inches in Georgia in 2017, as compared to the approximate 20 inches in Kansas on the same year. Despite that precipitation is increasing overall across the United States, this trend does not hold true for all states, such as Kansas. Plots like the above could demonstrate to stakeholders the importance of assessing region-specific trends to make decisions on a state-by-state basis.

Overall, presentations intended for stakeholders and executives at the company should address just one or two main points, isolate the most relevant statistics, and present wide-scale data, such as trends over a specific region over the last forty years. In this case, the most important goal should not be complete information exhaustion, but the prioritization of concision and clarity. The priority should also be to adequately present a particular point, support it with data, explain how it relates to the company structure, and suggest what should be done as a result. Data presentation at this level is focused on narrative structure and the presentation of clear information to guide major decisions.

In addition to watching presentations, executives responsible for important company decisions should be updated on trends in climate conditions and company profits on a regular basis. This type of data presentation should be more thorough than a presentation, with the purpose of providing insight on important general trends from the last week or month. This will cover any gaps in their knowledge and prepare them to interact and understand issues occurring in all areas of the company.

In this case, the data should take on the form of summaries and visuals of trends for each catastrophic event in various regions, as well as any particularly unusual and notable irregularities. In other words, they should receive a report which consists of the trends in wildfire, hurricane, winter storms, and droughts in each region in the US from the past month. Then there should also be an explanation or plot to represent weather conditions which deviate from the norm, which will differ on a month-by-month basis. The reports should also include predictions for future climate conditions and losses.

Additionally, these reports may provide insight regarding policy decisions in particular regions. For instance, suppose that there is a proposition to cease or reduce policies in the Southwestern United States. If this is a relevant current topic for the company, a regular report might include data to help guide the executive's decisions.

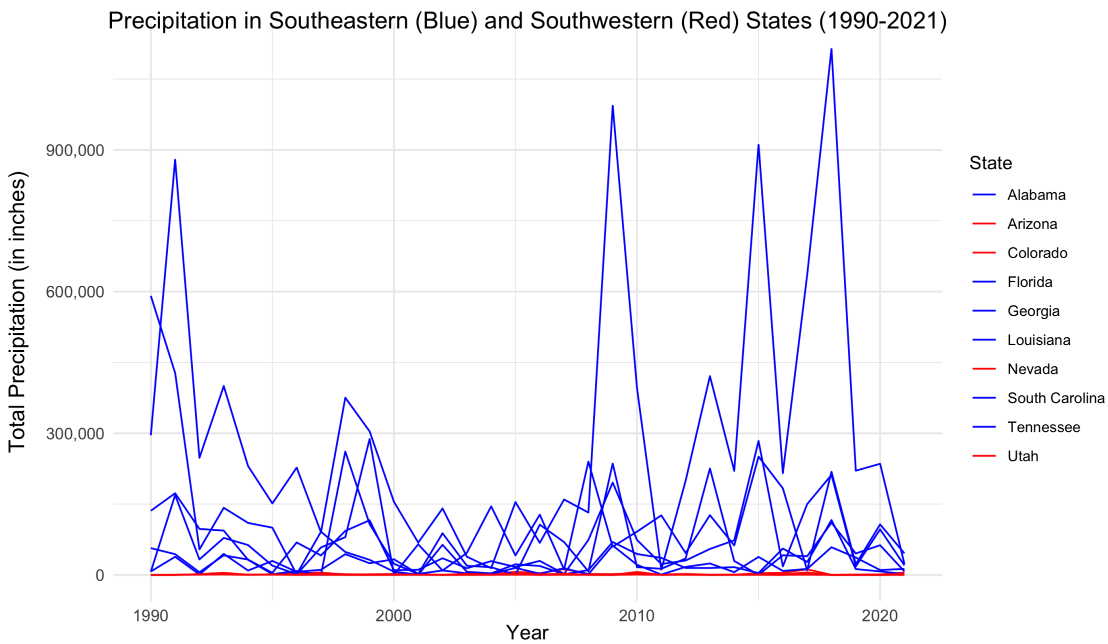

Figure 4.

The line plot above compares the precipitation, in inches, for each year from 2017 to 2020 between Southwestern and Southeastern states. The Southeastern states, Alabama, Florida, Georgia, Louisiana, South Carolina, and Tennessee, are represented by the blue lines, while the Southwestern states, including Arizona, Colorado, Nevada, and Utah are represented in red. This plot clearly represents the much greater rates of precipitation in the Southeastern states, overall, as compared to the Southwestern states, during this time frame. This can help inform decisions regarding the policies in the Southeast and Southwest regions.

The box plot below presents a different visualization of similar information.

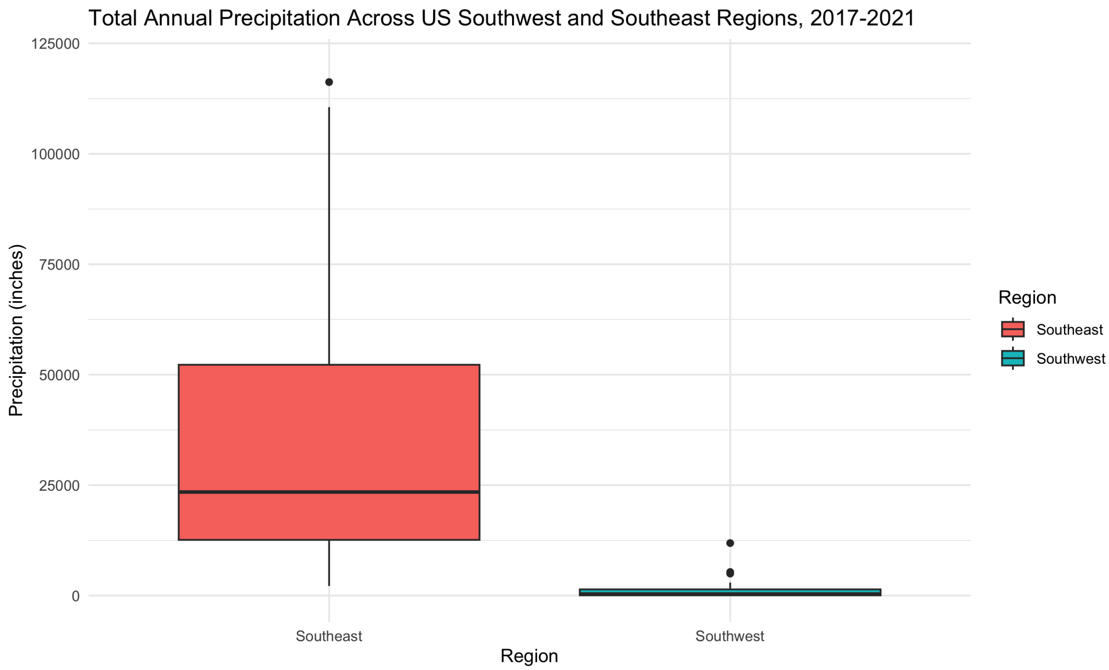

Figure 5.

The boxplot above compares the distribution of total annual precipitation recorded across different states within the Southwest and Southeast regions of the United States from 2017 to 2021. Each boxplot illustrates the distribution of annual precipitation values for a specific region. The x-axis distinguishes between the Southwest and Southeast regions, and the y-axis represents the total annual precipitation in inches.

Again, it is evident that the Southeast region of the US experienced far more precipitation in the years from 2017 to 2020, as compared to the Southwest region. Overall, the median precipitation out of the total annual precipitation measurements in the Southeast region was just below 25,000 inches, while the median precipitation for the Southwest region was below 1,000 inches. However, the Southeast region also displays a greater range of data, which suggests that the precipitation in this region is more variable.

Overall, the goal of these regular executive reports are to educate the people on the most recent trends in climate. The time frame may vary; a company may choose to produce quarter or mid year reports instead. The important part is that there are thorough, regular reports consisting of all necessary information in a variety of regions and catastrophic events over the past months in order to keep employees at the executive level up to date on these trends. Unlike presentations, they do not require a narrative, but should still contain information on a mass scale.

Finally, the needs of underwriters are extremely important, because they make policy decisions on a case-to-case basis. Underwriters benefit from access to very specific micro-level climate data. They need access to a dynamic interface which allows them to toggle between states, regions, years, and catastrophic events, and see relevant data to help them make decisions with respect to a specific policy.

For this reason, I propose the method of displaying the data in the form of a map, in which each color represents the extremity for a given catastrophic event.

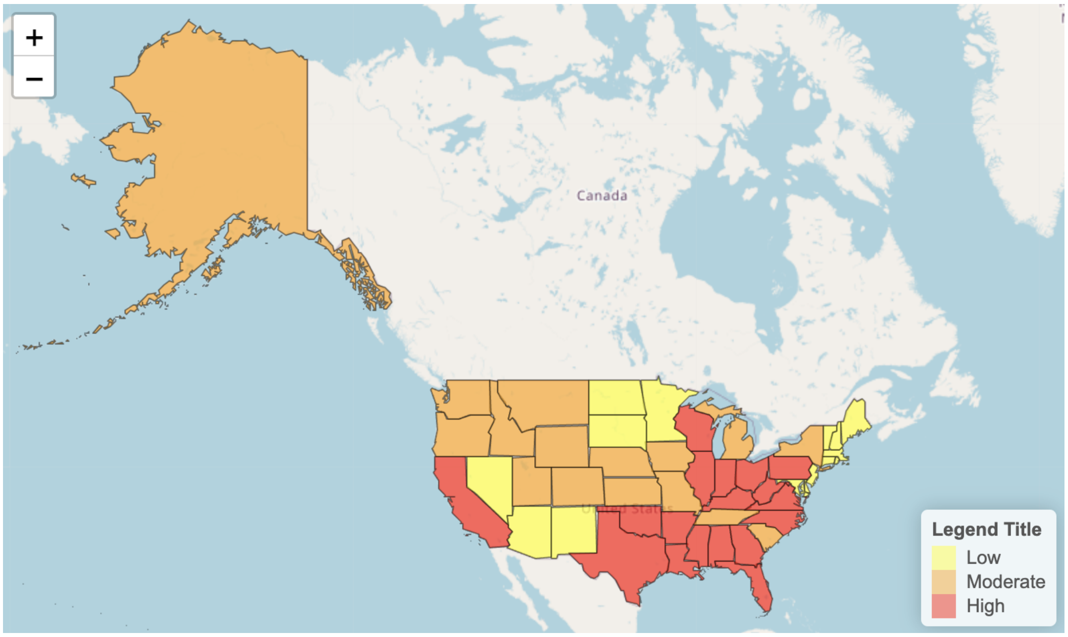

Figure 6. Map of the United States, Categorized by Total Precipitation in 2017.

The map above is an example of a visualization which would be useful to underwriters. The map illustrates precipitation levels across states. In this context, 'precipitation' refers to the observed precipitation for each state in 2017, in inches. States experiencing 'low' precipitation are represented in yellow, those with 'moderate' precipitation are in orange, and 'high' precipitation states are denoted in red. The low bracket refers to states which received a maximum of 3,785 inches, the moderate bracket refers to states which received a minimum of 3,785 inches and a maximum of 17,825 inches, and the high bracket refers to states which received over 17,825 inches of precipitation.

An interface designed for underwriters might allow them to choose a particular catastrophic event, or time period, and see a map visualization based on this data. Brackets would vary based on the context. Additionally, it might be useful for underwriters to see visualizations to compare the severity of catastrophic event in different states for different time periods.

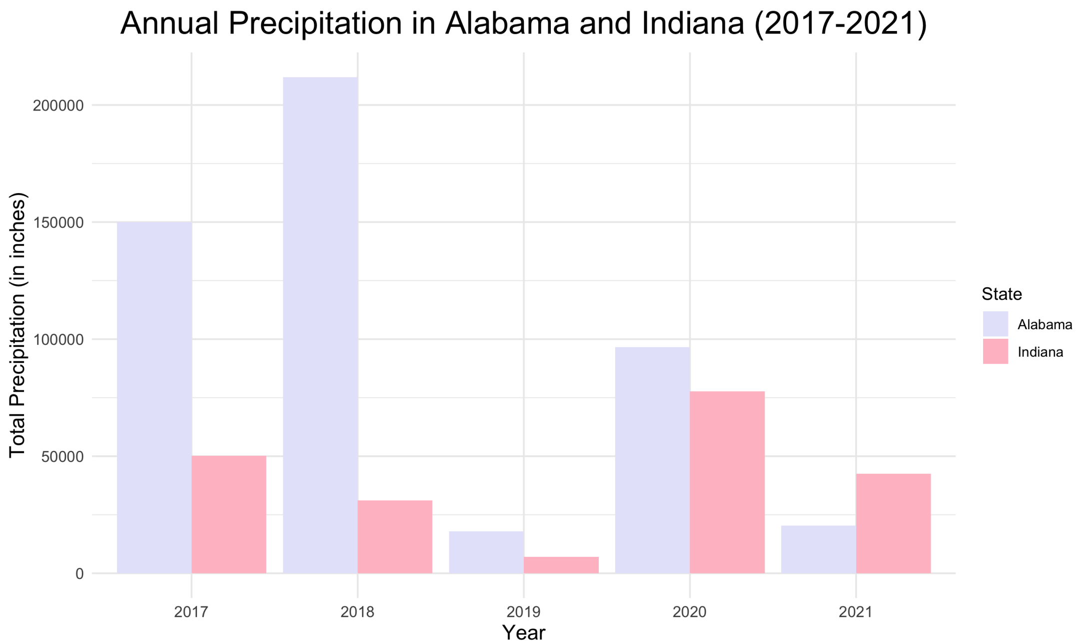

Figure 7.

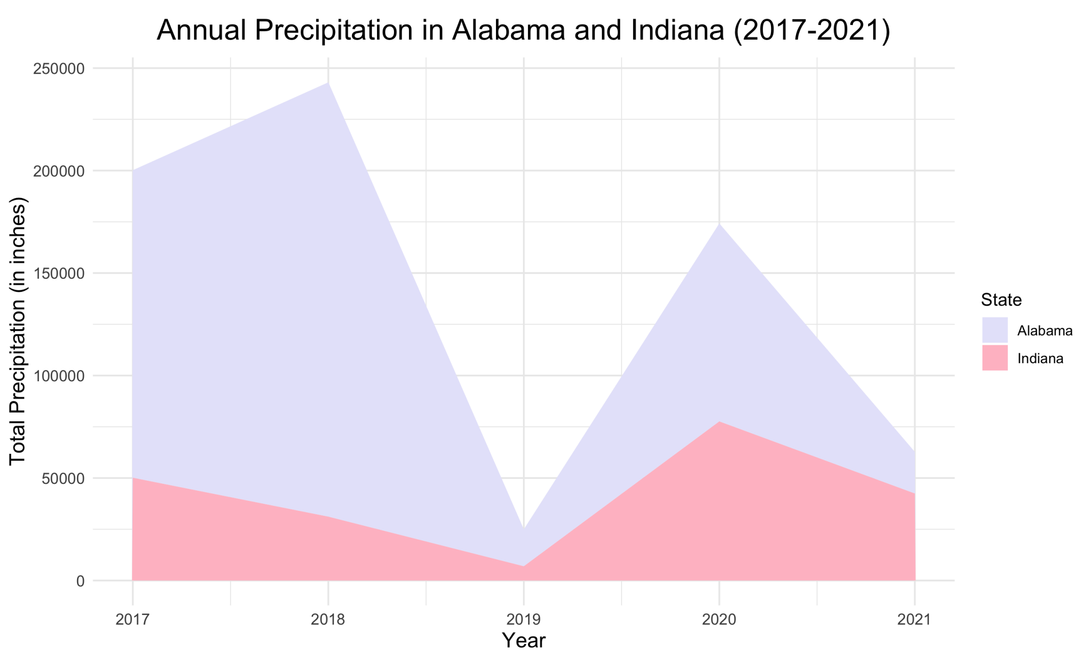

Figure 8.

The plots above depict the annual precipitation in Alabama and Indiana, in inches, from 2017 to 2021. Based on the first plot, the bar plot, it is evident that Alabama tends to experience more rain than Indiana. 2017 and 2018 were particularly rainy states for Alabama. Alabama reaches a peak of over 200,000 inches of rain in 2018, while Indiana experiences less than 50,000 inches that same year. Additionally, it appears that Alabama tends to experience a greater range of precipitation measurements.

The second plot, a stacked area plot, also compares the precipitation observed in Alabama and Indiana for each year from 2017 to 2021. Again, it is evident that Indiana experiences less rainfall overall, with lower variability. Bar plots and stacked area plots offer unique own benefits. Bar plots are simple and straightforward to interpret, are suitable for both categorical and continuous response data, and illustrate comparisons well. Stacked area plots are particularly useful for visualizing temporal patterns over continuous intervals, provide a cumulative view of the data, and feature smooth transitions. Both Figure 7 and Figure 8 are examples of plots which could offer underwriters a comparison of the precipitation in two states in recent years.

Underwriters would also likely benefit from access to tables. The table below depicts the months with the highest recorded precipitation measurements out of all the states from 2017 to 2021.

Highest Precipitation Months Across all States from 2017 to 2021

Total Precipitation, in inches

State	Month	Year	Precipitation
Georgia	April	2020	1,203.38
Missouri	April	2017	1,064.48
Texas	April	2019	1,041.52
Texas	April	2017	781.66
Virginia	April	2020	765.87
Texas	August	2017	2,021.36
Virginia	August	2020	1,131.17
Georgia	August	2020	1,093.32
Georgia	August	2021	1,073.99
Georgia	August	2018	888.37
Georgia	December	2018	1,560.62

Table 1.

In the table above, the state with the highest recorded precipitation month measurement was Georgia. Georgia experienced a striking 1,203.38 inches of rain in April of 2020. Second was Missouri, which experienced 1,064.48 inches of rain in April of 2017. Overall, it seems that April and August tend to be the rainiest months, while Georgia, Virginia, and Texas may be some of the highest precipitation states. Tables like the above might be helpful for underwriters to get a sense of past data. Other helpful tables might, for example, display the catastrophic events which incurred the most damage in 2021 in a particular state. Overall, it is important for underwriters to have access to the most information possible when they are making decisions.

In addition to historical data, executives and underwriters need to see predictions of future climate conditions and future claims to help inform their decisions. This is where catastrophic modeling comes into play. As climate change progresses, it is increasingly important for data scientists and actuaries to collaborate effectively, and to communicate their data analysis and predictions clearly to the rest of the company.

There are many forms of models that data scientists may use. Certain tech companies offer an interface that assists with the implementation of a cat modeling system within an insurance company. For this project, I developed a multiple linear regression model to predict incurred property loss based on the climate conditions of the past three years. In reality, insurance companies are likely to have a large complex system of many models, but this is an example of a simplified approach.

The multilinear linear model predicts the response variable, incurred loss, based on climate conditions such as hail, rain, and snow. Loss is one of the most important quantities to predict at an insurance company because making decisions to mitigate loss, and charging higher premiums in order to proportionally accommodate potential loss, is arguably the most important function of an insurance company. The predictors of loss in this model are weather conditions from past years, based on the well-studied reality that catastrophic conditions are responsible for massive destruction of property.

Before implementing the model, I performed some exploratory analysis on the climate variables and tested the assumptions of linear regression.

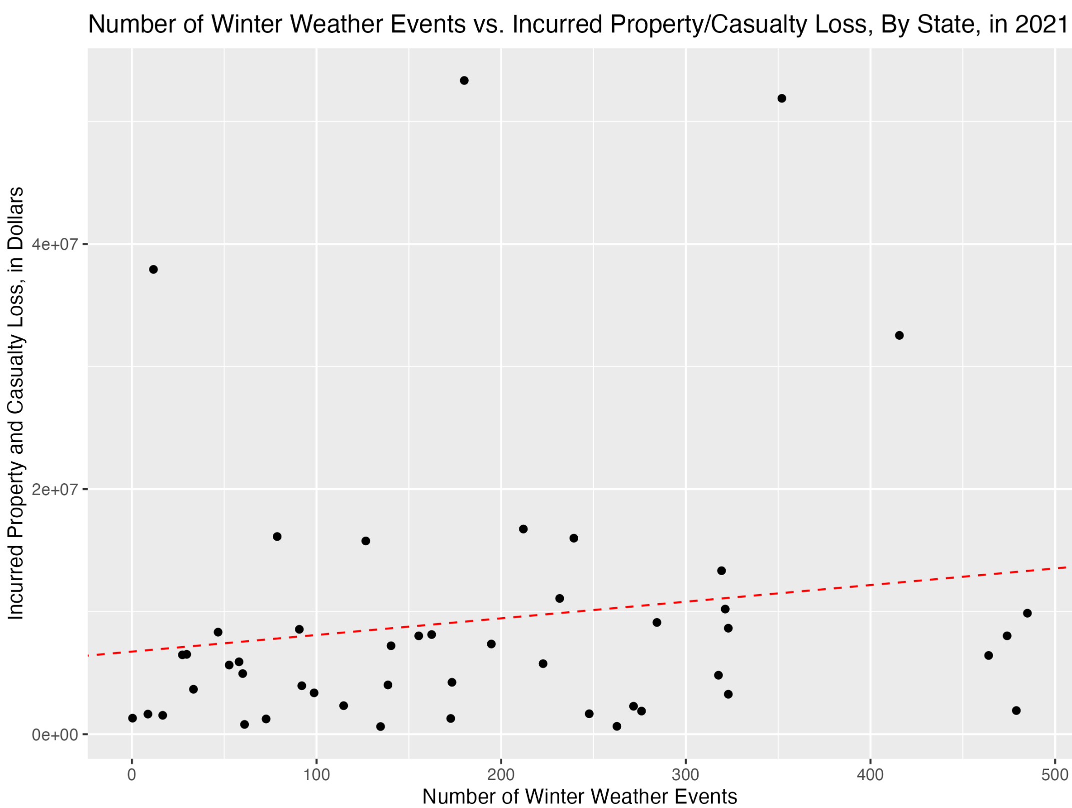

Figure 9.

As part of the exploratory analysis, Figure 9 presents a scatter plot depicting the number of winter weather events by state, along with the corresponding incurred property and casualty losses in dollars for the year 2021. Overall, there is a positive linear trend, suggesting that an increase in winter weather events is associated with an increase in property and casualty loss. This makes sense, because winter weather can be very destructive.

There were a few model specifications. A log transform was performed on the response variable to limit the effect of outliers. A few extreme outliers were also removed to improve the constancy of variance. Finally, forward step selection was performed to find the optimal combination of features. The final linear model predicted loss based on winter weather, hail, tornado, wind, floods, wildfire, and dust storms.

Overall, the linear model performed fairly well, as shown below.

```
Call:
lm(formula = log(TotalLossAmounts + 1) ~ Winter + Hail + Tornado +
    Wind + Flood + Wildfire + DustStorm, data = model_2021_df_ou)

Residuals:
     Min       1Q   Median       3Q      Max
-1.78050 -0.43403 -0.04878  0.45549  1.91626

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) 14.1239263  0.2730044  51.735  < 2e-16 ***
Winter       0.0018213  0.0011220   1.623  0.11240
Hail        -0.0029463  0.0010561  -2.790  0.00804 **
Tornado      0.0097607  0.0064316   1.518  0.13698
Wind         0.0021607  0.0006759   3.197  0.00272 **
Flood        0.0002270  0.0012093   0.188  0.85208
Wildfire     0.0037758  0.0018468   2.045  0.04752 *
DustStorm    0.0469755  0.0236528   1.986  0.05391 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.8234 on 40 degrees of freedom
Multiple R-squared:  0.4968,    Adjusted R-squared:  0.4087
F-statistic: 5.641 on 7 and 40 DF,  p-value: 0.0001412
```

As shown above, the statistically significant variables are hail, wind, and wildfire, which are associated with respective p-values of 0.008, 0.002, and 0.04. Dust storms approached significance, with a p-value of 0.053. Upon reversing the log transformation, the coefficient for wind was 0.002, suggesting that a 1,000-unit increase in wind across a state's average over the last three years is linked to a 2 increase in property and casualty loss in that state in the following year. The coefficient for wildfire was 0.004, which suggests that there is also a positive relationship between wildfire and property loss. The coefficient for hail was -0.003, which suggests that there is a negative association between hail and property loss. This is very counterintuitive, so there are likely some confounders at play.

The adjusted R-squared of 0.4087 indicates that approximately 40.87 percent of the variability in the dependent variable is explained by the model, considering the number of predictors.

The F-statistic, with a value of 5.641, tests the overall significance of the regression model. The statistically significant p-value of 0.00014 suggests that the model is meaningful, and the predictors collectively have a significant effect on the dependent variable.

In summary, the model shows a moderate level of explanatory power (adjusted R-squared), and the F-statistic indicates that the model is statistically significant. Improvements in the model, such as using a response variable based on only property data alone, as opposed to property and casualty, is likely to improve the significance of the results. Once improved, this model can be used to make predictions of property loss in the future year, based on the past three years of weather conditions.

For an effective data system, companies must develop the best methods for communicating model predictions to employees. For instance, underwriters might have access to a map for which states are colored based on their respective level of predicted property loss. Predictions also play a role in the determination of the pricing model, which underwriters rely on to effectively charge premiums.

5 Discussion

Overall, the purpose of this project was to present the importance of effective data communication and present the main characteristics of an effective data system for a property insurance company. It is important to note that the optimal systems and decisions will vary by company. Though this report presents guidelines for the most crucial attributes of a data system, companies may choose their own methods within this overall foundation. Different companies focus on different primary goals. For instance, insurance companies which serve affluent clients, often residing in highly catastrophic zones such as coastal areas, may be more likely to decide to pull out of a state altogether, as opposed to simply limiting policies.

A company must also decide for itself what routines it plans to implement. As stated, regular updates are vital for executives, who benefit most from data on a mass scale. But based on their preferences, executives may request a monthly, weekly, or quarterly report on data.

Meanwhile, underwriters will always need access to granular data. However, there are different ways that a company may achieve this goal. There are different options for interfaces and software. As long as there is a thorough, comprehensible system that is regularly updated and allows users to navigate to find the data they need, the details are up to the company.

Regardless of specifics, the bottom line is that data systems are indispensable for property insurers. They are useful for driving optimal business decisions and ensuring that all employees are in alignment.

There are also some limitations of this project to note. For one, the visualizations in this report are based only on precipitation data. Insurance companies will, of course, want to develop visualization strategies for all different types of weather conditions. Future research may involve an exploration of visualizations for certain catastrophes, especially wildfires, which are much more difficult to quantify than precipitation.

Additionally, the model in this report uses a response variable based on incurred loss in both property and casualty insurance. Ideally, it would make more sense to have used property loss data, as property insurance is the main focus of this report. However, insurance data can be hard to come by, as it is not always publicly available. Fortunately, the model still produces useful findings despite this limitation. Future studies may develop models with a response variable based on property loss. Ultimately, insurance companies have access to all of their internal data, so they have a better opportunity to develop models based on all the information they need.

In general, the attributes of a functional data system in an insurance company which are presented in this report should serve as general guidelines. Companies should continue to experiment with what works best for them, and fine-tune their systems over time.

6 Conclusion

Climate change entails the increasing prevalence of wildfires, hurricanes, and other catastrophic events, along with extremes like drought and high temperatures. Over time, the top insurance companies, especially in property insurance, will be those equipped with the best models and systems of communication.

In this project, I introduced a comprehensive model for data communication that insurance companies can implement in the coming years. Several essential characteristics define an effective model, including the establishment of systems to communicate historical data and trends, present future projections, and cater to the needs of employees across all levels of the company.

Recognizing diverse needs among employees, I proposed various data communication methods and outlined some of the key concepts. For executive presentations, it is recommended to use visualizations that present mass data at once and provide clear insight, across large geographic regions. For underwriters, it is better to provide a reliable interface through which they can access data on past weather conditions and losses, as well as future projections. The system should enable them to toggle between regions, years, and catastrophic events, in order to facilitate informed decision-making based on their particular case. I also suggest the use of maps as visualizations, using color gradients to represent a range of values across different states. This increases awareness of particularly catastrophic areas or zones which are projected to experience significant loss in the coming years.

As time progresses, it is in the best interest of insurance companies to optimize their processes, experimenting with different systems to identify the most effective options. Future research in this field might involve comparing various CAT models to determine their accuracy or developing more streamlined platforms for data communication. Insurance companies can leverage their data not only for risk assessment and premium setting but also to encourage the next generation of homeowners to take proactive measures on their homes. These measures include installing wildfire-resistant panels, earthquake retrofitting, and incorporating flood-prevention mechanisms. It may not be entirely true in the context of climate change, but as they say in Norway, “There is no bad weather - only bad clothing.” Insurance is about being prepared.

7 Citations

“A Firm Foundation: How Insurance Supports the Economy.” Insurance Information Institute, Insurance Information Institute, Inc., Feb. 2023, `www.iii.org/publications/a-firm-foundation-how-insurance-supports-the-economy/a-50-state-commitment/incurred-losses-by-state`.

Albin, Joshua. ”The Effects of Catastrophes on Insurance and Reinsurance Markets.” JayScholar, JayScholar, 28 Apr. 2021, `http://jayscholar.etown.edu/cgi/viewcontent.cgi?article=1005&context=`

mathstu.

Bouwer, Laurens M. "Projections of Future Extreme Weather Losses Under Changes in Climate and Exposure." Wiley Online Library, John Wiley And Sons, Inc., 7 Sept. 2012, onlinelibrary.wiley.com/doi/10.1111/j.1539-6924.2012.01880.x.

Brooks, Kristopher J. "Another Major Insurer Is Halting New Policy Sales in California." CBS News, CBS Interactive, www.cbsnews.com/news/allstate-insurance-state-farm-california/. Accessed 1 Dec. 2023.

"Climate Div." National Centers for Environmental Information , https://www.ncei.noaa.gov/data/climdiv/. Accessed 27 Nov. 2023.

Diaz de Leon , Pedro. "Top Risk Factors Facing the Insurance Industry in 2023." Cherry Bekaert, 17 July 2023, www.cbh.com/guide/articles/top-risk-factors-facing-the-insurance-industry-in-2023/.

Hall, Mary. "What Affects Profit Margins in the Insurance Sector?" Investopedia, Investopedia, 30 Aug. 2023, www.investopedia.com/ask/answers/052515/what-usual-profit-margin-company-insurance-sector.asp#:~:text=Insurance%20companies%20generate%20revenue%20through,losses%20due%20to%20insurance%20claims.

Jones, Aled W., and Alexander Phillips. "Voluntary Business Engagement in Climate Change: A Study of the Climatewise Principles." Journal of Cleaner Production, Elsevier, 16 July 2016, www.sciencedirect.com/science/article/abs/pii/S0959652616309520?via%3Dihub.

Lemick, Dustin. "How Do Insurance Companies Make Money?" BriteCo, BriteCo Inc., 25 June 2023, brite.co/blog/jewelry-insurance/how-do-insurance-companies-make-money/#:~:text=Insurance%20companies%20make%20money%20primarily,partnering%20with%20agents%20and%20brokers.

Lin, Ning, and Eric Shullman. "Dealing with Hurricane Surge Flooding in a Changing Environment: Part I. Risk Assessment Considering Storm Climatology Change, Sea Level Rise, and Coastal Development." SpringerLink, Springer Berlin Heidelberg, 21 Feb. 2017, https://link.springer.com/article/10.1007/s00477-016-1377-5.

Marr, Bernard. "How Big Data Is Changing Insurance Forever." Forbes, Forbes Magazine, 16 Dec. 2015, www.forbes.com/sites/bernardmarr/2015/12/16/how-big-data-is-changing-the-insurance-industry-forever/?sh=109d9dc6289b.

Odening, Martin, and Zhiwei Shen. "Challenges of Insuring Weather Risk in Agriculture." Agricultural Finance Review, Emerald Group Publishing Limited, 1 July 2014, www.emerald.com/insight/content/doi/10.1108/AFR-11-2013-0039/full/html.

Saffron , O'Neill J., and Smith Nicholas . "Climate Change and Visual Imagery ." WIREs Wiley Interdisciplinary Reviews, John Wiley And Sons, Inc., 18 Oct. 2013, wires.onlinelibrary.wiley.com/doi/full/10.1002/wcc.249.

Son, Rackhun, et al. "Climate Diagnostics of the Extreme Floods in Peru During Early 2017 ." SpringerLink, Springer Berlin Heidelberg, 24 Dec. 2019, link.springer.com/article/10.1007/s00382-019-05038-y#:~:text=In%20early%202017%2C%20coastal%20northeast,%2C%202%2013%20April%202017).

"Storm Events." National Centers for Environmental Information , www.ncei.noaa.gov/pub/data/swdi/stormevents/csvfiles/. Accessed 27 Nov. 2023.

Swain, Mamata. "Redesigning Crop Insurance for Coping with Climate Change in India." ARF Journals, ARF India, 10 Apr. 2023, www.arfjournals.com/image/catalog/Journals%20Papers/IJAEB/2023/No%201%20(2023)/6_Mamata%20Swain.pdf.