The Charged Tempest: Assessing Hurricane Impacts on Power Grid Integrity and Social Consequences

Team Number: 16808
Modeling The Future Challenge 2024

April 2024

Contents

1 Executive Summary

Hurricanes, typhoons, and tropical cyclones are all natural disasters capable of causing extreme damage and disturbance to communities, businesses, and individual lives. There are many types of damages a hurricane could cause, such as human fatalities and casualties, property damages, power outages, infrastructure destructions, environmental impact, etc. Additionally, we can characterize the cause of damage of these types of disasters into two natural aspects, wind and flooding. This paper, given the limited availability of data, will specifically analyze the impact of wind damage on the power grid, which is a sector that can be more easily quantified. Our model utilizes two data sets, EAGLE-I and HURDAT2. The EAGLE-I dataset will serve as our major dependent variable as it provides information on the scale of a particular outage at a particular time in a particular county. On the other hand, the HURDAT2 dataset provides information on the path and Windfield information of a list of historical hurricanes to help us model and predict the path of an approaching hurricane. Throughout our modeling and risk analysis, we have identified an increasing frequency and severity of hurricanes in recent years, the difference in the strength of the power grid based on counties' response to a particular natural disaster (exp: Brazoria is more vulnerable than Harris), and the meteorological prediction of path of an incoming hurricane for immediate response from local governments.

Ultimately, our model will provide insights into both long-term and short-term recommendations to help counties and states (Texas) lessen the impact of hurricanes on their local power grid through methods such as Business Disruptions insurance, Encouragement of Individual power generation like portal power generators in a more vulnerable area, and transition to underground power lines.

2 Background

Hurricanes are one of nature's most powerful and destructive forces, capable of wreaking havoc on coastal regions, communities, economies, and ecosystems. Year after year, these storms leave trails of destruction for areas affected. In 2023, the total damage from Atlantic storms alone is estimated at around 4 billion USD.

Therefore, this project aims to further analyze this phenomenon, including both path predictions and impact calculations within affected areas. By examining the meteorological factors driving hurricane formation and developing effective trajectory and vulnerability models we assess the consequences of windstorms on businesses and individuals within the effective areas. In particular, we will focus on the issue of damage to electrical grids, an integral component of industries, public institutions, and modern-day life.

Texas is extremely susceptible to hurricanes due to its proximity to the Gulf of Mexico. Tropical storms have significantly affected the power grid, causing widespread outages and substantial damage. For instance, Hurricane Nicholas in 2021 left over 500,000 residents without power as it swept across the central and east Texas coast. Similarly, according to the Electric Reliability Council of Texas (ERCOT), Hurricane Harvey caused severe damage and led to the loss of 10,000 megawatts of electricity. Several power plants went offline due to the flooding interfering with fuel supply. High voltage transmission lines near the coast were damaged by winds exceeding 130 miles per hour. This emphasizes the importance of vital infrastructure such as hospitals, water treatment facilities, construction, agriculture, education, and other utilities that may be vulnerable to downed power, especially those living in these buildings. The World Economic Forum highlighted the vulnerability of Texas' power grid, deeming it as a wake-up call for the necessity of constructing a cleaner, more resilient, and improved grid that can better withstand extreme weather conditions.

We developed a two-part methodology. We analyzed vulnerability by county using historical hurricane outage data. Our second section focuses on creating a real-time hurricane Bidirectional Long Short-Term Memory Artificial Intelligence path predictor to forecast hurricane track, wind speed, and wind radius distribution. This will allow us to provide real-time recommendations for counties that may be hit, especially for those that are also vulnerable, calculated from the first section of our methodology.

The following sections will research the principles of hurricane dynamics, explore the advancements in tracking and prediction technologies, and investigate the implications of power outages caused by these storms on coastal communities. Through several comprehensive case studies on significant hurricanes, our work will explore strategies for enhancing our resilience against the challenges caused by hurricanes and subsequent power outages. This includes the use of insurance in assisting recovery efforts, building more resilient infrastructure, and adapting various tools. Ultimately, our aim is to provide insights and recommendations that will inform decision-makers, empower communities, and mitigate the risks associated with these natural disasters moving forward.

Specifically, this paper will aim to answer three essential questions:

1. Which counties are most prone to hurricane influence?
2. What causes some counties to experience more severe outages than others during hurricane events?
3. How can counties improve resilience against power outages and reduce potential damages in future events?

Our paper also provides the following recommendations:

1. Increase the number of underground cables to mitigate the chance of damage
2. Buy portable power generators
3. Business interruption insurance and windstorm coverage insurance

3 Data Methodology

3.1 EAGLE-I Power Outage Data

The Environment for the Analysis of Geo-Located Energy Information (EAGLE-ITM) is maintained at Oak Ridge National Laboratory (ORNL) for the Department of Energy (DOE). EAGLE-I is DOE's operational and scalable data and information platform for real-time wide-area situational awareness of the energy sector, providing a centralized platform for monitoring power distribution outages for over 146 million customers: just over 92% coverage of US and territories.

The dataset, collected by the EAGLE-I program at ORNL, includes eight years of power outage information from 2014 to 2022, at 15-minute intervals, at the county level. With its contemporary temporal range and county-wise spatial coverage, this dataset served as our primary source to quantify the impact, helping formulate both independent and dependent variables.

Variable	Variable Name	Description
FIPS code	Fips_code	The FIPS code of the county in which the power outages occurred.
County Name	County	The county name in which the power outages occurred
State Name	State	The state in which the power outage occurred
Outages	Sum	The number of customers without power.
Datetime	Run_start_time	Date & timestamp in GMT in the format "MM/DD/YY 00:00".

Additional information for each observation: EAGLE-I collects power outage information from all covered utilities at 15-minute intervals, and this timestamp marks the beginning of the collection run.

3.2 HURDAT2

To inform our research, we accessed the hurricane pathing data from the National Oceanic and Atmospheric Administration's (NOAA) comprehensive hurricane database, HURDAT2. To enhance the predictive capabilities of our Long Short-Term Memory (LSTM) model, we extracted and analyzed a set of critical features from the HURDAT dataset. This dataset tracks hurricane paths separated by 6-hour time steps for several hundred hurricane instances, where each hurricane instance can be regarded and grouped as a separate time series. These included the latitude and longitude to pinpoint the hurricane's center, the entry time for temporal context, maximum wind speed and rate of change for intensity analysis, and minimum central pressure for understanding the system's strength. Additionally, we incorporated zonal (east-west) and meridional (north-south) speeds to track directional movement. We better segmented the hurricane radius based on wind speed differences to categorize the storm's structure and potential impact.

Variable	Variable Name	Description
Longitude	lon	Geospatial Coordinate Variable, given in data
Latitude	lat	Geospatial Coordinate Variable, given in data
Entry Time	entry_time	Time (updated every 6 hours), given in data
Maximum Wind Speed	max_winds	Maximum wind speed recorded per entry, given in data
Max Wind Speed Change	delta_wind	Change in max. wind speed since last record calculated as $\text{max\_wind}_t - \text{max\_wind}_{t-1}$
Minimum Wind Pressure	min_pressure	Minimum wind pressure within an entry time, given in data
Zonal speed	zonal_speed	Latitudinal rate of change in hours, calculated as $\frac{lat_t - lat_{t-1}}{1 \text{ timestep}}$
Meridonal speed	meridonal_speed	Longitudinal rate of change in hours, calculated as $\frac{lon_t - lon_{t-1}}{1 \text{ timestep}}$
Wind radius descriptor 1	34kt_ne	Radius of wind speed 34 knots at the northeast quadrant
Wind radius descriptor 2	34kt_se	Radius of wind speed 34 knots at the southeast quadrant
Wind radius descriptor 3	34kt_nw	Radius of wind speed 34 knots at the northwest quadrant
Wind radius descriptor 4	34kt_sw	Radius of wind speed 34 knots at the southwest quadrant
Wind radius descriptor 5	50kt_ne	Radius of wind speed 50 knots at the northeast quadrant
Wind radius descriptor 6	50kt_se	Radius of wind speed 50 knots at the southeast quadrant
Wind radius descriptor 7	50kt_nw	Radius of wind speed 50 knots at the northwest quadrant
Wind radius descriptor 8	50kt_sw	Radius of wind speed 50 knots at the southwest quadrant
Wind radius descriptor 9	64kt_ne	Radius of wind speed 64 knots at the northeast quadrant
Wind radius descriptor 10	64kt_se	Radius of wind speed 64 knots at the southeast quadrant
Wind radius descriptor 11	64kt_nw	Radius of wind speed 64 knots at the northwest quadrant
Wind radius descriptor 12	64kt_sw	Radius of wind speed 64 knots at the southwest quadrant

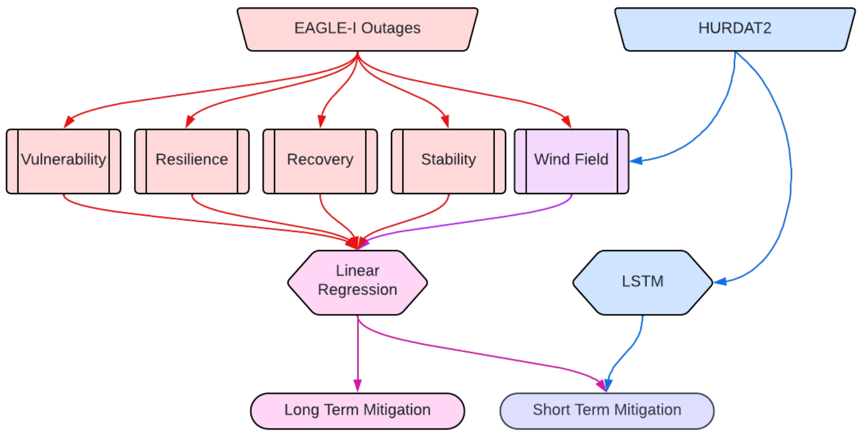
Variables highlighted in blue were also outputted as predictions with our model. We did not include the 64kt variables as an output variable because most of the samples had a value of 0 for this. Scaled data using RobustScaler.

4 Mathematics Methodology

4.1 Architecture Overview

Our model architecture integrates two principal data sources to assess the impact of hurricanes on power grid integrity and develop mitigation strategies. The initial input is derived from the EAGLE-I Outages database, encompassing metrics of the power grid's Vulnerability, Resilience, Recovery, and Wind Field—the latter also informed by the HURDAT2 hurricane tracking data from NOAA. To process this input effectively, we applied a dual-analytic approach.

We first used Linear Regression to determine the relationship between the grid's defensive attributes and the severity of disruptions caused by hurricanes. This statistical method (via coefficients) allowed us to quantify these relationships and form a framework for long-term mitigation strategies to enhance the grid's strength and integrity for future events. Concurrently, we utilized an LSTM model, which is adept at handling time-series data, to predict more immediate, short-term impacts on the grid based on the dynamic hurricane data derived from HURDAT2.



(a) Model Architecture

The culmination of our analysis is twofold: we generate long-term mitigation strategies to strengthen the power grid systematically, and we devise short-term mitigation tactics for prompt, responsive actions during hurricane events. The integration of Linear Regression and LSTM models enables us to construct a robust, predictive mechanism that guides both immediate responses and informs infrastructural and policy decisions for sustained grid integrity. This comprehensive approach aims to minimize power outages and maintain grid stability in the face of hurricane challenges.

We used linear regression due to its time and resource efficiency and strong interpretability. Interpretability is especially important for this section because we must fully understand the inherent processes within our vulnerability and resilience analysis. We used a Long Short-Term Memory AI model because of its accuracy and precision. It was initially intended as a Large Language Model; however, it's efficacy for sequential data can be transferred towards time series analysis, given the proper adaptations. These models can effectively capture complex relations such as hurricane pathing and wind fields. Furthermore, its black-box nature allows us to quickly design a successful model without the intricacies and convolutions behind the math of hurricanes, which include fluid dynamics, specific wind fields, etc. It is more difficult to interpret because of factors such as its hyper-abstract weight distribution and multidimensional vector calculations; however, the pathing area of this project is much more results-focused, depending on fast and accurate predictions of hurricanes, which allows us to use a model with higher accuracy but lower interpretability.

4.2 Assumption and Justification

1. **Assumption:** We assume that the spatial resolution of county-level power outage data and the temporal resolution of hurricane data (updated every 6 hours) accurately correlate hurricane char-

acteristics with power outages.

Justification: EAGLE-I provides detailed outage data at 15-minute intervals, offering precise timing of outages, while HURDAT2's 6-hourly updates on hurricane parameters allow for matching with outage events, it reasonable to assume that most significant changes in hurricane behavior and its effect on the power grid can be captured within these intervals.

2. **Assumption:** We assume communities/regions within a county experience uniform damage and maintain consistent infrastructure repair and power restoration capabilities after hurricanes.

   Justification: Assuming uniform damage and consistent capabilities for infrastructure repair and power restoration within a county after hurricanes simplifies modeling and planning by providing a computationally manageable framework and a baseline for analysis, given the often county-level aggregated data. This assumption also helps emergency planning by preparing for a baseline scenario, ensuring resource allocation can meet at least a generalized need.

3. **Assumption:** We assume that the primary cause of power outages during hurricane events is the hurricane, excluding other potential causes like human error or unrelated weather events.

   Justification: The assumption that hurricanes are the primary cause of power outages during their events simplifies the analysis by focusing on the most significant and direct impact on infrastructure rather than considering less prevalent causes like human error or other weather phenomena. This approach allows for targeted modeling and response strategies to the overwhelming effects of hurricanes on the power grid.

4. **Assumption:** We assume that our vulnerability statistic, which is calculated outages, can be representative of overall property damages expected from hurricanes.

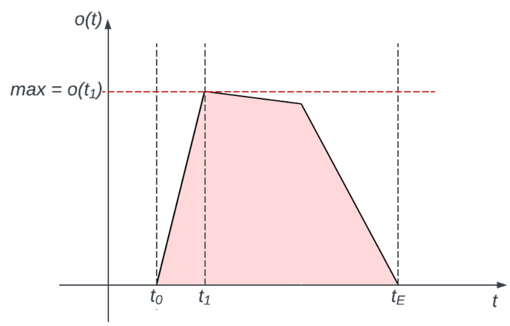
   Justification: Power outages are caused by the severity of a storm, which is also causes building damage and property loss. Outages will cause buildings to be more prone to further destruction, such as increasing the risk of damage from heat, moisture, and fire.

4.3 Hurricane Outages Characterization

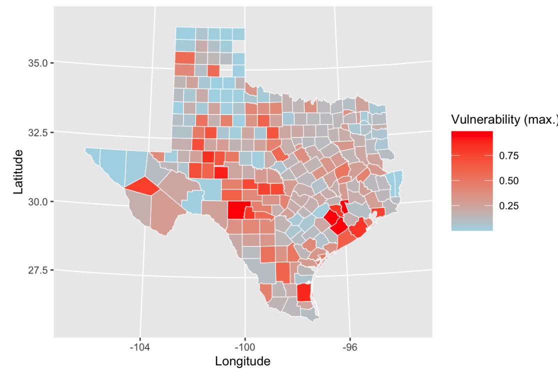
We characterize a hurricane event and its impact on outages through two phases: Damage Propagation and Infrastructure Restoration. Within this framework, we have developed three matrices to assess the power grid's strength and integrity at the local/county level: Vulnerability, Resilience, and Recovery. Additionally, we recognize that no historical hurricanes have affected every single county in Texas. Consequently, we have chosen the February 2021 North American Winter Storm, which impacted almost all counties in Texas, as our primary source for intrastate reference for all three indices.

4.3.1 Vulnerability

The vulnerability index is quantitatively established by the metric $o(t_1)$ illustrated in the figure below as the maximum value on the outage axis, signifying the most substantial number of power outages scaled to the local population at time t_1. This point represents the apex of the vulnerability phase, indicating the initial brunt of the disaster where the power grid's performance dips most significantly. By benchmarking the height of this peak, the vulnerability index effectively measures the immediate impact of the hurricane on the electrical infrastructure's capacity to serve the affected population.



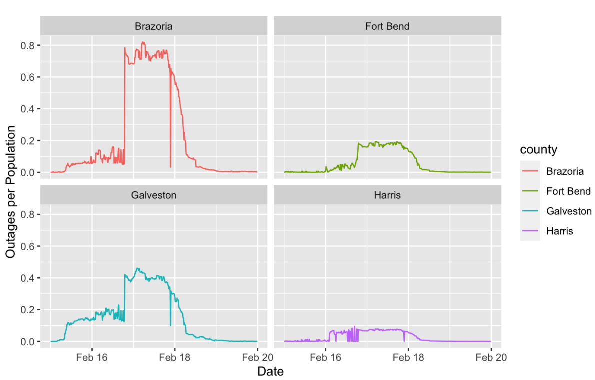
(b) Vulnerability Definition



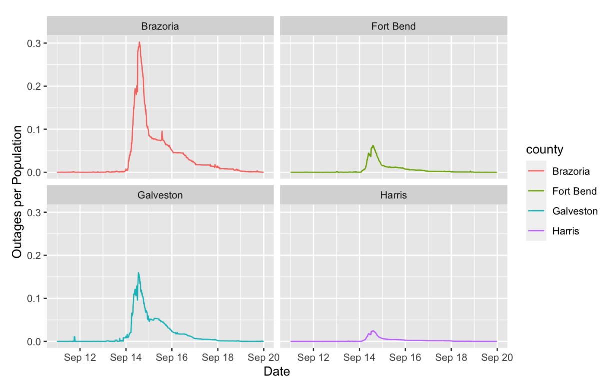
(c) Vulnerability index

The following figures from the February 2021 North American Winter Storm and Hurricane Nicholas exemplified clear disparities in power grid strength among the different counties in Texas. The variation in the height of the peaks of outages per population during these events indicates a difference in the level of vulnerability of each county's power grid.

For example, during the winter storm, Brazoria and Galveston Counties experienced higher peak outages per population, pointing to a greater vulnerability or weaker grid strength compared to Fort Bend and Harris Counties, which exhibited lower peaks. To prove this pattern also applies to other extreme natural disasters, during Hurricane Nicholas, Brazoria County's peak suggests a higher rate of outages relative to its population compared to the other counties, reinforcing the notion of inconsistent power grid integrity across different regions.



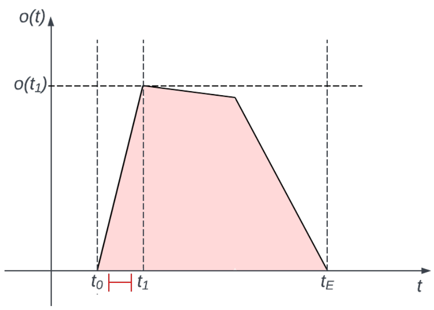
(d) 02/2021 Winter Storm



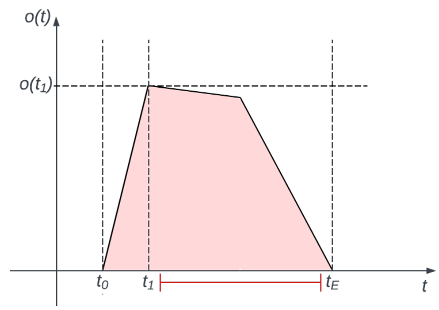
(e) 09/2021 Hurricane Nicholas

4.3.2 Resilience, Recovery, & Stability

The resilience index is characterized by the duration from the initial impact of the hurricane at time t_0 to the point of maximum outages t_1 as indicated by the horizontal distance between these two points on the time axis. This index reflects the grid's ability to withstand escalating damages before reaching a critical threshold of outages. The recovery index, in contrast, is represented by the time span from t_1 when the outages peak, to t_E the moment when the system stabilizes and returns to minimal outage levels. This index evaluates the speed and capability of the grid's restoration processes following the peak of the hurricane's impact. The stability index represents the maximum slope of hurricane outages between t_0 and t_1. It tracks the fastest rate at which outages happen, thus instead of showing how long a county can withstand damages like resilience, it shows how bad the worst part of a hurricane is on a county.

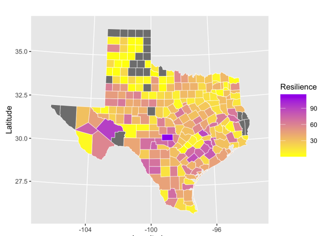


(f) Resilience Definition

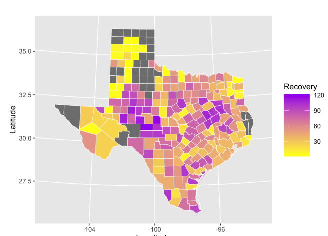


(g) Recovery Definition

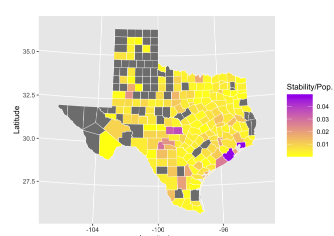
Again, based on the 2021 North American Winter Storm which covers most of all counties in Texas, we can characterize both indices based on each county's response to this natural disaster. The resilience map highlights the differential robustness of power grids across counties, with some showing high resilience scores, suggesting a strong capacity to withstand the initial shock of the storm. In contrast, the recovery map provides insight into the speed and effectiveness with which counties could restore power. This time, counties with lower recovery scores indicate a more rapid return to normalcy post-disruption, underscoring the effectiveness of their restoration strategies and infrastructure resilience. Although there is a weak statistical correlation between the two indices, in general, based on the visualization, counties that cannot withstand damages gradually (lighter color in resilience) typically also have a longer time in restoration (darker color in recovery). We can see that stability is more sporadically placed geographically



(h) Resilience Spatial index



(i) Recovery Spatial index

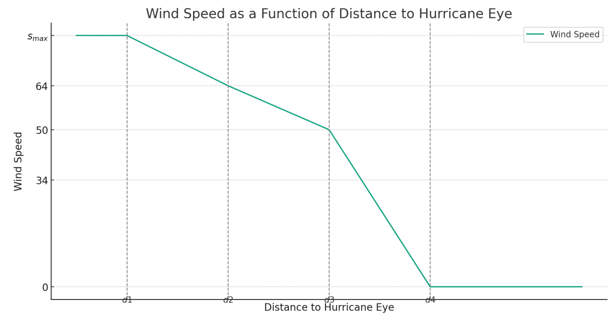


(j) Stability Spatial index

4.3.3 Wind Field

Using the variable s_{max} to represent the hurricane's maximum wind speed. We use d_1,d_2,d_3, and d_4 to represent the distance when the hurricane is at s_{max},64,50, and 34 mph respectively. Thus, the following piece-wise linear function is used to approximate the hurricane's wind speed at distance x.

$$\begin{cases} s_{max} & 0 \le x \le d_1 \\ \frac{64-s_{max}}{d_2-d_1}(x-d_1)+s_{max} & d_1 \le x \le d_2 \\ \frac{-14}{d_3-d_2}(x-d_2)+64 & d_2 \le x \le d_3 \\ \frac{-16}{d_4-d_3}(x-d_3)+50 & d_3 \le x \le d_4 \\ 0 & d_4 \le x \end{cases}$$



4.3.4 Multivariable Linear Regression

To maximize the statistical significance of our linear regression, we tested many permutations of multiplication of our features. We used Python to combine and test thousands of such permutations to see how the different features affected each other. We found that by multiplying vulnerability, population, and wind field, we were able to maintain both a high r-squared and adjusted r-squared value. We found that this regression has a very low p-value for the combined features and stability. Notice that both population and the vulnerability index have to do with the size and growth of a county. This indicates that higher developed areas are at significantly higher risk, and should focus on building infrastructure.

Given all the provided information above, we have the following estimated regression equation by fitting and testing many separate permutations of the variables to minimize variance while maintaining comprehensible results:

	Estimate	Std. Error	t value	Pr(> $\|t\|$)
(Intercept)	9.979e-01	1.771e+02	0.006	0.99551
combine_factors	5.872e-03	1.341e-04	43.786	< 2e-16 ***
stability	2.011e-01	5.069e-02	3.967	7.68e-05 ***
recover	1.257e-03	4.246e-04	2.960	0.00313 **

Table 1: Coefficients

Residual standard error	4446
Degrees of freedom	1269
Multiple R-squared	0.6506
Adjusted R-squared	0.6497
F-statistic	787.6
p-value	< 2.2e-16

Table 2: Model Summary

The fit of this regression model, which predicts hurricane outages based on vulnerability, resilience, recovery, stability, population, and wind field influence, can be evaluated using several key statistical measures: Multiple R-squared, Adjusted R-squared, F-statistic, and the p-value of the F-statistic.

1. The Multiple R-squared value is 0.6506, indicating that the model explains approximately 65% of the variance in hurricane outages. Correspondingly, the Adjusted R-squared value, which accounts for the number of independent variables and their degrees of freedom, is 0.6497. Both measures suggest that the model explains well over half of the variance. Considering the complexity of the modeled phenomena, this is a significant portion of the data.

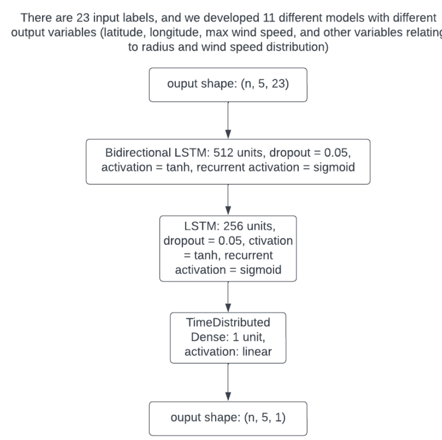
2. The F-statistic is 787.6, with a p-value of less than 2.2e-16. This significant F-statistic and the extremely low p-value indicate that the model is statistically significant, meaning a meaningful relationship exists between the independent and dependent variables. The low p-value strongly suggests that the observed relationships are not due to chance.

While the Multiple R-squared value indicates that the model explains a modest portion of the variance in hurricane outages, the statistical significance evidenced by the F-statistic and its p-value supports the conclusion that the predictors chosen for the model have a justified and meaningful relationship with the outcome.

4.4 Pathing Characterization

Our pathing model is a bidirectional time-distributed long short-term memory model.

The input to the model consists of tabular data sequences of length 5. Each sequence represents a temporal sequence of observations in a time series context. Each sequence represents a 5-day observation period. The Bidirectional LSTM layer is the core component of the architecture. It processes the input sequences in two directions: forward and backward. This is achieved by duplicating the LSTM layer and processing the input sequence with one LSTM layer in the original order and another LSTM layer in the reverse order. By doing so, the Bidirectional LSTM layer captures information from past and future time steps.

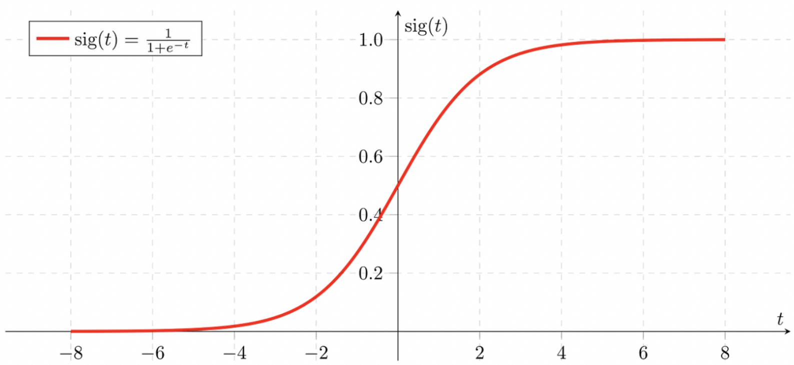


The forward and reverse LSTMs are combined with a concatenation function. The next layer in this model is a standard LSTM model with 256 units and a dropout rate of 0.05. The LSTM layer feeds into the Time Distributed Dense function. The Time-Distributed layer is applied after the Bidirectional LSTM layer. It allows the model to apply a dense layer to each time step of the LSTM output independently. This is particularly useful when the task involves making predictions at each time step of the sequence, such as sequence labeling or sequence generation tasks. The dense layer applied is a linear function of one neuron.

Our entire pathing model is a Bidirectional LSTM layer of 512 units, a standard LSTM layer of 256 units, and finally a Time Distributed fully connected layer with only one unit. We used a Mean Squared Error loss metric. We used the Adaptive Momentum Estimation optimizer with learning rate 0.01 and weight decay of 0.0001 to reduce overfitting. This optimizer combines AdaGrad and RMSProp to reduce convergence time and increase model performance.

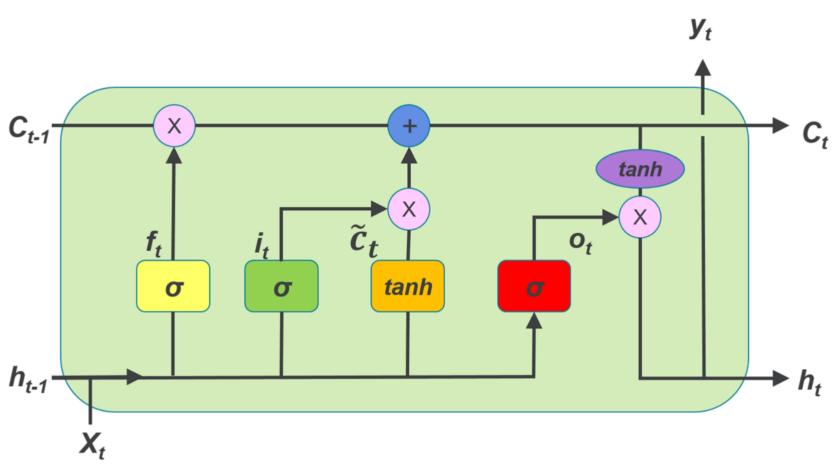
4.4.1 LSTM

The Long-Short-Term Memory Machine (LSTM) is a type of recurrent neural network (RNN). RNNs have a self-looping recurrent flow to preserve previous inputs for future predictions. However, this makes the model extremely sensitive to vanishing and exploding gradients when training because the self-looping function causes gradients to be affected exponentially. To fix this, we used an LSTM model, which resolves this issue with gating mechanisms to control the flow of information and weights. LSTM models consist of three gates: input, forget, and output. The LSTM gate activation functions are sigmoid.



(k) Sigmoid

These gates are what allow it to represent sequential data robustly. It has displayed great ability in Natural Language Processing; however, tabular data can also be sequential when each sample is a group of rows. This allows the model to reference time as another entity by separating each time stamp.



(l) LSTM

Here's a breakdown of the key components and variables in an LSTM unit:

1. **Cell State (C_t)**: The cell state is the core component of the LSTM, carrying information throughout the processing of the sequence. It undergoes minimal changes at each step, allowing the network to add or remove information to the cell state in a controlled manner through gates.

2. **Input Gate (i_t)**: The input gate controls how much of the new information from the current input (x_t) and the previous output (h_{t-1}) is added to the cell state. A sigmoid layer determines which values will be updated.

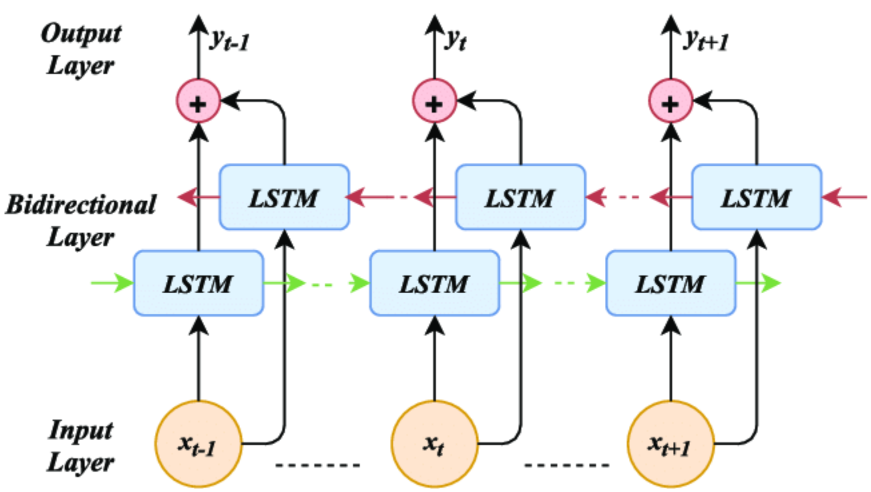
3. **Forget Gate (f_t)**: This gate decides which information is discarded from the cell state. A sigmoid layer examines h_{t-1} and x_t, outputting numbers between 0 and 1 for each number in the cell state C_{t-1}, where 0 means "completely forget" and 1 means "completely retain."

4. **Output Gate (o_t)**: The output gate determines what the next hidden state (h_t) should be. This hidden state contains information on previous inputs, filtered by the output gate. A sigmoid layer decides which parts of the cell state are outputted, and a tanh layer produces a vector of new candidate values, C_t, that could be added to the state.

5. **Candidate Cell State ($\tilde{C}_t$)**: This vector represents new candidate values that could be added to the state, calculated using a tanh layer. It is combined with the input gate to update the cell state.

6. **Hidden State (h_t)**: The hidden state is the output of the LSTM unit, derived from the cell state but filtered through the output gate. It can be used for predictions or passed to the next time step of the LSTM.

4.4.2 Bidirectional LSTM



(m) Bidirectional LSTM

1. **Input Layer**: Inputs are fed into both LSTM layers in sequence. Each input at time step t is denoted as x_t.

2. **Bidirectional Layer**:

   (a) **Forward LSTM**: Processes the sequence from the beginning to the end to capture information in the forward direction. At each time step t, the forward LSTM takes input x_t and the hidden state from the previous time step $h_{t-1}^{forward}$, and outputs a hidden state $h_t^{forward}$.

   (b) **Backward LSTM**: Processes the sequence from the end to the beginning to capture information in the backward direction. At each time step t, the backward LSTM takes input x_t and the hidden state from the following time step $h_{t+1}^{backward}$, and outputs a hidden state $h_t^{backward}$.

   (c) **Output Layer**:
   The outputs of the forward and backward LSTM layers at time step t are combined, often by summation or concatenation, to form the final output y_t. This output considers information from both past ($h_t^{forward}$) and future ($h_t^{backward}$) states relative to the current time step.
   The combined output y_t is input for the final Time Distributed Dense output layer. Incorporating forward and backward information allows BiLSTMs to make more informed decisions than standard unidirectional LSTMs. This is particularly useful in tasks that include sequential data in both directions. This combines the power of LSTMs with bidirectional processing to enhance the model's capability of capturing both the past and future context of the input. An example for why this is important can be explained by a textual example. If one applies a simple text analysis LSTM on the sentence: "Apple is something I like to eat", the model will not understand what "Apple" means, due to lack of context from the words after it. A BiLSTM will have a forward pass that lacks context for "Apple", but the backward pass calculates what the forward pass missed, and is concatenated.

4.5 Time Distributed Dense

Our final layer of the model is a Timer Distributed Dense layer. This layer combines the Time Distributed and Dense layers, allowing our Dense function to be applied across all time steps.

1. Sequential Data Handling: Each time step can be treated as an independent data point in sequences. The Time Distributed Dense Layer allows you to apply the same dense layer to each time step of the sequence separately, enabling the network to learn temporal patterns and dependencies effectively.

2. Parameter Sharing: By using the same set of parameters (weights and biases) for each time step, the layer encourages parameter sharing across time, which can help the network generalize better and learn from the temporal dynamics of the data.

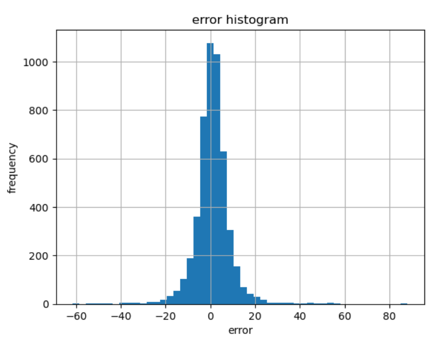
3. Output: The output of each time step passes through the dense layer independently, producing a sequence of output vectors with the same length as the input sequence. This layer allows us to output a 5-day forecast of the hurricane parameters.

4.6 LSTM Results

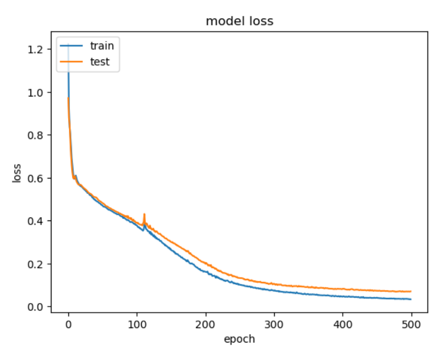
Our pathing model produces exceptional results for all intended outputs. All loss values are relatively low, and the model generalizes well across all sets.

Model Output	Training Loss	Validation Loss	Testing Loss
Max wind	0.0197	0.0744	0.0847
Lat	0.0052	0.0198	0.028
Lon	0.0048	0.0144	0.04
34kt_ne	0.023	0.0198	0.015
34kt_se	0.0048	0.0144	0.02
34kt_nw	0.03	0.066	0.054
34kt_sw	0.0431	0.1378	0.12
54kt_nw	0.0296	0.0652	0.052
54kt_ne	0.0397	0.161	0.08
54kt_se	0.0471	0.1608	0.17
54kt_nw	0.0501	0.1929	0.182
54kt_sw	0.0744	0.5217	0.458

The wind error distribution is fairly tight centered around 0, and the training graph appears extremely healthy.



Wind Error Histogram



Wind Loss Training Graph

Statistic	Value
count	432.000000
mean	11.666988
std	11.117696
min	0.005133
25%	3.472953
50%	7.876393
75%	16.408729
max	52.664535

Wind Error statistics

25%, 50%, and 75% define Q1, Q2, and Q3 respectively for calculated Mean Squared Error between model wind forecasts and the associated true values.

4.7 Sensitivity Analysis

We can conduct a sensitivity analysis based on the coefficients produced from the linear regression model. This analysis will allow us to assess and evaluate the impact of changing different factors. Using the "big four" (Brazoria, Fort Bend, Galveston, and Harris) as examples, we evaluate the impact of changing the other three independent variables by +10%. We didn't include the -10% sensitivity evaluation because the nature of linear regression will output the same number in terms of absolute value.

Combined factors represent the columns *Vulnerability*, *Population*, and *Wind Field*, because these variables are multiplied together in the equation, and hence when individually increased by 10% results in the same value as each other.

Variable	Brazoria	Fort Bend	Galveston	Harris
Combined Factors	9.5750%	9.8763%	9.6305%	9.4789%
Stability	0.3254%	0.0791%	0.2566%	0.4341%
Recover	0.0996%	0.0444%	0.1127%	0.0869%
All	32.1180%	32.8141%	32.2463%	31.8961%

The following are observations of this analysis:

1. The combined factors (Vulnerability, Population, Wind Field) have a positive coefficient indicating an increase in any of the three factors would lead to an increase in outages. Additionally, the Wind Field factor is designated to be similar among the three counties due to their close geographical proximities. Therefore, the discrepancies between counties are likely attributed to a combination of differences in population (fixed for short-term) and power grid vulnerability.

2. Harris County experienced much greater increases upon changes in the stability index in comparison to other counties. Correspondingly, an improvement (or decrease) in the stability factor would benefit Harris County the most. This makes intuitive sense because Stability is a measurement of the maximum positive slope of outages during a hurricane event. Therefore, as a population hub, that slope would naturally be higher and any shifts would bring about the most change to Harris.

3. When modifying all the factors collectively, all counties react similarly in terms of direction but not in terms of magnitude, which likely is indicative of their potential for improvement.

4.8 Strengths and Weaknesses

1. This paper integrates various data sources (EAGLE-I Power Outage Data and HURDAT2) and a combination of linear regression and LSTM models for predicting and analyzing hurricane-induced power outages. This allows our paper to offer insights into both short-term and long-term mitigation strategies.

2. This paper uniquely combines meteorological data with engineering evaluations (vulnerability, resilience, recovery, stability) and data science principles to address a critical societal challenge with power outages. It demonstrated a much more complete and in a sense realistic strategy to tackle complex issues.

3. The main weakness of this dataset comes from insufficient data. While the ideal research paper on hurricanes in America would be about Florida, the data is simply not reliable and accessible enough. Additionally, many of the choices made in the paper, such as choosing to track power outages, were made to ensure we had the best data available. This doesn't take away from the results of the paper, rather it shapes the way the paper was made.

5 Risk Analysis

5.1 Risk Overview

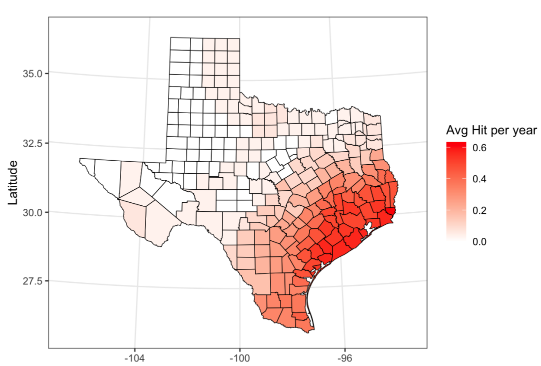
In the context of power outages, risk analysis of hurricanes examines the potential risks associated with hurricane-induced disruptions to the power supply. Frequency refers to how often hurricanes occur that lead to power outages, highlighting the regularity of these disruptions. Severity addresses the potential impact or harm of hurricane-induced power outages, such as economic losses, health and safety risks, and the effect on critical infrastructure and services like power plants and transmission lines. The distribution of risks assesses how the risks of power outages due to hurricanes are spread across different regions, identifying areas more prone to such disruptions.

The frequency and severity of risks posed by hurricanes leading to power outages are defined as follows. For each category, we also analyze the distribution of these risks over time and across locations and how historical trends might influence future predictions.

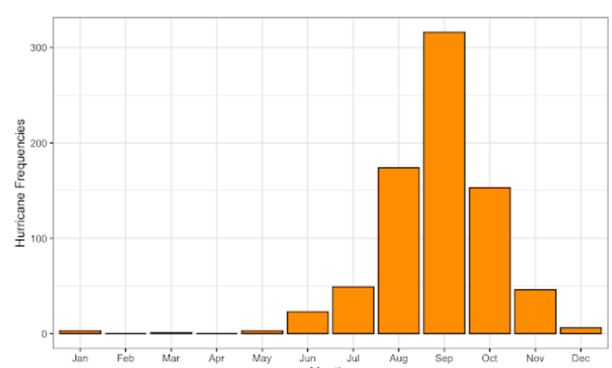
5.2 Frequency

Space: the following heatmap displayed the frequency of hurricane hits across the state of Texas. The color intensity represents the number of hurricanes that have hit each county. The scale on the right indicates that darker shades of red correspond to a higher frequency of hurricane hits, with the darkest shade representing 15 or more hits. The concentration of darker red shades along the coastline, particularly on the southeastern side of Texas, suggests that these areas are more frequently affected by hurricanes. Counties that are inland generally show a lighter shade, indicating fewer hurricane hits, which is consistent with hurricanes losing strength as they move over land.

Time: the bar chart representing hurricane frequencies by month. This chart shows a clear seasonal pattern, with the peak of hurricane occurrences in September, followed by August and October. The months of May through November show higher hurricane activity, which corresponds to the Atlantic hurricane season, while the months from December through April show minimal activity.

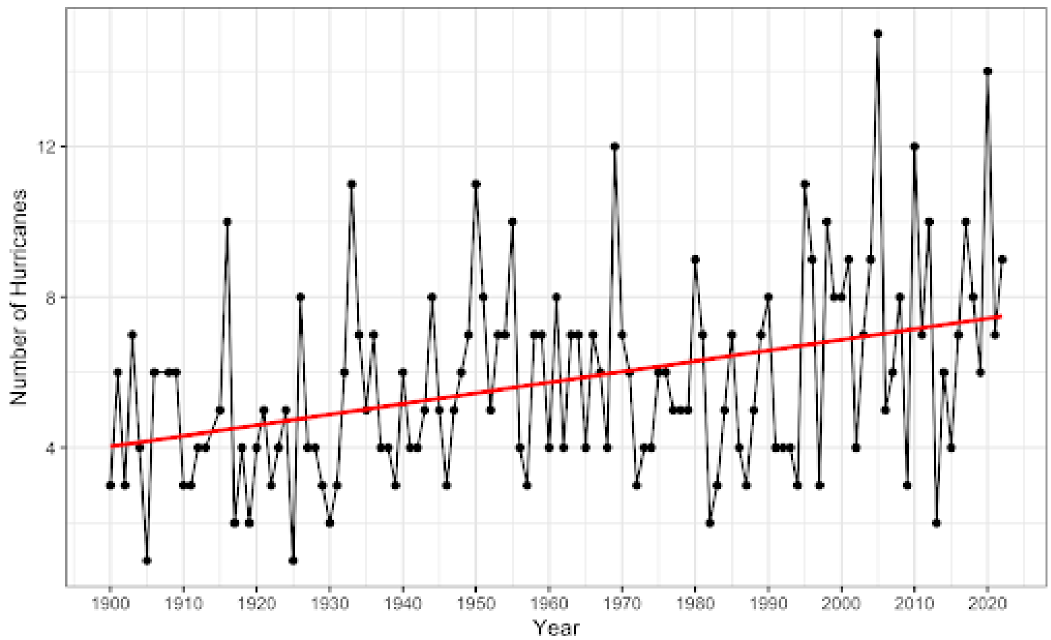


(n) Hurricane Hits On Texas Counties



(o) Hurricane Occurrence by Month

General Trend: Additionally, looking at the historical hurricane occurrence data, we observe a time series plot that records the annual frequency of hurricanes from the year 1900 to 2020. The plot reveals three significant patterns and trends:



(p) Historical Hurricane Count

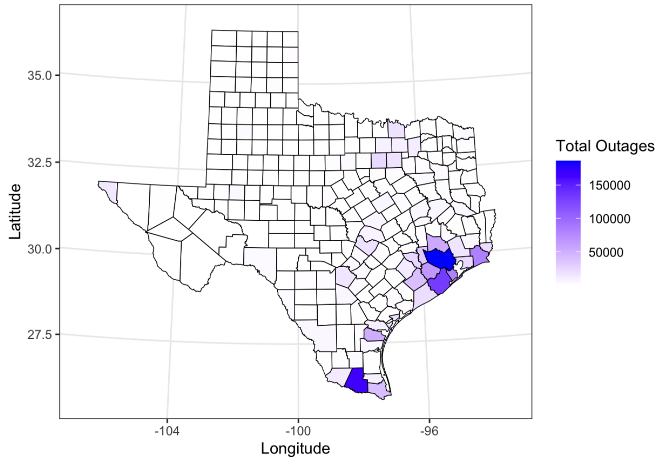
1. There is an evident upward trend in the number of hurricanes over the 120-year period, as indicated by the red trend line. This could suggest an increase in the frequency of hurricanes likely due to climate change, or mean the advancement of meteorological technology improves hurricane reporting and tracking.

2. The individual data points exhibit considerable fluctuation from year to year, likely indicating that the number of hurricanes is subject to significant variability. While the overall trend is upward, individual years may see spikes or drops in hurricane frequency.

3. There are periods where the number of hurricanes is consistently above the trend line, indicating phases of heightened hurricane activity. These clusters of years with high hurricane counts could be associated with multi-year climatic phenomena such as the warm phase of the Atlantic Multidecadal Oscillation (AMO) or other ocean-atmosphere interactions.

5.3 Severity

Space: The heatmap indicates the distribution of total power outages throughout Texas. The color gradient signifies the extent of outages, with darker shades representing more outages. The map shows a concentration of darker colors in the southeastern counties, which are coastal areas typically more vulnerable to hurricanes. This suggests that these regions experience higher power disruptions due to hurricanes, while the inland regions, marked by lighter shades, have fewer outages.

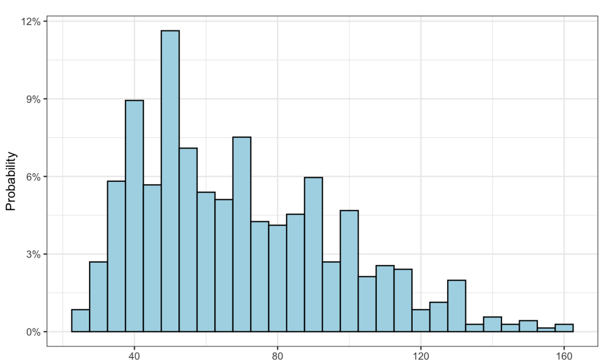


(q) Total Outages Count since 2019

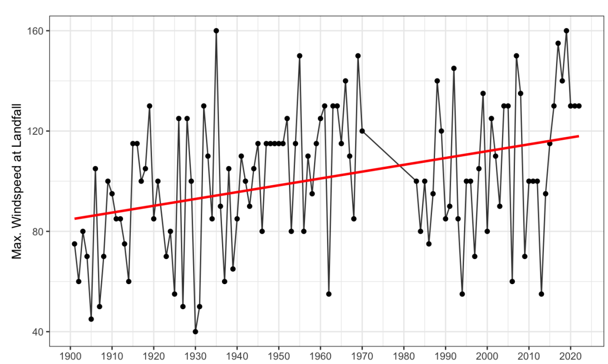
Time & Trend: The scatter plot with a trend line represents the maximum windspeed of hurricanes at landfall over a period extending from 1900 to 2020. Each data point corresponds to a hurricane event with the maximum windspeed in a particular year, with the red trend line illustrating a general increase in windspeed over time. This trend indicates that hurricanes making landfall have become more severe, with higher windspeeds, as time progresses.

There has been a clear increase in the maximum windspeed of hurricanes at landfall over the last century, as evidenced by the upward slope of the trend line in the scatter plot. This suggests a growing intensity in hurricane activity over time.

The histogram shows that while the most common windspeeds at landfall are between 50 and 60 knots, a broad range of windspeeds have been recorded, indicating variability in hurricane severity.



(r) Max. Wind Speed Distribution



(s) Max. Wind Speed Trend

5.4 Expected Value

In this context of hurricane damages, the expected value represents the anticipated number of power outages occurring within a specific county. For consistency and simplicity purposes, we've defined an outage as the maximum number of outages observed throughout the day. We then categorize these maximum daily outages into intervals based on deciles. Each interval represents a range of outage occurrences, with the size of the interval increasing notably for higher deciles.

To illustrate, let's again consider the adjacent Texas counties from the start of the paper: Brazoria, Fort Bend, Galveston, and Harris.

Interval	Probability (Brazoria)
(1, 36]	10.048%
(36, 65]	9.783%
(65, 98]	9.995%
(98, 152]	9.995%
(152, 285]	9.995%
(285, 519]	9.995%
(519, 842]	9.995%
(842, 1360]	9.995%
(1360, 2300]	9.995%
(2300, 318000]	9.995%

(a) Brazoria

Interval	Probability (Fort Bend)
(2, 48]	10.069%
(48, 71.2]	9.856%
(71.2, 100]	10.282%
(100, 141]	9.750%
(141, 216]	9.963%
(216, 372]	10.016%
(372, 617]	9.963%
(617, 994]	9.963%
(994, 2330]	10.016%
(2330, 172000]	10.016%

(b) Fort Bend

Interval	Probability (Galveston)
(1, 8]	8.519%
(8, 33]	10.185%
(33, 57]	9.583%
(57, 91]	10.093%
(91, 154]	9.861%
(154, 251]	10.093%
(251, 423]	10.000%
(423, 871]	9.907%
(871, 1930]	10.046%
(1930, 165000]	9.954%

(c) Galveston

Interval	Probability (Harris)
(1, 70]	7.977%
(70, 512]	9.995%
(512, 788]	9.995%
(788, 1120]	10.042%
(1120, 1570]	9.948%
(1570, 2150]	9.995%
(2150, 2840]	9.995%
(2840, 4060]	9.995%
(4060, 6500]	9.995%
(6500, 456000]	9.995%

(d) Harris

Intervals and percentages for various counties

Using the information above, we can compute the expected values for each county by multiplying each interval's midpoint by its corresponding probability and then summing up these products across all intervals.

County	Brazoria	Fort Bend	Galveston	Harris
Expected Value	565.4252	489.1474	380.9495	1960.004

6 Recommendations

6.1 Insurance

The impact of power outages due to hurricanes is generally more significant on businesses than on individuals. It can potentially hinder operational capabilities, cause indefinite closures, disrupt production schedules, and severely damage facilities. Therefore, we recommend acquiring adequate business interruption insurance to prepare for hurricane-related electrical issues. Our business insurance will compensate for the loss of income and cover operational expenses incurred during the suspension of business activities, which will likely result from direct and collateral damage to power infrastructure. This is especially critical for companies operating within hurricane-prone regions, where the propensity for such disruptions necessitates a comprehensive evaluation of potential risks and consequential financial vulnerabilities.
Furthermore, implementing business interruption insurance as a precaution helps facilitate a more resilient recovery framework. It enables businesses to maintain crucial financial obligations, such as employee wages and lease payments, in the aftermath of a hurricane-induced power outage. This proactive financial planning strategy contributes to broader economic stability by minimizing downtime and accelerating the restoration of necessary business operations post-disruption. Thus, the widespread utilization of business interruption insurance is a fundamental step in enhancing organizational adaptability and sustainability in the face of hurricane-related power challenges.

Furthermore, windstorm insurance will also benefit companies in regions prone to hurricanes for several reasons. Windstorms, including hurricanes, tornadoes, and severe thunderstorms, can cause significant damage to commercial properties, including buildings, equipment, and inventory. Windstorm insurance helps cover repairing or replacing damaged property, ensuring business owners can quickly resume operations after a storm. Repairing or rebuilding a business property after a windstorm can be expensive. Windstorm insurance provides financial protection by covering the costs associated with property damage, reducing the financial burden on business owners and helping them avoid substantial out-of-pocket expenses. Windstorms can cause property damage and bodily injury or property damage to third parties in addition to property damage. Windstorm insurance typically includes liability coverage, protecting business owners from legal claims and lawsuits arising from injuries or damages caused by the storm. Knowing that their business is adequately protected against the financial consequences of windstorm damage gives business owners peace of mind. With the right insurance coverage, they can focus on running their business without worrying about the potential impact of severe weather events. For both insurance plans, we will cover:

$$\frac{\text{Vulnerability[County]} + \text{Length Of Exposure}}{1.7 \cdot (\text{Max Vulnerability} + \text{Length Of Exposure})}$$

Vulnerability-calculated outages can be representative of overall damages expected from hurricanes as justified in our Assumptions and Justifications section. The longer the exposure, the greater the damages and the more our insurance will cover. This formula covers for potential property damage extrapolated from outage vulnerability and factors in the length of exposure, so therefore it is sufficient for windstorm property insurance. The greater the damage, the greater the business interruption period is, so therefore this formula applies to business interruption insurance as well.

We made sure that if a county is less vulnerable to damages, then it will receive a lower compensation because of lower expected damages compared to more vulnerable counties. Furthermore, we multiplied the denominator by a factor of 1.7 to ensure that insurance will overcompensate counties, thus allowing insurance companies to profit.

Example coverage (daily) of Brazoria, Fort Bend, Galveston, and Harris for Hurricane Nicholas (2021).

1. Brazoria: 57.34%
2. Fort Bend: 51.95%
3. Galveston: 53.055%
4. Harris: 50.55%

As the example coverages have demonstrated, for counties that are more vulnerable to hurricane damages receive greater coverage, while counties less vulnerable to damage receive lower coverage. This allows insurance companies to provide assistance to companies in high-risk counties while maintaining a profit.

6.2 Behavior Change

Power generators are devices designed to provide electricity by converting mechanical energy into electrical energy. This means that it can serve as a crucial backup power source during outages and provide temporary relief to areas that will lack access to the main power grid for extended periods of time. This small but significant shift in behavior amongst commercial entities and residential populations can serve as an important mitigative adaptation against disruptions caused by power outages. Incorporating power generators into contingency planning ensures an uninterrupted power supply, reducing the risks of operational disruptions, economic losses, and harm to individuals' well-being during catastrophic events. Providing every household with a power generator will drastically reduce the power outages caused by natural disasters. Those who own a power generator will have a lower natural disaster insurance premium to incentivize purchasing these items. The endorsement of power generators is a measure with immediate efficacy and an asset that fortifies businesses and communities against natural disasters' unpredictability and infrastructural vulnerabilities.

6.3 Modifying Outcomes

The transition to underground power lines within Texas constitutes an initiative to enhance the electrical grid's resilience against the detrimental impacts of hurricanes. This infrastructural modification, characterized by the underground installation of electrical distribution systems, significantly mitigates the exposure of power lines to environmental hazards such as high-velocity winds, airborne debris, and arboreal falls. We anticipate that this approach would substantially reduce the incidence and extent of power disruptions, thereby increasing the reliability of electricity supply across communities with different recovery quotas. Despite the initial capital outlay required for the implementation of underground power lines, the prospective benefits, including diminished maintenance exigencies and a reduction in emergency response expenditures post-disaster, highlight this investment's economic and safety advantages. It offers a sustainable solution to ensure uninterrupted power provision and fortify community resilience.

7 Conclusion

Urgent action is needed to address climate change and mitigate the growing threats of hurricanes as they grow more frequent and intense, posing heightened risks to coastal communities. Such could be indicated through a clear connection between climate change. Therefore, the state and local government needs a targeted approach to use their money more efficiently to support the resilience against these natural disasters. In addition, individuals could also take action to mitigate this risk by purchasing insurance and self-sufficient electricity generators so that they are much more prepared when facing these hurricanes. Lastly, our model provides path predictions to help state and local governments estimate the expected outages a particular hurricane may bring. This will help them strengthen their preparedness and accelerate a more immediate restoration process.

8 Acknowledgement

We sincerely thank our advisor, Mrs. MacDonald, for her longstanding support of our team, club, and school. Mrs. MacDonald is the one who shared this great opportunity with us, prepared us before the competition, and stuck with us throughout the contest.

We also thank our mentor Mrs. Witcraft for her constructive and specific comments on our paper and for providing us with clear and insightful advice throughout and after the two mentor meetings.

Last but not least, we would like to thank the Actuarial Foundation and all its partners and sponsors for making this program available to high schoolers at no cost at all. Throughout this competition, we gained insights into how to use math, computer science, and statistics to solve real-world problems and address the issues we care about.

9 References Cited

[1] Yale Climate Connections. (2023, November). The unusual 2023 Atlantic hurricane season ends. Retrieved from https://yaleclimateconnections.org/2023/11/the-unusual-2023-atlantic-hurricane-season-ends/

[2] Texas Tribune. (2021, September 13). Texas Tropical Storm Nicholas updates. Retrieved from https://www.texastribune.org/2021/09/13/texas-tropical-storm-nicholas-update s/

[3] Vault Electricity. (n.d.). Hurricane Harvey's impact on Texas electricity grid. Retrieved from https://www.vaultelectricity.com/hurricane-harveys-impact-on-texas-electricity-grid/

[4] World Economic Forum. (2021, March). What can Texas power crisis show us about building more resilient grid America? Retrieved from https://www.weforum.org/agenda/2021/03/what-can-texas-power-crisis-show-us-about-building-more-resilient-grid-america/

[5] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," in Neural Computation, vol. 9, no. 8, pp. 1735-1780, 15 Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.

[6] Landsea, C. W., & Franklin, J. L. (2013). Atlantic Hurricane Database Uncertainty and Presentation of a New Database Format. Monthly Weather Review, 141(10), 3576-3592. https://doi.org/10.1175/MWR-D-12-00254.1

[7] Masters, J., & Henson, B. (2023, November 30). The Unusual 2023 Atlantic Hurricane Season Ends. Yale Climate Connections. Retrieved from https://yaleclimateconnections.org/2023/11/the-unusual-2023-atlantic-hurricane-season-ends/

[8] Texas Tribune Staff. (2021, September 13). More than 100,000 Texas electricity customers without power after Nicholas. Restoring it to everyone could take days. Texas Tribune. Retrieved from https://www.texastribune.org/2021/09/13/texas-tropical-storm-nicholas-updates/

[9] Poudyal, A., Dubey, A., Iyengar, V., & Garcia-Camargo, D. (2022, July). Spatiotemporal impact assessment of hurricanes on electric power systems. In 2022 IEEE Power & Energy Society General Meeting (PESGM) (pp. 1-5). IEEE.

[10] Ouyang, M., & Duenas-Osorio, L. (2014). Multi-dimensional hurricane resilience assessment of electric power systems. Structural Safety, 48, 15-24.

[11] Tansakul, Varisara, Myers, Aaron, Tennille, Sarah, Denman, Matthew, Hamaker, Alec, Huihui, Jonathan, Medlen, Kyle, Allen, Karl, Redmon, Daniel, Chinthavali, Supriya, Coletti, Mark, Grant, Joshua, Lee, Matt, Maguire, Dakotah, Newby, Scott, Stahl, Chelsey, Bhaduri, Budhu, & Sanyal, Jibo. EAGLE-I Power Outage Data 2014 - 2022. United States. doi:https://doi.org/10.13139/ORNLNCCS/1975202