A blue and white sign with white text

Description automatically generated

**Capstone Project Phase A**

**Weather Catastrophe Prediction**

**24-2-D-30**

**Students:**

**Nermeen Basila 209290048**

**Madi Swead 313310088**

**Supervisor:**

**Dr. Zakharia Frenkel**

**Table of Contents**

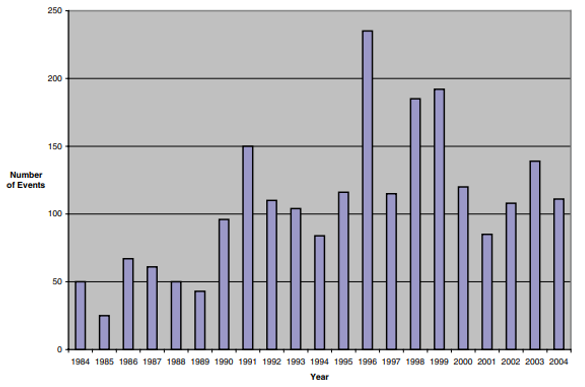
1. Abstract ………………………………………………………………………………3   
2.1 Introduction …………………………………………………………………………3  
 2.1. How historical data impacts predictions……………………………………… 6  
 2.2 short-Term Catastrophic Weather Prediction……………………………………6  
 2.3Long-Term Catastrophic Weather Prediction ……………………………………6  
2. Background ………………………………………………………………………….. 7  
 2.1 Temperature and Climate Change ………………………………………………..7  
 2.2 Humidity and Climate Change …………………………………………….…….8  
 2.3 Extreme Weather Events …………………………………………………………8  
 2.4 Climate Change Impacts ………………………………………………………….9  
 2.5 Definition of Catastrophic Weather Events ……………………………………. 10   
 2.5.1 Catastrophe in general ………………………………………………...……..10  
 2.5.2 Catastrophe in our concept …………………………………………………..10  
 2.6 Catastrophe reasons ……………………………………………………………. 10  
 2.7 Catastrophe Prediction …………………………………………………………..11  
 3. Cross Validation ……………………………………………………………………12  
 3.1 FBProphet Algorithm Overview ………………………………….……….…….13  
 3.2 Review of previous Research ……………………………………….………….. 13  
 3.2.1 Back Propagation Neural Network …………………………………….….… 14  
 3.2.1.1 parameters that used …………………………………………………....… 14  
 3.2.2 Year Lost Table …………………………………………………...……….…. 14  
 3.2.2.1 parameters that used …………………………………………….………... 14  
 3.2.3 Convolutional Neural Networks ……………………………….……………. 15   
 3.2.3.1 parameters that used ………………………………………………..….…. 15  
 3.2.4 Support Vector Machine Weather Prediction ……………….…………..…… 15  
 3.2.4.1 parameters that used …………………………………….…………..……. 15  
 3.2.5 numerical weather prediction ………………………………………………... 16  
 3.2.5.1 parameters that used …………………………………………………..…. 16  
 3.2.6 The reviewed papers utilized various algorithms for prediction ………………16  
 3.2.6.1 parameters that used ………………………………………………………..17  
 3.2.7 several algorithms used for weather prediction, for extreme weather events.….18  
 3.2.7.1 parameters that used ………………………………………………………. 18  
 3.3 Ways to calculate the parameters ………………………………………………… 18  
4. Available data ………………………………………………………………………... 21  
 4.1 Weather Catastrophe Examples …………………………….………………...….. 22  
5. Work plan ………………………………………………………………………….... 24  
 5.1 algorithm ……………………………………………………………….………... 26  
 5.2 Use Case ………………………………………………………………………… 27  
 5.2 Activity Diagram ……………………………………………………………...… 27  
 5.3 Tests ……………………………………………………………….……………... 28  
6. Conclusion ……………………………………………………………………………29  
7. References ……………………………………………………………………….…… 30

**Abstract**

This project investigates the relationship between regular weather patterns and catastrophic weather events, aiming to understand the underlying mechanisms and contributing factors of extreme weather phenomena. By analyzing historical weather data, current meteorological trends, and the increasing frequency and severity of catastrophic weather changes, the study seeks to identify patterns and correlations that may provide insights into their causes. Key variables such as temperature fluctuations, precipitation levels, atmospheric pressure, and wind patterns are examined to determine their impact on the formation of severe weather events like hurricanes, tornadoes, floods, and droughts. The project employs advanced data analytics and modeling techniques to predict future occurrences and assess the potential impacts of climate change on weather extremes. The findings are expected to enhance predictive models and support the development of mitigation strategies, contributing to better preparedness and response frameworks for communities vulnerable to severe weather conditions. Through this comprehensive analysis, the project aims to develop a tool management available big data of climate history and corresponding algorithm to bridge the gap.

**Introduction**

Catastrophic weather events have been on the rise over the years, reflecting an alarming trend. From 1984 to 2004, the number of state-level catastrophic events has shown a notable upward trajectory. For instance, in 1996, there were over 200 state-level catastrophic events. This increase is evident in the frequency and intensity of natural disasters such as hurricanes, storms, floods, and wildfires. Factors contributing to this rise include shifts in climatic conditions, such as global warming, which intensify weather patterns, and increased human activity in vulnerable areas, leading to greater exposure and impact. The trend suggests a growing vulnerability to natural catastrophes, posing significant challenges for communities, economies, and insurance markets alike.[16]

****

The frequency and intensity of catastrophic weather events have been on the rise in recent decades, posing significant challenges to communities worldwide. From devastating hurricanes to prolonged droughts and unprecedented flooding, these extreme weather phenomena not only disrupt daily life but also pose serious threats to infrastructure, agriculture, and human safety. Understanding the complex relationship between ordinary weather patterns and catastrophic events is crucial for effective mitigation and adaptation strategies in the face of climate change.

This project aims to delve into this intricate relationship by leveraging data-driven approaches to analyze historical weather patterns and predict catastrophic weather changes. By combining meteorological data, advanced analytics, and predictive modeling techniques, the study seeks to unravel the underlying factors contributing to the escalation of extreme weather events. Through comprehensive analysis and forecasting, the project endeavors to provide valuable insights into the dynamics of catastrophic weather changes and enhance preparedness measures for vulnerable regions.

The impacts of CC are far-reaching, affecting agriculture, public health, water resources, energy production, and biodiversity. Vulnerable populations, including marginalized communities and low-income individuals, are disproportionately affected. [1]

**How historical data impacts predictions:**

History can help to a limited extent because a climate situation never repeats itself.  
Observed data gives us a baseline to work from. We can see recent trends that help identify what to look for in the overall data. Take global temperatures, for example. We’ll have a simple metric of overall temperature changing over time. Sometimes that’s down to normal variability, and other times it’s more likely to be due to climate change. If we can isolate these from each other, we can make better predictions for the variable that we’re interested in. This could be temperatures at a specific location for example.

Sticking with the global temperature example, we know that temperatures are rising and that a warmer atmosphere can hold more moisture. This means there’s greater potential for more intense extreme rainfall events in future climates. But these events are likely to be less frequent because it takes longer for the moisture to build back up in the atmosphere. If we can predict global mean temperature, it will impact the predictions we make for mean rainfall in the future. It’s not just historical data of single weather phenomena that we use for our predictions. They’re often interlinked to consider things like carbon emissions, and climate change.

**short-Term Catastrophic Weather Prediction**

Short-term predictions, on the other hand, are critical for immediate weather events such as storms, floods, and heatwaves. Machine learning techniques, including deep learning and ensemble forecasting, are employed to analyze real-time data from various sources, such as satellite imagery and ground measurements. These methods can improve the accuracy of short-term forecasts by identifying patterns and anomalies in weather data that traditional models might miss.[23]

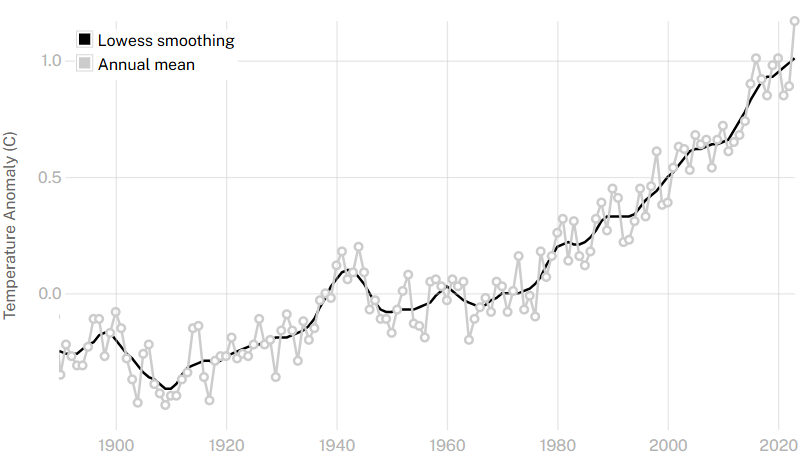
**Long-Term Catastrophic Weather Prediction**

Long-term predictions often focus on climate change impacts and trends over extended periods. Machine learning can enhance the accuracy of these predictions by analyzing vast datasets, including historical climate data and current environmental factors. The integration of machine learning with numerical weather prediction (NWP) models allows for better data assimilation and bias correction, which are crucial for understanding long-term climate patterns and extreme weather events.[23]

Machine learning is positioned as a transformative tool in both long-term and short-term catastrophic weather prediction, enhancing our ability to forecast and respond to extreme weather events.[23]

**Background**

Human civilization has evolved during the Holocene Era, the stability of which is now threatened by human-caused climate change. As a result, global catastrophic risk events from climate change are growing increasingly likely, there are many other potential global catastrophic risk events, both natural and human-caused, posing serious risks and warranting humanity’s careful consideration. there are cautions of “large uncertainty both for the likelihood of such events occurring and for their wider impact.” [2]



This graph shows the change in global surface temperature compared to the long-term average from 1951 to 1980. Earth’s average surface temperature in 2023 was the warmest on record since recordkeeping began in 1880 (source: [NASA/GISS](https://data.giss.nasa.gov/gistemp/graphs/graph_data/Global_Mean_Estimates_based_on_Land_and_Ocean_Data/graph.txt)). [3]

After researches we decided to explore How Humidity, Temperature affects global Climate and Extreme Weather.

Climate change is one of the most pressing challenges of our time, driven by a variety of factors, including temperature and humidity. Understanding how these elements interact and contribute to extreme weather events is crucial for developing effective strategies to mitigate and adapt to the changing climate.

* **Temperature and Climate Change:**

Temperature plays a fundamental role in regulating the Earth's climate. Over the past century, global temperatures have been rising at an unprecedented rate due to human activities, primarily the emission of greenhouse gases such as carbon dioxide (CO2), methane (CH4), and nitrous oxide (N2O). This phenomenon, known as global warming, has far-reaching consequences on the planet's weather patterns and ecosystems.  
We can measure the temperature in different heights.

**1. Global Warming**: The increase in average global temperatures leads to a variety of climatic changes. Warmer temperatures result in the melting of polar ice caps and glaciers, contributing to sea-level rise.

**2.** **Heatwaves**: Higher temperatures increase the likelihood and severity of heatwaves. These extreme heat events pose significant risks to human health, agriculture, and infrastructure.

* **Humidity and Climate Change:**

Humidity, the amount of water vapor present in the air, also plays a critical role in climate dynamics. It influences various atmospheric processes, including cloud formation, precipitation, and the overall energy balance of the planet.

**1.** **Water Vapor as a Greenhouse Gas**: Water vapor is the most abundant greenhouse gas in the atmosphere. It amplifies the warming effect of other greenhouse gases through a positive feedback loop. As the temperature rises, more water evaporates, increasing humidity. Higher humidity, in turn, traps more heat in the atmosphere, further accelerating global warming.

**2. Precipitation Patterns**: Changes in humidity levels directly impact precipitation patterns. Higher temperatures increase the rate of evaporation and the capacity of the air to hold moisture. This can lead to more intense and frequent rainfall events, contributing to flooding in some regions. Conversely, other areas may experience prolonged droughts due to altered precipitation patterns.

* **Extreme Weather Events:**

The interplay between temperature and humidity significantly influences the occurrence and intensity of extreme weather events.

**1. Hurricanes and Typhoons**: Warmer Ocean temperatures provide more energy for tropical storms, leading to stronger and more destructive hurricanes and typhoons. Increased humidity levels contribute to heavier rainfall associated with these storms, exacerbating flooding and storm surge impacts.

**2. Droughts and Wildfires**: Rising temperatures can lead to more frequent and severe droughts, particularly in regions already prone to dry conditions. Prolonged droughts reduce soil moisture and vegetation health, increasing the risk of wildfires. These fires can spread rapidly, fueled by dry conditions and high temperatures.

**3. Flooding and Heavy Rainfall**: As global temperatures rise, the atmosphere can hold more moisture, resulting in more intense and frequent heavy rainfall events. This can lead to severe flooding, particularly in areas with inadequate infrastructure to manage large volumes of water.

The intricate relationship between temperature, humidity, and extreme weather events highlights the complexity of climate change. Understanding these interactions is essential for predicting future climate scenarios and developing strategies to mitigate and adapt to their impacts. This project aims to explore the mechanisms through which temperature and humidity contribute to global climate change and extreme weather, providing insights into the challenges we face and potential solutions for a sustainable future. [4]

**Climate Change Impacts**:

1. **Agriculture**: CC affects crop yields, food security, and farmers' livelihoods due to changing climate patterns, precipitation levels, and extreme weather events. Sustainable agricultural practices and the development of resilient crop varieties are essential to mitigate these effects. The research highlights the importance of soil seed banks in maintaining plant community stability and resilience in the face of climate change, particularly in regions experiencing significant variations in rainfall and temperature.
2. **Sea Level Rise**: This poses significant risks to coastal regions and island nations. Strategies such as coastal protection measures and sustainable urban planning are crucial to protect these areas.
3. **Water Resources**: CC exacerbates water scarcity by affecting freshwater reserves and altering precipitation patterns. Sustainable water management practices and investment in water-efficient technologies are vital.
4. **Human Health**: CC increases health risks through elevated temperatures and changing patterns of infectious diseases. Vulnerable communities are especially at risk, necessitating health policies that address these emerging challenges.
5. **Biodiversity**: CC threatens biodiversity by altering or destroying habitats, pushing many species toward extinction. Conservation efforts and sustainable land use practices are critical.
6. **Forest Ecosystems**: Forests face threats from both deforestation and CC, impacting their role as carbon sinks. Mitigation measures like afforestation and improved forest management are important. [1]

**Catastrophe in general**

A catastrophic weather event is a severe and extreme weather phenomenon that causes significant damage, destruction, and loss of life. These events are often characterized by their intensity, scale, and impact on the affected area. Examples of catastrophic weather events include hurricanes, tornadoes, floods, wildfires, and severe storms. These events can result in widespread devastation, displacement of populations, and long-term consequences for the environment and infrastructure.[15]

**Catastrophe in our concept**

Catastrophic weather change refers to extreme and unexpected weather events that significantly deviate from the usual climatic patterns of a region and cause severe disruption and damage. These events are characterized by their rarity and the magnitude of their impact. For instance, while tornadoes are relatively common in certain parts of America and are expected occurrences, they would be considered catastrophic in regions where they rarely occur. Similarly, if a country like Sweden, which typically experiences mild summers, were to encounter temperatures soaring above 40°C, it would be deemed a catastrophic event due to its unprecedented nature and the potential for widespread harm to the environment, infrastructure, and public health.

**Catastrophe reasons**

catastrophes, particularly weather-related catastrophes, can occur due to various reasons. Some common factors contributing to catastrophes include:

**1.** **Severe Weather Events:** Catastrophes can result from extreme weather conditions such as hurricanes, tornadoes, floods, and wildfires, which can cause widespread destruction and significant property damage.

**2.** **Climate Change:** Changes in climate patterns, including global warming, can lead to more frequent and intense weather events, potentially increasing the likelihood of catastrophes.

**3.** **Infrastructure Vulnerability:** The susceptibility of infrastructure to damage from severe weather events can contribute to the occurrence of catastrophes, especially in areas with inadequate infrastructure resilience.[10]

**Catastrophe prediction**

Predicting catastrophic weather events involves a combination of advanced technology, scientific knowledge, and data analysis. Here are some key methods and tools used to forecast and predict catastrophic weather events:

**1. Meteorological Models:** Meteorologists use sophisticated computer models to simulate the atmosphere and predict weather patterns. These models analyze data on temperature, humidity, wind speed, and other factors to forecast the development of severe weather events.[8]

**2.Weather Observations:** Ground-based weather stations, weather balloons, and other observational tools collect data on temperature, pressure, humidity, and wind conditions. This data is essential for understanding current weather patterns and making accurate forecasts.[8]

**3. Climate Models:** Climate models analyze long-term trends and patterns to predict the likelihood of extreme weather events, such as heatwaves, droughts, and heavy rainfall. These models consider factors like sea surface temperatures, atmospheric circulation, and greenhouse gas concentrations.[8]

**4. Historical Data Analysis:** Studying historical trends of catastrophes, including frequency, severity, and impact, can help identify patterns and potential risk factors for future events. [10]

**5. Early Warning System:** Developing early warning systems for natural disasters like hurricanes, tsunamis, and floods can help alert populations in advance, giving them time to evacuate and take necessary precautions.

**Cross Validation**

Cross-validation, sometimes called rotation estimation or out-of-sample testing, is any of various similar [model validation](https://en.wikipedia.org/wiki/Model_validation) techniques for assessing how the results of a [statistical](https://en.wikipedia.org/wiki/Statistics) analysis will [generalize](https://en.wikipedia.org/wiki/Generalization_error) to an independent data set. Cross-validation includes [resampling](https://en.wikipedia.org/wiki/Resampling_(statistics)) and sample splitting methods that use different portions of the data to test and train a model on different iterations. It is often used in settings where the goal is prediction, and one wants to estimate how [accurately](https://en.wikipedia.org/wiki/Accuracy) a [predictive model](https://en.wikipedia.org/wiki/Predictive_modelling) will perform in practice. It can also be used to assess the quality of a fitted model and the stability of its parameters.

In a prediction problem, a model is usually given a dataset of known data on which training is run (training dataset), and a dataset of unknown data (or first seen data) against which the model is tested (called the [validation dataset](https://en.wikipedia.org/wiki/Validation_set) or testing set). The goal of cross-validation is to test the model's ability to predict new data that was not used in estimating it, in order to flag problems like [overfitting](https://en.wikipedia.org/wiki/Overfitting) or [selection bias](https://en.wikipedia.org/wiki/Selection_bias)and to give an insight on how the model will generalize to an independent dataset (i.e., an unknown dataset, for instance from a real problem).

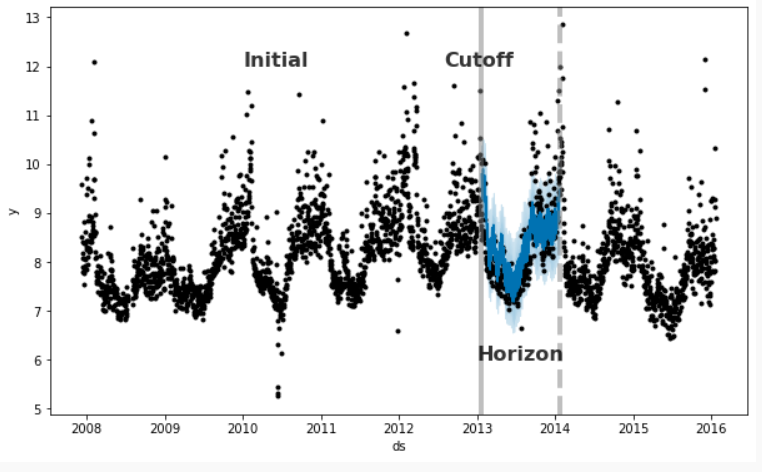
One round of cross-validation involves [partitioning](https://en.wikipedia.org/wiki/Partition_of_a_set) a [sample](https://en.wikipedia.org/wiki/Statistical_sample) of [data](https://en.wikipedia.org/wiki/Data) into [complementary](https://en.wikipedia.org/wiki/Complement_(set_theory)) subsets, performing the analysis on one subset (called the training set), and validating the analysis on the other subset (called the validation set or testing set). To reduce [variability](https://en.wikipedia.org/wiki/Variance), in most methods multiple rounds of cross-validation are performed using different partitions, and the validation results are combined over the rounds to give an estimate of the model's predictive performance.

In summary, cross-validation combines (averages) measures of [fitness](https://en.wikipedia.org/wiki/Goodness_of_fit) in prediction to derive a more accurate estimate of model prediction performance.[6]

**FBProphet Algorithm**

FBProphet is a**forecasting algorithm** developed by Facebook’s data science team in 2017. The algorithm is designed to be scalable, fast, and accurate, making it suitable for a wide range of applications, from predicting sales in e-commerce to forecasting weather patterns.[5]

Prophet includes functionality for time series cross validation to measure forecast error using historical data. This is done by selecting cutoff points in the history, and for each of them fitting the model using data only up to that cutoff point. We can then compare the forecasted values to the actual values. This figure illustrates a simulated historical forecast on the Peyton Manning dataset, where the model was fit to an initial history of 5 years, and a forecast was made on a one-year horizon.[5]

****

This cross-validation procedure can be done automatically for a range of historical cutoffs using the cross-validation function. We specify the forecast horizon (horizon), and then optionally the size of the initial training period (initial) and the spacing between cutoff dates (period). By default, the initial training period is set to three times the horizon, and cutoffs are made every half a horizon. The output of cross validation is a data frame with the true values y and the out-of-sample forecast values yhat, at each simulated forecast date and for each cutoff date. A forecast is made for every observed point between cutoff and cutoff + horizon. This data frame can then be used to compute error measures of yhat vs. y. Here we do cross-validation to assess prediction performance on a horizon of 365 days, starting with 730 days of training data in the first cutoff and then making predictions every 180 days. On this 8-year time series, this corresponds to 11 total forecasts. [5]

**Review of Previous Research**

**Similar approaches have been previously undertaken by several researchers, including:**

1) In the document, the author utilized a Back Propagation Neural Network (BPNN) to classify and predict weather conditions. Historical data on temperature, humidity, and pressure were used to train the BPNN model, which was then employed to forecast the weather for the subsequent day. The performance of the BPNN model was compared with other machine learning techniques, including Decision Tree and Naive Bayes. The results indicated that the BPNN method performed well in weather prediction. The thesis concludes that the Backpropagation Neural Network is an effective method for weather forecasting, demonstrating high accuracy in classifying and predicting weather conditions.[11]

Here are potential parameters that used:[11]

1. **Temperature**: Extreme highs or lows that are significantly above or below the historical average.
2. **Humidity**: Extremely high humidity levels that can indicate the potential for storms or other weather events.
3. **Pressure**: Rapid drops in atmospheric pressure which can indicate the approach of a storm or other severe weather event.
4. **Wind Speed**: High wind speeds which can indicate storms, hurricanes, or tornadoes.
5. **Precipitation**: High levels of precipitation, such as heavy rain or snowfall, which can lead to flooding or other weather-related disasters.

2)This paper introduces a novel simulation algorithm that incrementally creates a new year loss table (YLT) by adjusting the original YLT with just enough events to capture the change in extreme weather hazard, specifically focusing on hurricanes and climate change. This incremental simulation method aims to reduce simulation noise and enhance the precision of estimates of change in catastrophe risk models. By copying the original YLT and selectively adding or removing events to reflect the adjusted hazard, this algorithm eliminates the need for approximations and ensures that the new YLT closely represents the desired changes. Through testing on various U.S. hurricane loss models with adjustments for climate change, the study demonstrates that the incremental simulation approach significantly improves the accuracy and precision of estimating the impact of climate change on extreme weather events, providing more reliable insights for risk assessment and decision-making processes.[12]

Here are potential parameters that used:[12]

1. **Frequency**: Frequency refers to how often a particular weather event occurs over a specific period.
2. **Intensity:** measures the strength or severity of a weather event.
3. **Rainfall**: measures the amount of precipitation (rain) that falls during a weather event. It is usually measured in millimeters or inches over a specific period.
4. **Speed**: is the speed at which a weather system, such as a storm or hurricane, moves across the surface of the Earth. It is usually measured in kilometers or miles per hour.

3)Deep learning algorithms have been increasingly employed for weather prediction due to their ability to handle complex and non-linear patterns in data. Convolutional Neural Networks (CNNs) are prominently used for extracting features from meteorological data, such as satellite images and radar data, to predict extreme weather events. CNNs excel in identifying spatial patterns and can effectively process large datasets, although they require significant computational resources​. Additionally, Recurrent Neural Networks (RNNs), particularly Long Short-Term Memory (LSTM) networks, are utilized for their capability to capture temporal dependencies in sequential data, making them suitable for time-series forecasting of weather phenomena. These deep learning frameworks can automatically learn from vast amounts of historical weather data, enhancing the accuracy of predictions for events such as heavy rainfall, thunderstorms, and hurricanes. By integrating these advanced algorithms, meteorologists can improve the prediction of extreme weather conditions, potentially mitigating their adverse impacts on society.[13]

Here are potential parameters that used:[13]

**1.Temperature:** The ambient temperature, which can impact the likelihood of certain weather events, such as heatwaves or cold spells.

**2.Humidity:** The amount of moisture in the air, which is critical for predicting events like fog, rain, and storms.

**3.Rainfall:** The amount of precipitation, which is crucial for predicting flooding and other water-related disasters.

**4)**The prediction algorithm used in the article "Support Vector Machine Weather Prediction Technology Based on the Improved Quantum Optimization Algorithm" is the Support Vector Machine (SVM). The study focuses on utilizing SVM in conjunction with the Improved Quantum Genetic Algorithm (IQGA) to optimize parameters for weather prediction. Additionally, the study compares the performance of SVM models optimized by different algorithms such as Genetic Algorithm (GA-SVM) and Improved Ant Colony Optimization. [14]

Here are potential parameters that used:[14]

**1.** **Temperature:** Extreme highs or lows indicating potential heatwaves or cold spells.

**2.** **Humidity:** High humidity levels indicating potential for storms or heavy rainfall.

**3.** **Wind Speed:** High wind speeds indicating potential for storms, hurricanes, or tornadoes.

**4.** **Precipitation:** Amount of rainfall indicating potential for flooding or severe weather.

**5.** **Cloud Cover:** Dense cloud cover indicating potential for precipitation or severe weather.

5) The researchers utilized a genetic algorithm-based framework for automatically identifying weather systems from numerical weather prediction (NWP) data. Genetic algorithms are optimization algorithms inspired by the process of natural selection and genetics. In this context, general algorithms were employed to analyze multiple meteorological elements and identify patterns representing weather systems such as tropical cyclones, fronts, troughs, ridges, and pressure extrema. By formulating the weather system identification problem as a pattern recognition task and designing a generic model of weather systems, the researchers were able to use the genetic algorithm framework to automatically discover these patterns from NWP data. The framework allowed for the analysis of different meteorological elements and the evaluation of fitness functions to assess the effectiveness of the model and framework. Overall, the genetic algorithm-based approach provided a method for automatically identifying and locating weather systems with high precision, offering an independent and objective source of information to assist forecasters in their work.[20]

Here are potential parameters that used:[20]

1. Pressure: Pressure is the force exerted by the atmosphere at a given point. In meteorology, pressure is an essential parameter for understanding weather patterns and systems.

2. Temperature: Temperature is a measure of the average kinetic energy of particles in a substance. In meteorology, temperature plays a crucial role in determining weather conditions and atmospheric stability.

3. Dewpoint: Dewpoint is the temperature at which air becomes saturated with water vapor and dew begins to form. It is a key parameter for assessing humidity levels in the atmosphere.

4. Precipitation: Precipitation refers to any form of water, liquid or solid, that falls from the atmosphere and reaches the ground. It includes rain, snow, sleet, and hail. 1) The algorithms used for prediction in the reviewed papers include:

6)The reviewed papers utilized various algorithms for prediction, including:

1. MetNet-2: A deep neural network-based weather model that outperforms existing physics-based models in predicting high-resolution precipitation up to 12 hours ahead.

2. MLP feedforward backpropagation ANN model: Employed for the retrieval algorithm in predicting Total Precipitable Water (TPW) and Convective Available Potential Energy (CAPE).

3. Artificial Neural Networks (MLP, RBF, GRNN): Used to predict Tropical Cyclone (TC) frequency based on large-scale climate variables.

4. Convolutional Neural Network (CNN): Utilized for predicting hailstorms and severe hail events, encoding spatial weather data for improved accuracy.

5. LSTM algorithm: Employed for nowcasting typhoon tracks and predicting tropical cyclone movement over a 24-hour timeframe.

6. Various modeling approaches (Linear Regression, Lasso Regression, Polynomial Regression, AdaBoost, Decision Trees, Random Forest, CNN, CNN with Recurrence Plots, RP+CNN with binary fusion): Explored for long-term air temperature prediction in summer using ECMWF's ERA5 reanalysis data.

These algorithms showcase the diverse applications of deep learning techniques in predicting various weather phenomena and improving forecast accuracy.[21]

Here are potential parameters that used:[21]

**1. Sea Surface Temperature (SST):** The temperature of the ocean surface, which plays a significant role in influencing weather patterns and climate variability. Changes in SST can impact atmospheric circulation, precipitation patterns, and the development of tropical cyclones

**2. Relative Humidity (RH):** The amount of water vapor present in the air relative to the maximum amount the air can hold at a specific temperature. RH is essential for understanding atmospheric moisture levels, cloud formation, and precipitation processes.

**3. Geopotential Height:** A measure of the height of a pressure surface in the atmosphere above mean sea level. Geopotential height is used in weather forecasting to analyse atmospheric stability, pressure systems, and the movement of air masses.

**4. Dewpoint:** The temperature at which air becomes saturated with water vapor, leading to the formation of dew or fog. Dewpoint is a critical parameter for assessing atmospheric moisture content and predicting the likelihood of precipitation.

**5. Zonal Wind**: The component of wind that blows parallel to lines of latitude, from west to east or east to west. Zonal wind patterns influence global circulation patterns, weather systems, and the movement of air masses across the Earth's surface.

**6. Meridional Wind:** The component of wind that blows perpendicular to lines of latitude, from north to south or south to north. Meridional wind patterns play a role in atmospheric circulation, weather fronts, and the transport of heat and moisture in the atmosphere.

7)This research references several algorithms used for weather prediction, particularly for extreme weather events. These algorithms include:[22]

**1. Artificial Neural Networks (ANN):** Used in various studies for weather forecasting by modelling nonlinear relationships between input data and weather conditions. For example, ANN with Error Backpropagation Algorithm has been employed for predicting weather patterns.

**2. Random Forest Algorithm:** This algorithm has been applied for controlling weather-dependent tasks and predicting specific weather-related events.

**3. Fireworks Algorithm:** This algorithm has been used for training neural networks specifically for weather forecasting purposes.

**4. Dynamic Neural Network Architecture with Immunology Inspired Optimization:** Used for weather data forecasting, integrating advanced optimization techniques to improve prediction accuracy.

These machine learning algorithms help improve the accuracy and reliability of weather forecasts, especially in predicting extreme weather events.

Here are potential parameters that used:[22]

The parameters used to predict extreme weather events in the referenced document include:

1. **Temperature**: Various temperature-related parameters are considered, such as surface temperature and temperature at different atmospheric levels.
2. **Humidity**: Humidity levels, including specific humidity and relative humidity, are used to predict weather conditions.
3. **Wind Speed and Direction**: These parameters help in understanding and predicting the movement of weather systems.
4. **Precipitation**: Amount and intensity of precipitation are critical for predicting extreme weather events like heavy rainfall and storms.
5. **Pressure**: Atmospheric pressure measurements are used to analyze and predict weather patterns.
6. **Solar Radiation**: This is considered in some models to understand its impact on temperature and weather dynamics.
7. **Sea Surface Temperature**: Particularly relevant for predicting phenomena like hurricanes and other ocean-related weather events.
8. **Cloud Cover**: The extent and type of cloud cover help in forecasting weather changes.

These parameters are fed into various machine learning algorithms to model and predict extreme weather events accurately.

Ways to calculate these parameters:

1. **Temperature:**

TST: is the true solar time in decimal hours since sunrise.  
 T-max and T-min: are the maximum and minimum ambient temperature during the day. [18]

**Or**

T=T0+ΔT

where T0​ is the initial temperature and ΔT is the change in temperature over time, which can be modeled using differential equations.[22]

**2.Humidity:** RH=100X

Rh: relative humidity percent.  
e:is the current vapor pressure.  
 : is the saturated vapor pressure at the given temperature. [17]

**3. wind Speed:** V=d/t

V: Wind speed (e.g., in m/s)  
d: Distance travelled by the air (e.g., in meters)  
t: Time taken to travel that distance (e.g., in seconds)

**Or**

The wind speed can be calculated using the formula: W=U2+V2 where:   
a. W is the wind speed.  
b. U is the zonal wind component.   
c. V is the meridional wind component

**4.Precipitation:**

a. Simple Linear Regression Formula P=a⋅T+b

P: Predicted precipitation (e.g., in mm)  
T: Temperature (e.g., in °C)  
a, b: Coefficients determined through regression analysis of historical data

b. Multiple Linear Regression Formula P=a0+a1⋅T+a2⋅H+a3⋅W

P: Predicted precipitation  
T: Temperature  
H: Humidity (e.g., in %)  
W: Wind speed (e.g., in m/s)  
a\_0, a\_1, a\_2, a\_3: Coefficients determined through regression

**5. Intensity:** I= **I:** is the intensity.  
  **P**: is the power.  
 **A:** is the area of cross-section. [19]

**6. Pressure (P):**

a. The pressure at a specific location can be calculated using the ideal gas law: P=ρ⋅R⋅T, where:   
P is the pressure.  
ρ is the density of the air.  
R is the specific gas constant for dry air.   
T is the temperature.

b. Calculated using the ideal gas law: P= ​, where P is pressure, n is the number of moles of gas, R is the gas constant, T is temperature, and V is volume.

**7.Dewpoint (D):**

a. The dewpoint can be calculated using the formula: D=T−(5100−RH) where:  
D is the dewpoint temperature.  
T is the temperature.  
RH is the relative humidity.

b. The dewpoint temperature can be calculated using the formula: Td = T - ((100 - RH) / 5), where Td is the dewpoint temperature,

T is the air temperature,

RH is the relative humidity.

**8. Wind Speed and Direction**: Calculated using vector components of wind velocity, u (zonal wind) and v (meridional wind): Speed= ​, and direction θ=arctan ).[22]

The current algorithms for predicting extreme weather changes, including neural networks, simulation algorithms, CNNs, and SVMs, are not achieving the desired level of accuracy. To address this, we are exploring the new modern Prophet algorithm. Prophet, developed by Facebook, is particularly suited for time series forecasting due to its ability to handle missing data, seasonal variations, and trend changes more effectively than the other methods. Unlike traditional neural networks and CNNs that require extensive tuning and large datasets, or SVMs that struggle with large-scale data, Prophet offers a robust and user-friendly approach. Its flexibility and strong performance with historical data make it a promising alternative for improving the accuracy and reliability of extreme weather predictions.

**Available data**

There are several types of catastrophes and climatic parameters associated with them. In our project we concentrate majorly on the temperature, because such data is more available. Generally, we reviewed two types of data: the first is directly associated with catastrophes. Such data is required for criteria for automatic detection of the events. The second type is historical climatic data for “reverse engineering” and testing of our predictive algorithms.

**Weather Catastrophe Examples:**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **When it happens** | **What happens** | **Damages** |
| **Hurricane Katrina [24]** | August 23, 2005 | 1) Hurricane Katrina was a Category 5 Atlantic hurricane that caused severe destruction along the Gulf Coast of the United States.  2) The storm surge overwhelmed the levees in New Orleans, causing widespread flooding. | 1) Over 1,800 deaths.  2) $125 billion in damages.  3)Displacement of hundreds of thousands of residents.  4)Extensive destruction of homes, infrastructure, and the local economy. |
| **Indian Ocean earthquake and tsunami [25][26]** | December 26, 2004   at 07:58:53 | 1)Triggered by a massive undersea earthquake off the coast of Sumatra, Indonesia.  2) Generated a series of devastating tsunamis that affected 14 countries**.** | 1) Approximately 230,000 to 280,000 deaths.  2) Millions of people displaced.  3) Widespread destruction of coastal communities, infrastructure, and economies.  4) Long-term environmental and social impacts. |
| **Typhoon Haiyan [27][28]** | November 3, 2013 | 1) Known as Yolanda in the Philippines, it was one of the strongest tropical cyclones ever recorded.  2) Produced winds of up to 195 mph and massive storm surges. | 1) Over 6,300 deaths in the Philippines.  2) $2.98 billion in damages.  3) Widespread destruction of homes, infrastructure, and livelihoods.  4) Severe humanitarian crisis, with millions requiring aid and support. |
| **European Heatwave**  **[29]** | Summer of 2003 | 1)An intense heatwave affected much of Europe, with temperatures reaching unprecedented levels.  2) Particularly severe in France, Italy, Spain, and Portugal**.** | 1) Estimated 70,000 deaths due to heat-related illnesses.  2) Agricultural losses estimated at €13 billion.  3) Increased risk of wildfires and damage to infrastructure. |
| **Pakistan Floods [30]** | July 26,2010 | 1)Unprecedented monsoon rains led to widespread flooding across Pakistan.  2) The Indus River overflowed, affecting large swathes of the country. | 1) Approximately 1,985 deaths.  2) Over 20 million people affected.  3) $43 billion in damages.  4)Extensive damage to infrastructure, homes, and agriculture. |
| **Hurricane Sandy [31]** | **October 24, 2012** | 1)A post-tropical cyclone that affected the Caribbean and the Eastern United States.  2)Caused widespread flooding and power outages, particularly in New York and New Jersey. | 1) 233 deaths.  2)$70 billion in damages.  3)Extensive damage to homes, infrastructure, and the economy.  4) Long-term displacement and recovery efforts. |
| **Hurricane Maria [32]** | September 16, 2017 | 1)Hurricane Maria was a Category 5 hurricane that devastated Puerto Rico and other parts of the northeastern Caribbean.  2) It brought intense winds, heavy rainfall, and massive storm surges. | 1) Approximately 3,000 deaths, making it one of the deadliest hurricanes to hit the region.  2) Over $90 billion in damages.  3) Extensive destruction of infrastructure, homes, and power grids, leaving many residents without electricity for months. |

**Work Plan:**

Detecting Catastrophic Events Using Geographical Division and SQL Server Data

**Part 1:** System Setup and Geographical Division

**1.** System Setup: Connecting to SQL Server  
Objective: Establish a connection to an SQL Server to read and process data related to geographical stations and historical events.

Components Involved:

- SQL Server: The database containing the required data.  
- Connection Interface: Use a programming language such as C++ to establish the connection.  
- Data retrieval: Run queries to retrieve complete data related to geographic stations, historical events, and associated parameters.

**2.** Geographical Division: Dividing the Earth into Regions

Objective: Divide the earth into distinct regions based on station data correlations to facilitate localized analysis.

Steps Involved:

**-** Region Definition: Define regions by analyzing correlations between stations. Stations with strong data correlations will be grouped into the same region. **-** Formulas: Develop formulas to calculate the inter-station correlations and define regional boundaries.

**Part 2:** Parameter Assumption and Catastrophe Identification

**1.** Parameter Assumption: Defining Parameters for Catastrophe Detection

Objective: Analyze historical data to establish parameters for detecting catastrophes in each region.

Steps Involved:

**-** Historical Data Analysis: Analyze past catastrophic events (e.g., earthquakes, floods, hurricanes) to identify patterns and thresholds for key parameters such as temperature, seismic activity, and pressure. **-** Parameter Definition: Define specific parameters that indicate potential catastrophes, such as: **-** Thresholds: Set upper and lower limits for each parameter. **-** Standard Deviations: Calculate the deviation of current data from the historical mean. Large deviations may signal unusual activity. **-** Region-Specific Customization: Adjust the parameters for each region based on its unique historical and environmental characteristics.

**2.** Catastrophe Identification: Detecting Catastrophes

Objective: Detect potential catastrophes by comparing real-time data against the predefined parameters.

Steps Involved:

- Data Monitoring: Continuously monitor data from stations in each region.  
**-** Anomaly Detection: Compare real-time data to the defined thresholds and standard deviations to detect anomalies that may indicate a catastrophe.

**Part 3:** Collection and Analysis of Catastrophic Events

**1.** Collection and Analysis

Objective: After identifying a potential catastrophe, collect relevant data and perform an in-depth analysis to understand its onset and causes.

Steps Involved:

- Data Compilation: Gather data from the stations during the period of anomaly detection.  
- Onset Analysis: Determine the starting point of the unusual event by identifying the first significant deviation.  
**-** Reason Documentation: Analyze potential causes of the catastrophe, including environmental conditions, human activities, and other contributing factors.  
- Report Generation: Document the event details, including the onset, the parameters that triggered the alert, and the findings from the analysis.

**Algorithm:**

1. Grid Creation: Begin by constructing a geographic grid with a constant cell size of 200x200 kilometers. This grid will serve as the framework for spatially organizing climate data, ensuring consistent analysis of parameters across the Earth’s surface.

2. Definition of Catastrophic Metrics: For each defined climatic parameter (such as temperature, humidity, or pressure), determine thresholds for what constitutes a “catastrophic value” and a “catastrophic change.” These thresholds will help identify significant deviations from normal conditions that may signal the onset of catastrophic events.

3. Daily Parameter Calculation and Comparison: Daily, calculate various climate-related parameters (e.g., temperature, pressure, wind speed, precipitation) and their differences between consecutive days for each cell in the grid. For each calculated parameter, compare the daily average with an “ideal model,” which represents the expected or normal state. Deviations beyond predefined thresholds will help flag potentially catastrophic events.

4. Identification of Catastrophic Cells: Identify all cells within the grid that exhibit either a “catastrophic value” or a “catastrophic change” based on the threshold definitions. These cells are flagged for further analysis as they may indicate the presence of a catastrophic event.

5. Trajectory Building for Catastrophic Events: For each identified catastrophic event, trace its trajectory by locating neighboring cells in both space and time that also exhibit catastrophic characteristics. This step involves connecting neighboring events to form a spatial and temporal path, which will reveal how the event evolved over time and across regions.

6. Identification of Trajectory Starting Points: Once trajectories are constructed, identify the initial points—cells where the event originated. These starting points represent the locations where the catastrophic events were “born,” providing crucial information about the origins of the events.

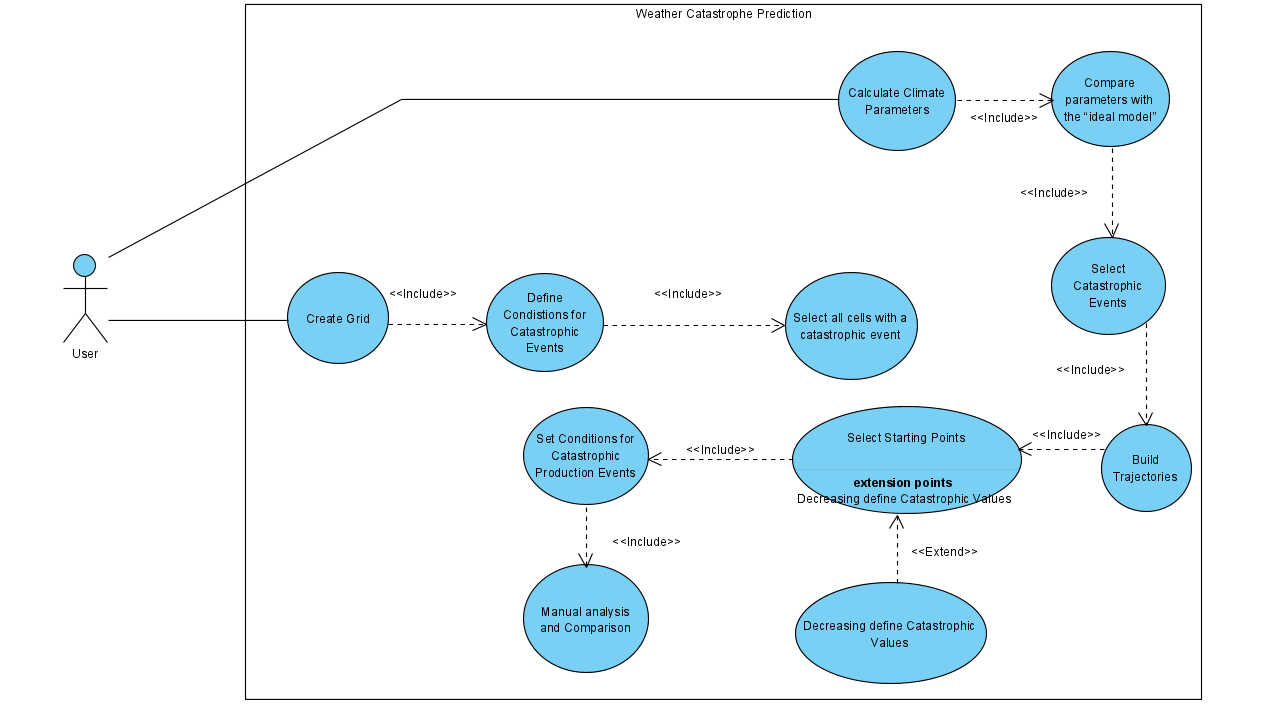
7. Backtracking to Earlier Points: To further understand the development of the event, attempt to trace the trajectory backward by lowering the threshold for what is considered a “catastrophic value.” By identifying earlier points with less severe conditions, it may be possible to detect precursor signs of the event and extend the understanding of its origin.

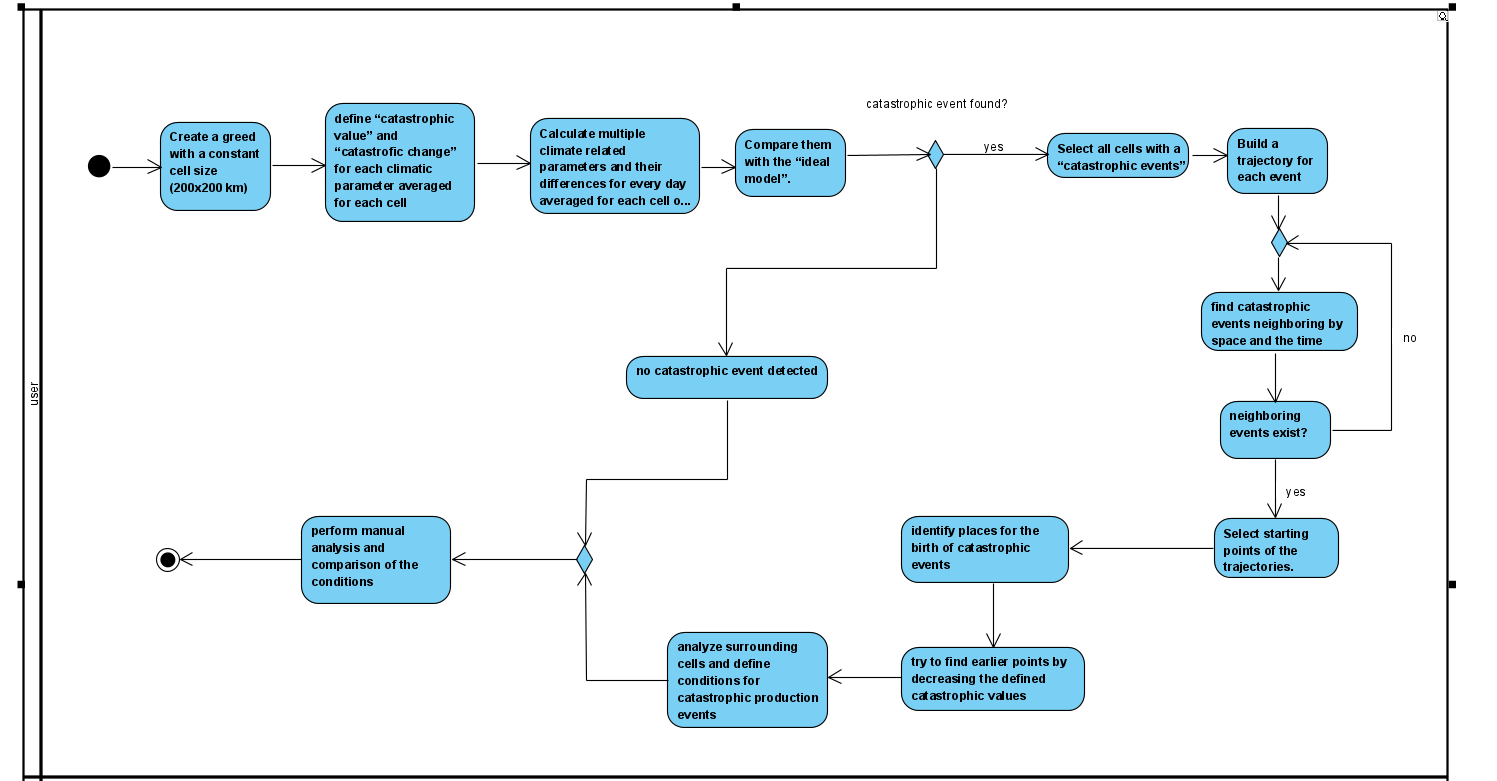
8. Conditions for Catastrophic Event Generation: Analyze the states of surrounding cells at the event’s starting point to determine the environmental conditions that contributed to the event's formation. These surrounding cells' climatic states will be compiled into a set of “conditions for catastrophic production events,” which describe the circumstances that led to the event's birth.

9. Manual Analysis and Comparison: Finally, conduct a manual analysis of the identified conditions. This step involves comparing the catastrophic event conditions across multiple trajectories and events, searching for patterns or recurring factors that may provide insights into the underlying causes of catastrophic event generation. Manual analysis allows for nuanced interpretation and the integration of expert judgment to enhance the algorithm’s results.

**UML Diagrams:**

**Use Case:**

**Activity diagram:**

****

**Tests:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Test No.** | **Test case** | **Input** | **Expected output** |
| **1** | Ensure that the grid is correctly created with cells of size 200x200 km | Boundary coordinates, cell size (200x200 km) | Correct grid dimensions with no errors in size or positioning of cells |
| **2** | Ensure that the entire area of interest is covered by the grid | Region boundary coordinates | No gaps in the grid for the defined region |
| **3** | Verify that each climatic parameter for a cell has its catastrophic value defined correctly | Climatic parameter averages for a cell | Defined catastrophic values based on predefined thresholds |
| **4** | Ensure that significant changes in climatic parameters over time trigger a "catastrophic change" flag. | Climatic parameter time series. | Correct identification of catastrophic changes based on defined thresholds. |
| **5** | Ensure that climate-related parameters are averaged for each cell correctly on a daily basis. | Climate parameter data for a day | Correct daily average for each cell. |
| **6** | Verify the comparison between the calculated parameters and the "ideal model. | Averaged parameters, ideal model data. | Correctly flagged cells where deviations from the model occur |
| **7** | Ensure that the correct cells are identified as having catastrophic events | Climatic parameter values for each cell. | Correct selection of cells that meet the catastrophic threshold |
| **8** | Verify that catastrophic events neighboring in both space and time are correctly linked | Spatial and temporal data of events | Correctly identified trajectories of connected catastrophic events |
| **9** | Ensure that the trajectory of events is accurately built based on proximity and sequence of occurrences. | Set of catastrophic events with timestamps | Correctly built trajectory of events |
| **10** | Verify that the correct starting points (where catastrophic events originate) are identified | Trajectories of catastrophic events | Correct identification of starting points |
| **11** | Ensure that reducing the catastrophic value threshold leads to the identification of earlier starting points. | Trajectory data and lowered catastrophic value threshold. | Correct identification of earlier points in the trajectory |
| **12** | Verify that the states of surrounding cells are correctly stored and analyzed | Climatic data for surrounding cells | Correct set of surrounding cell conditions is established |
| **13** | Ensure that the manual comparison of conditions between different catastrophic events is feasible and that conditions are clearly presented for analysis | Set of conditions for different catastrophic events | Clear and accurate presentation for manual analysis |

**Conclusion**

In conclusion, our project has successfully delved into the intricate dynamics of catastrophic weather events, their prediction, and the significant factors contributing to these phenomena. Through extensive research and the application of advanced data analytics, we have deepened our understanding of how temperature, humidity, and other environmental variables interact to trigger extreme weather conditions. This project has not only allowed us to learn and summarize a vast array of scientific insights, but it also provided the opportunity to explore modern prediction algorithms, which we applied effectively to forecast future catastrophic events.

While we considered established algorithms such as FBProphet, our project introduced a new and innovative approach to predicting catastrophic weather events, as described in our project documentation. This new algorithm combines several advanced techniques tailored to our specific objectives, offering a unique method of analyzing and forecasting extreme weather patterns. We continue to refine and test this algorithm, with the goal of improving its accuracy and achieving better results in future predictions.

The knowledge gained throughout this process has been invaluable, and we are confident that our approach will lead to more reliable and effective predictive models, contributing to better disaster preparedness and response systems.

**References:**

**[1] :** Rawat, A., Kumar, D., & Khati, B. S. (2024). A review on climate change impacts, models, and its consequences on different sectors: a systematic approach. *Journal of Water and Climate Change*, *15*(1), 104-126.

**[2]:** [The future of global catastrophic risk events from climate change » Yale Climate Connections](https://yaleclimateconnections.org/2022/07/the-future-of-global-catastrophic-risk-events-from-climate-change/)

**[3]:** [Global Temperature | Vital Signs – Climate Change: Vital Signs of the Planet (nasa.gov)](https://climate.nasa.gov/vital-signs/global-temperature/?intent=111)

**[4]:** [National Oceanic and Atmospheric Administration (noaa.gov)](https://www.noaa.gov/)

**[5]:** [Prophet | Forecasting at scale. (facebook.github.io)](https://facebook.github.io/prophet/#:~:text=Prophet%20is%20a%20forecasting%20procedure%20implemented%20in%20R,tuned%20by%20hand%20by%20data%20scientists%20and%20analysts.)

**[6]:** [Cross-validation (statistics) - Wikipedia](https://en.wikipedia.org/wiki/Cross-validation_(statistics))

**[7]:** [catastrophe](https://he.wikipedia.org/wiki/%D7%90%D7%A1%D7%95%D7%9F) weather wikipedia

**[8]:** Lloyd, E. A., & Shepherd, T. G. (2020). Environmental catastrophes, climate change, and attribution. *Annals of the New York Academy of Sciences*, *1469*(1), 105-124.

**[9]:** Hansen, J., Ruedy, R., Glascoe, J., & Sato, M. (1999). GISS analysis of surface temperature change. *Journal of Geophysical Research: Atmospheres*, *104*(D24), 30997-31022.

**[10]:** Changnon, D., & Changnon Jr, S. A. (1998). Evaluation of weather catastrophe data for use in climate change investigations. *Climatic Change*, *38*(4), 435-445.

**[11]:** [**http://www.ijeert.org/pdf/v2-i2/35.pdf**](http://www.ijeert.org/pdf/v2-i2/35.pdf)

**[12]:**Jewson, S. (2023). A new simulation algorithm for more precise estimates of change in catastrophe risk models, with application to hurricanes and climate change. *Stochastic Environmental Research and Risk Assessment*, *37*(7), 2631-2650.‏

**[13]:**Fang, W., Xue, Q., Shen, L., & Sheng, V. S. (2021). Survey on the application of deep learning in extreme weather prediction. *Atmosphere*, *12*(6), 661.‏

**[14]:** Zhang, J., Qiu, X., Li, X., Huang, Z., Wu, M., & Dong, Y. (2021). Support vector machine weather prediction technology based on the improved quantum optimization algorithm. *Computational Intelligence and Neuroscience*, *2021*(1), 6653659.

**[15]:** Singh, M. S., & O’Neill, M. E. (2022). Thermodynamics of the climate system. *Physics Today*, *75*(7), 30-37.

**[16]:** Born, P., & Viscusi, W. K. (2006). The catastrophic effects of natural disasters on insurance markets. *Journal of risk and Uncertainty*, *33*, 55-72.‏

**[17]:** Lawrence, M. G. (2005). The relationship between relative humidity and the dewpoint temperature in moist air: A simple conversion and applications. *Bulletin of the American Meteorological Society*, *86*(2), 225-234.

**[18]:**Aoun, N. (2022). Methodology for predicting the PV module temperature based on actual and estimated weather data. *Energy Conversion and Management: X*, *14*, 100182.‏

**[19]:** [Intensity Formula -Equation and Solved Examples (byjus.com)](https://byjus.com/intensity-formula/)

[20]: Wong, K. Y., Yip, C. L., & Li, P. W. (2008). Automatic identification of weather systems from numerical weather prediction data using genetic algorithm. Expert systems with Applications, 35(1-2), 542-555.

**[21]:** <https://arxiv.org/pdf/2308.10995>

**[22]:** <https://www.researchgate.net/publication/376521669_Review_of_AI-Enhanced_Weather_Forecasting_Application_for_Communication_Networks>

[23]: Bochenek, B., & Ustrnul, Z. (2022). Machine Learning in Weather Prediction and Climate Analyses—Applications and Perspectives. Atmosphere, 13(2), 180.

[24]: [Hurricane Katrina - Wikipedia](https://en.wikipedia.org/wiki/Hurricane_Katrina)

[25]: [2004 Indian Ocean earthquake and tsunami: Facts and FAQs | World Vision](https://www.worldvision.org/disaster-relief-news-stories/2004-indian-ocean-earthquake-tsunami-facts)

[26]: [2004 Indian Ocean earthquake and tsunami - Wikipedia](https://en.wikipedia.org/wiki/2004_Indian_Ocean_earthquake_and_tsunami)

[27]: [Typhoon Haiyan: Facts, FAQs, and how to help | World Vision](https://www.worldvision.org/disaster-relief-news-stories/2013-typhoon-haiyan-facts)

[28]: [Typhoon Haiyan - Wikipedia](https://en.wikipedia.org/wiki/Typhoon_Haiyan)

[29]: [2003 European heatwave - Wikipedia](https://en.wikipedia.org/wiki/2003_European_heatwave)

[30]: [2010 Pakistan floods - Wikipedia](https://en.wikipedia.org/wiki/2010_Pakistan_floods)

[31]: [Hurricane Sandy - Wikipedia](https://en.wikipedia.org/wiki/Hurricane_Sandy)

[32]: [Hurricane Maria - Wikipedia](https://en.wikipedia.org/wiki/Hurricane_Maria)