Climate Change Data Analysis

Submitted in partial fulfillment of the requirements of the degree of

Bachelor of Engineering

by

Name	Roll No
Fernandes Renita Christopher	312015
Kamaraj Lisha	312026
Manihar Naaz Fatima	312028
Upase Prajakta Jotiram	312059

Project Guide:

Prof. Ahlam Ansari

(Computer Engineering)

M. H. Saboo Siddik College of Engineering
University of Mumbai

2023-24

M.H. SABOO SIDDIK COLLEGE OF ENGINEERING

8, Saboo Siddik Road, Byculla, Mumbai - 400 008.

This is to certify that,

Roll No	Name
312015	Fernandes Renita Christopher
312026	Kamaraj Lisha
312028	Manihar Naaz Fatima
312059	Upase Prajakta Jotiram

Of Final Year (B.E. Semester VII) degree course in Computer Engineering, have completed the specified project report on,

Climate Change Data Analysis

As partial fulfillment of the project works in a satisfactory manner as per the rules of the curriculum laid by the University of Mumbai, during the Academic Year July 2023 - December 2023.

Internal Guide External Examiner

Project Report Approval for B. E.

This project report entitled **"Climate Change Data Analysis"** by **Fernandes Renita Christopher, Kamaraj Lisha, Manihar Naaz Fatima and Upase Prajakta Jotiram** is approved for the degree of Computer Engineering.

EXAMINERS

1. ______________________________

2. ______________________________

SUPERVISORS

1. ______________________________

2. ______________________________

Date:

Place: Mumbai

Acknowledgment

We would like to express our gratitude and appreciation to our parents for motivating and encouraging us throughout the career.

We wish to express our sincere thanks to our Principal Prof. Javed Habib, M.H. Saboo Siddik College of Engineering for providing us all the facilities, support, and wonderful environment to meet our project requirements.

We would also take the opportunity to express our humble gratitude to our Head of the Department of Computer Engineering Dr. Mohammed Ahmed Shaikh for supporting us in all aspects and for encouraging us with her valuable suggestions to make our project successful.

We are highly thankful to our internal project guide Prof. Ahlam Ansari whose valuable guidance helped us understand the project better, her constant guidance and willingness to share her vast knowledge made us understand this project and its manifestations in great depth and helped us to complete the project successfully.

We would also like to acknowledge with much appreciation the role of the Computer Department staff, especially the Laboratory staff, who permitted us to use the labs when needed and the necessary material to complete the project.

We would like to express our gratitude and appreciate the guidance given by other supervisors and project guides, their comments and tips helped us in improving our presentation skills.

Although there may be many who remain unacknowledged in this humble note of appreciation, there are none who remain unappreciated.

Table of Content

List of Figures

Abstract

The most basic stage in forecasting climate change is the analysis of climate data. Its primary goal is to improve knowledge of the atmosphere and how it interacts with the seas, the cryosphere, and the land surface using a variety of methods or approaches. These include mathematical data modeling, diagnostic analyses, and empirical research on climate data.

Climate change is an important subject that requires in-depth research and comprehension of its underlying causes and effects. The ability of big data analytics to handle large and complicated information is essential for understanding the subtleties of climate change. The goal of this small project is to employ big data analytics tools to examine data on climate change and derive insightful conclusions.

Several climate-related datasets, such as temperature records, precipitation data, sea level measurements, greenhouse gas emissions, and other pertinent variables, are first gathered as part of the project's multi-step methodology. In order to generate a structured data repository appropriate for analysis, these datasets are combined, purified, and altered.

1. Introduction

Climate change is a critical global challenge that has far-reaching implications for our planet. As the effects of climate change become increasingly evident, there is a growing need to understand, monitor, and mitigate its impact. Big data analytics plays a pivotal role in this endeavor, providing the tools and techniques necessary to analyze vast and complex datasets related to climate and environmental data.Climate change data analysis is vital for several reasons:

- Understanding Trends: Analyzing climate data helps us identify long-term trends, such as rising temperatures, shifting precipitation patterns, and sea-level rise.
- Predicting Impacts: By analyzing historical climate data and projecting future scenarios, we can predict the potential impacts of climate change on ecosystems, economies, and human societies.
- Supporting Policy and Decision-Making: Governments, organizations, and individuals use climate data analysis to make informed decisions and develop policies for climate mitigation and adaptation.

Climate change data is characterized by several attributes that make it well-suited for big data analytics. Climate data is vast and continually growing, with historical records and real-time measurements. Real-time data from satellites, weather stations, and other sources is generated at high speeds. Data comes in various formats, including structured, semi-structured, and unstructured data. Complexity is seen where climate data is multidimensional, containing information on temperature, precipitation, greenhouse gas levels, and more.Climate data analysis faces challenges such as data quality issues, integration of disparate datasets, computational complexity, and the need for domain expertise.The analysis of climate data can be applied in numerous ways-

- ❖ Climate Prediction: Modeling future climate scenarios to understand the potential consequences of climate change.
- ❖ Extreme Weather Events: Studying patterns to predict and mitigate the impact of extreme weather events such as hurricanes and droughts.
- ❖ Ecosystem Health: Monitoring and preserving ecosystems by analyzing climate data's impact on biodiversity.
- ❖ Energy and Resource Management: Optimizing energy usage and resource allocation based on climate data.

2. Aim & Objective

2.1. Aim

The aim of a climate change data analysis project can vary depending on the specific objectives and scope of the project, but generally, it involves using data to gain insights into various aspects of climate change. Here are some common aims for such projects:

- Assessing Climate Trends: Analyze historical climate data to identify trends and patterns related to temperature, precipitation, and other climate variables. This can help in understanding how the climate has been changing over time.
- Modeling Future Climate Scenarios: Use climate models and data to project future climate scenarios. This can help in predicting how the climate might change in the coming decades, which is crucial for planning and mitigation efforts.
- Impact Assessment: Analyze the impact of climate change on various sectors, such as agriculture, water resources, and public health. This can help in identifying vulnerabilities and developing adaptation strategies.
- Carbon Footprint Analysis: Evaluate the carbon emissions from different sources, such as industries, transportation, and energy production. This can help in understanding the contributions to greenhouse gas emissions.
- Ecosystem Analysis: Study the effects of climate change on ecosystems, biodiversity, and habitats. This can help in conservation and preservation efforts.
- Policy and Decision Support: Provide data-driven insights to policymakers and decision-makers to inform climate-related policies and strategies.
- Public Awareness and Education: Use data analysis to create visualizations and educational materials to raise public awareness about climate change.
- Mitigation Strategies: Identify and analyze the effectiveness of various mitigation strategies, such as renewable energy adoption, reforestation, and energy efficiency measures.
- Extreme Weather Events Analysis: Study the frequency and intensity of extreme weather events like hurricanes, droughts, and wildfires to understand their links to climate change.
- Social and Economic Impacts: Examine the social and economic impacts of climate change, including effects on vulnerable populations, food security, and economic stability.
- Global Climate Agreements Compliance: Evaluate whether countries are meeting their climate targets as per international agreements like the Paris Agreement.

2.2 Objective

The primary objective of conducting a comprehensive analysis of climate change data is to gain a deeper understanding of the complex climate system, its changes over time, and their implications. The specific objectives are as follows:

➢ Identify Trends : To detect and quantify long-term trends in global and regional climate variables, such as temperature, precipitation, and sea level, in order to understand the direction and magnitude of climate change.

➢ Assess Variability : To examine the natural variability in climate patterns and distinguish it from anthropogenic influences, providing insights into the causes of climate change.

➢ Detect Extreme Events : To identify and characterize extreme weather and climate events, including heatwaves, floods, and hurricanes, to assess their frequency, intensity, and impacts.

➢ Analyze Emissions : To assess greenhouse gas emissions, their sources, and their relationship to climate change, allowing for informed mitigation strategies.

➢ Evaluate Impacts : To understand the ecological, social, and economic consequences of climate change, including effects on ecosystems, human health, and infrastructure.

➢ Predict Future Scenarios : To develop predictive models that project future climate conditions under different emission scenarios, aiding in decision-making and adaptation planning.

➢ Inform Policy and Mitigation: To provide data-driven insights to policymakers and stakeholders, facilitating evidence-based decisions and actions to mitigate and adapt to climate change.

➢ Communicate Findings : To effectively communicate the results of the analysis through data visualization and reports, ensuring that the information is accessible and actionable for a wide audience.

By achieving these objectives, the analysis of climate change data aims to contribute to a more sustainable and resilient future by guiding policies and actions that address the challenges posed by a changing climate.

3. Problem Statement

Climate change is a pressing global issue, and there is a need for comprehensive data analysis to better understand its impacts and inform policy decisions. The problem is to conduct a detailed analysis of climate change data. Examine historical climate data to identify long-term trends in temperature, precipitation, and other relevant variables. Detects and analyzes extreme weather events, such as heatwaves, hurricanes, and droughts, to assess their frequency and intensity. Investigate regional variations in climate change effects and their potential consequences for specific geographic areas. Assess greenhouse gas emissions and their correlation with climate change trends. Evaluate the ecological, social, and economic impacts of climate change, including sea-level rise, agricultural productivity, and health effects. Develop predictive models to forecast future climate change scenarios under various emission reduction strategies. Create informative visualizations to communicate the findings effectively to policymakers, scientists, and the general public. The solution to this problem will help in making informed decisions and policies to mitigate and adapt to climate change, ultimately contributing to a more sustainable and resilient future.

4. Implementation

In a big data analysis project for climate change, large datasets are examined by utilizing big data analysis and advanced data analytics. It seeks to pinpoint global climate change trends, patterns, and possible effects. The project uses machine learning to assess various climate-related data sources and offers useful insights to help with policy formation and mitigation of climate change. It also frequently highlights the importance of open data principles for data repeatability and cooperation. The analysis is concentrated on the GFDL-CM4 climate model. Future climatic conditions are projected using this model.

Methodology :

Climate change data analysis involves the following methodology:

Data Source: Observed precipitation data, notably the CPC-CONUS dataset from the NOAA (National Oceanic and Atmospheric Administration), are used in the analysis of climate change data. Precipitation data for the contiguous United States is probably included in this collection.Two distinct datasets' amount of air temperature data are used in this section. It centers on one particular occurrence, the cold storm that hit Texas in February of 2021. This occurrence is crucial to the analysis.

Time Period: Given a period of four years of data being used, it is implied that the dataset spans a considerable amount of time.

Functions and Operations: xarray, a well-liked library for managing multidimensional arrays, is used to perform crucial functions and operations, particularly for processing geospatial and climate data.

The following are a few of the operations to be learned:

- **groupby:** This function is used for time-based grouping (e.g., monthly or yearly aggregations) where data is grouped according to specific criteria.
- **concat:** Data from several sources or eras can be combined via concatenation.
- **sel and isel:** These functions are employed to pick particular data according to variables, time, or location.

Leap Year Management: This will cover leap year management, as leap years have an extra day

(February 29th). Maintaining correct time series data requires doing this.

Output File Format: The processed and analyzed data are saved as NetCDF files in the output file format. Because it can hold metadata in addition to the data, NetCDF is a widely used format for meteorological and climate data.

Interpolation: One method for estimating values between known data points is interpolation. It probably refers to adding missing or intermediate temperature readings in this case.

Scaling Datasets: Kelvin (°K) to Celsius (°C) temperature values are being converted. This improves the data's readability and usefulness.

Rolling Average: This technique involves averaging numbers over a predetermined time period to smooth out data. Identifying patterns and trends is important.

Sample Plots: In order to better understand the data, it involves making visual representations, or plots. These maps most likely depict temperature variations in Texas during the cold storm.

Calculating Temperature Anomaly: The analysis determines the difference between the predicted temperature and the typical expected temperature for February for each member of the model's ensemble (variations of the forecast). We call this discrepancy the "temperature anomaly." This makes it easier to determine how much the actual temperature varied from the predicted value.

Ensemble Mean: The model's average of all temperature predictions is known as the ensemble mean. It offers a comprehensive picture of the predictions made by the entire model.

Monthly Temperature Changes: Throughout these time periods, the analysis examines variations in monthly temperatures. It wants to know how the temperature changes from month to month.

Historical and Future Periods: Two eras are taken into account. The first is the historical era, which represents the previous environment, and spans from 1980 to 2010. The SSP585 future projection, which covers the years 2070 through 2099, is the second. This prediction looks at a very extreme future and is meant to be an example only; it does not necessarily depict the actual course of global warming.

Result Visualization: The temperature anomaly for every member of the model's ensemble is displayed in the result visualization. It suggests that most of the fluctuations of the cold anomaly in February 2021 were accurately predicted by the model's forecasts.These measurements aid in understanding historical temperature trends as well as the most extreme future scenario.

5. Outputs

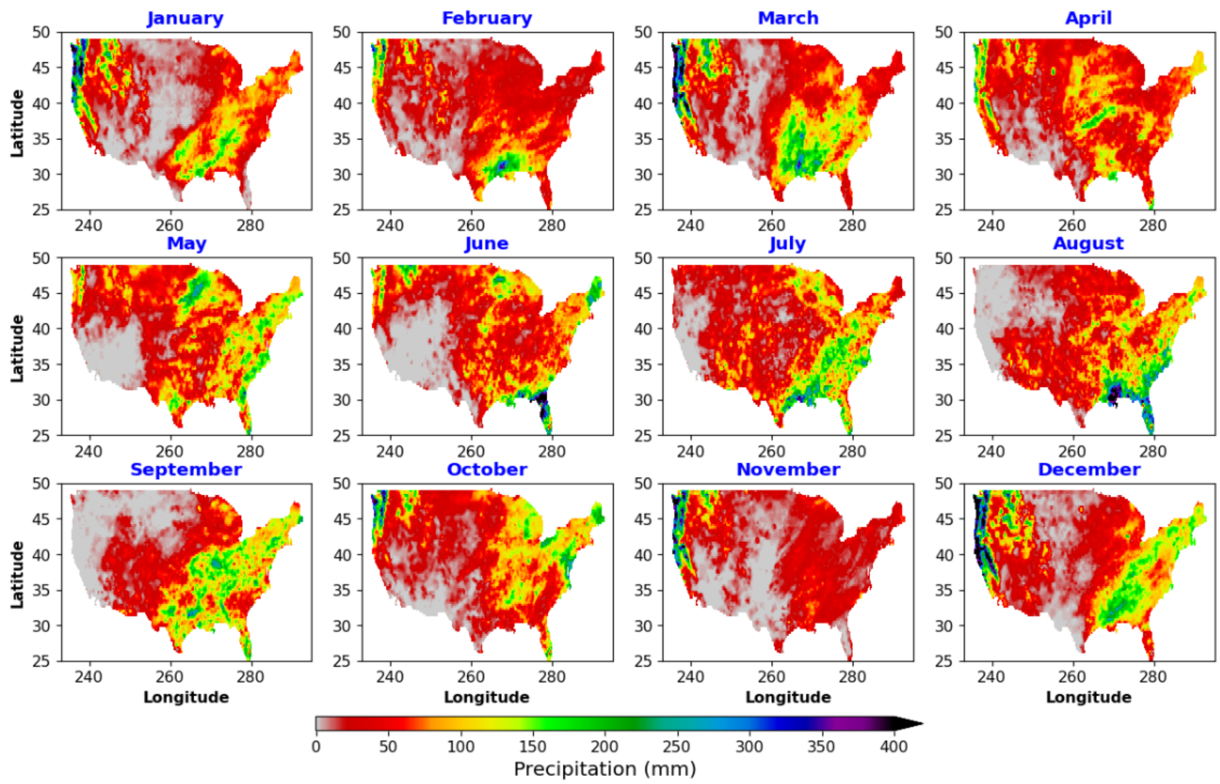

Fig 3.1 : Extract daily observed precipitation data

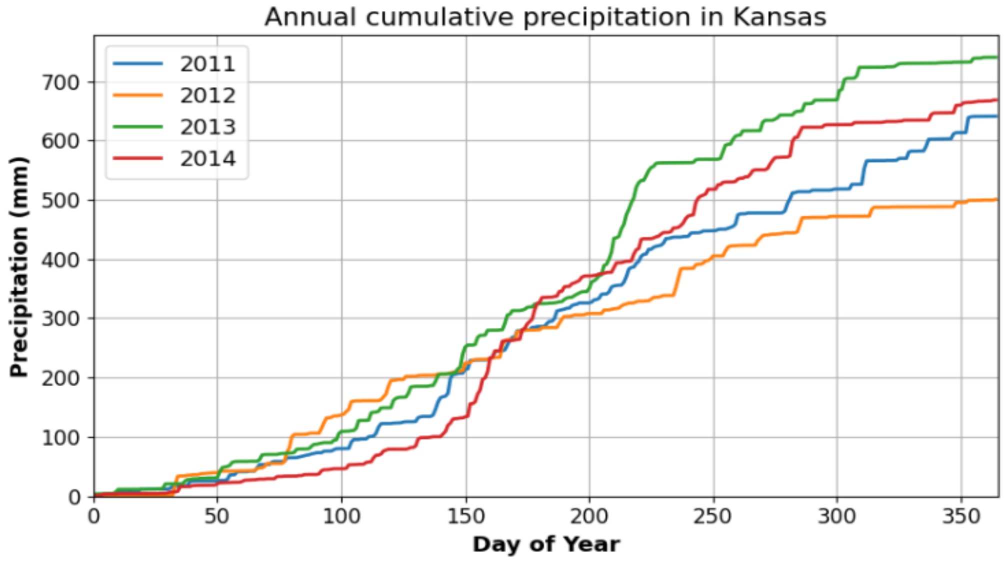

Fig 3.2 : Annual cumulative precipitation in Kansas

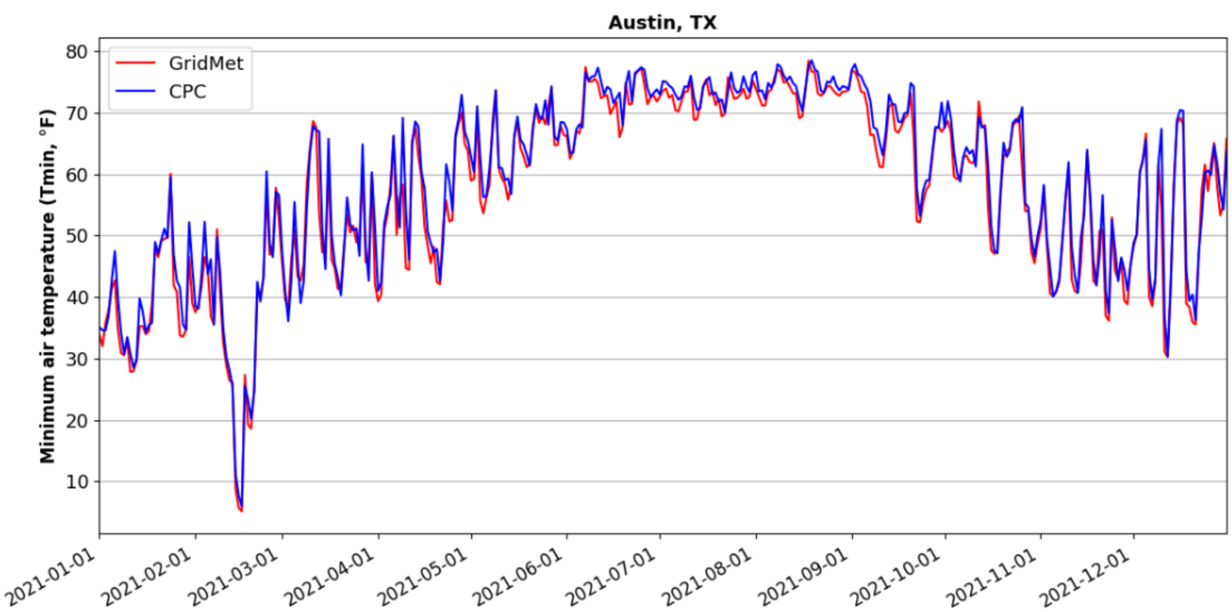

Fig 3.3: Minimum air temperature (Tmin, °F)

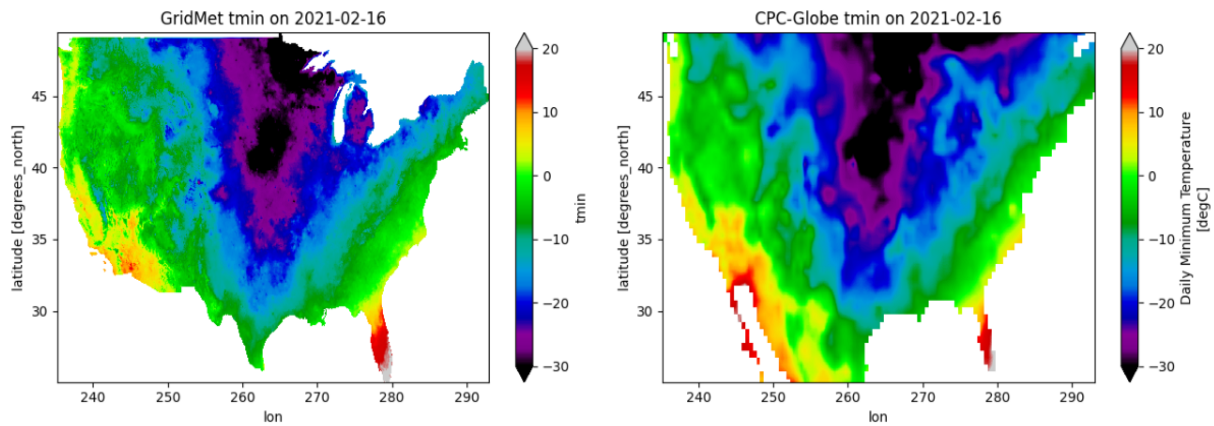

Fig 3.4: Daily Minimum air temperature (Tmin, °F) Two of the sample plots

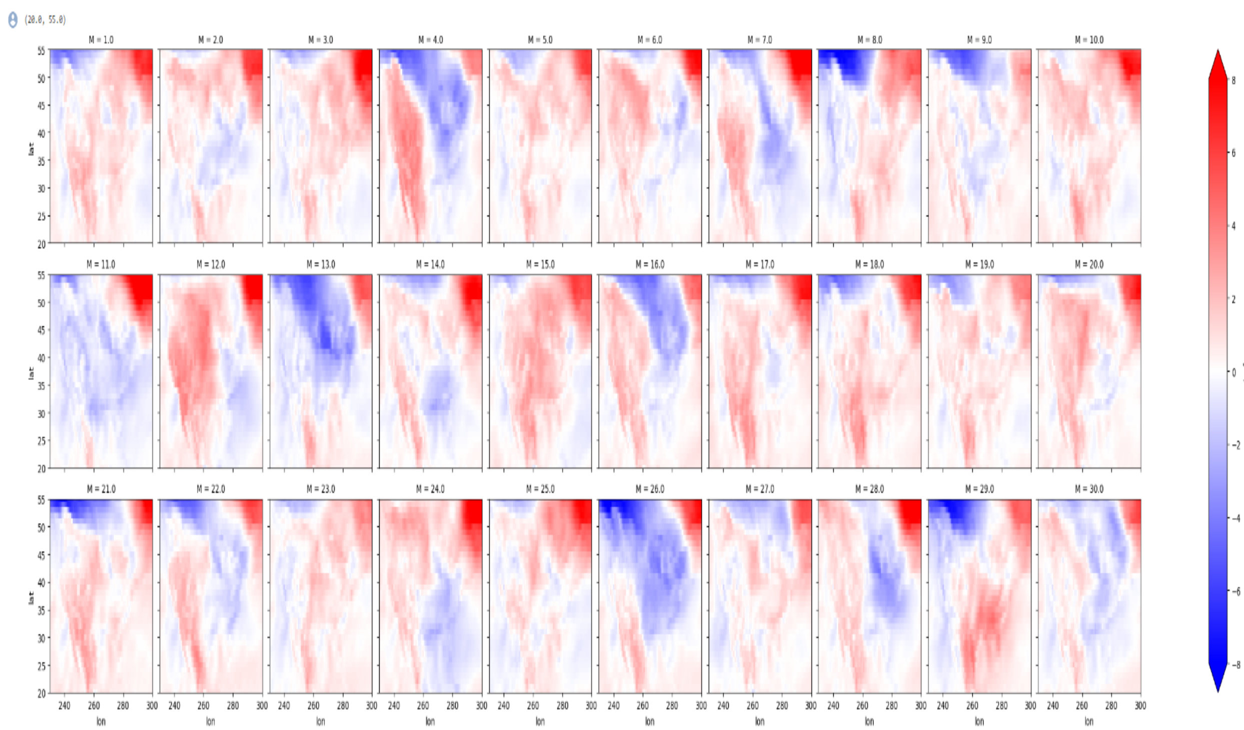

Fig 3.5: Seasonal Forecast (NMME)

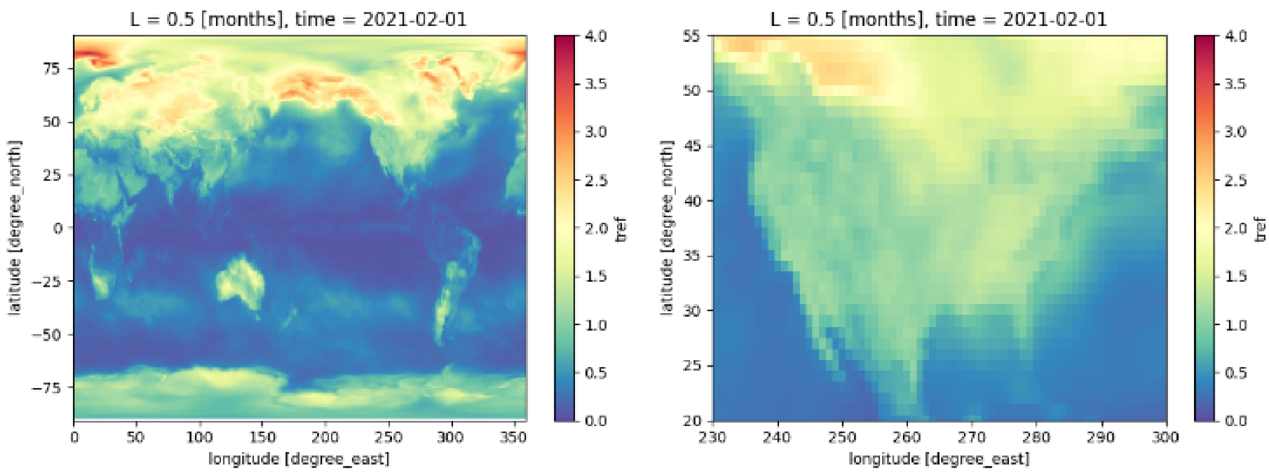

Fig 3.6: The standard deviation of forecasts

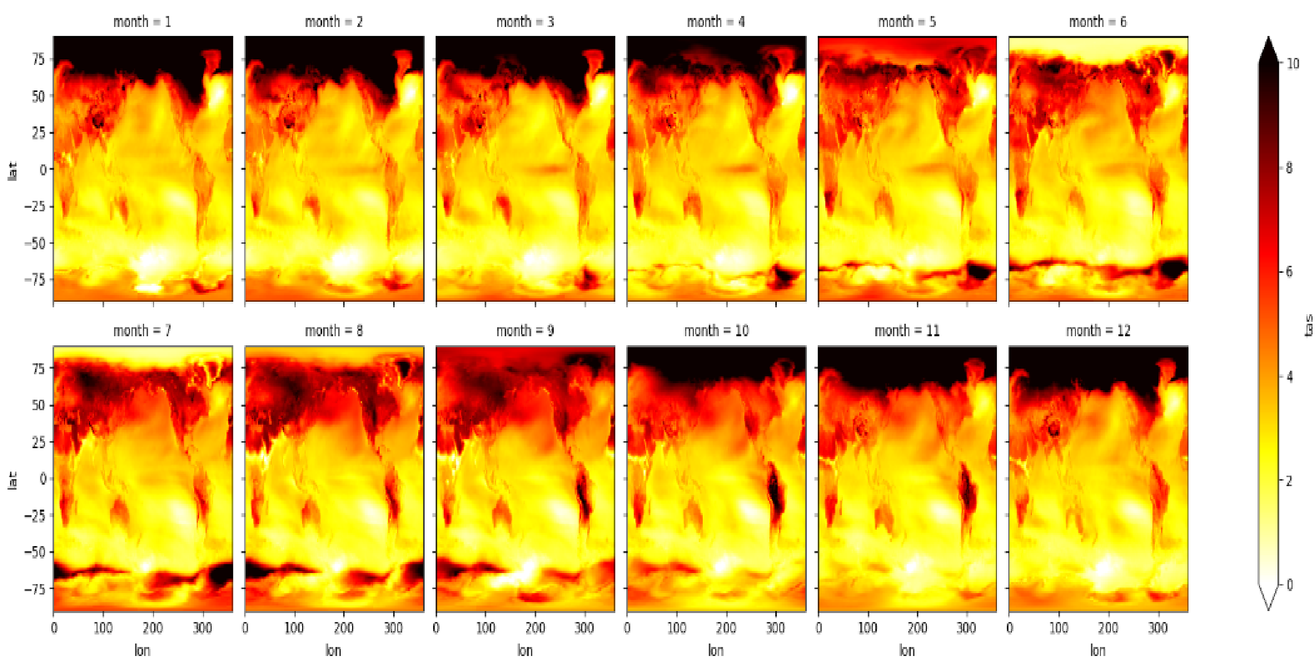

Fig 3.7: Climate Change Assessment (CMIP6)

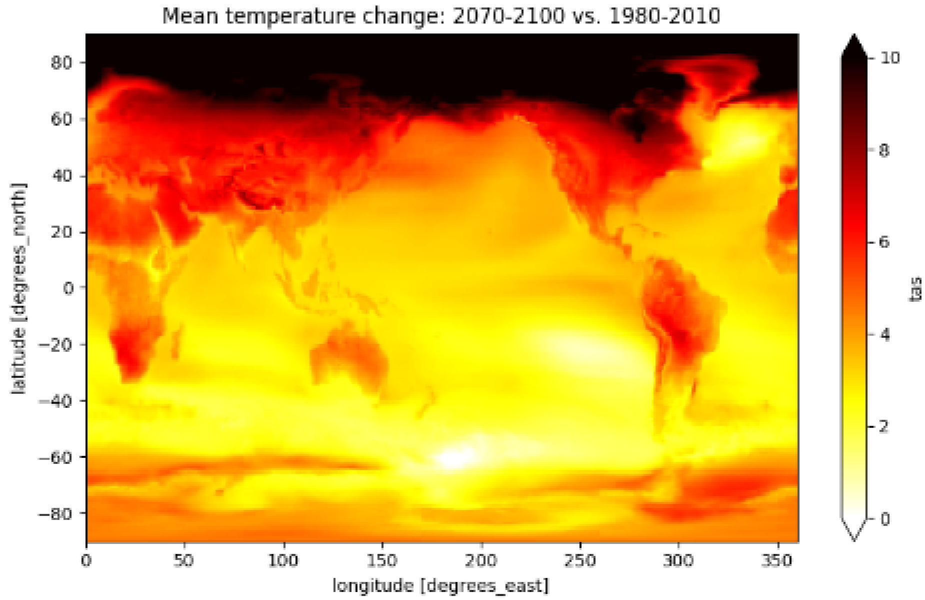

Fig 3.8: Mean temperature change: 2070-2100 vs. 1980-2010

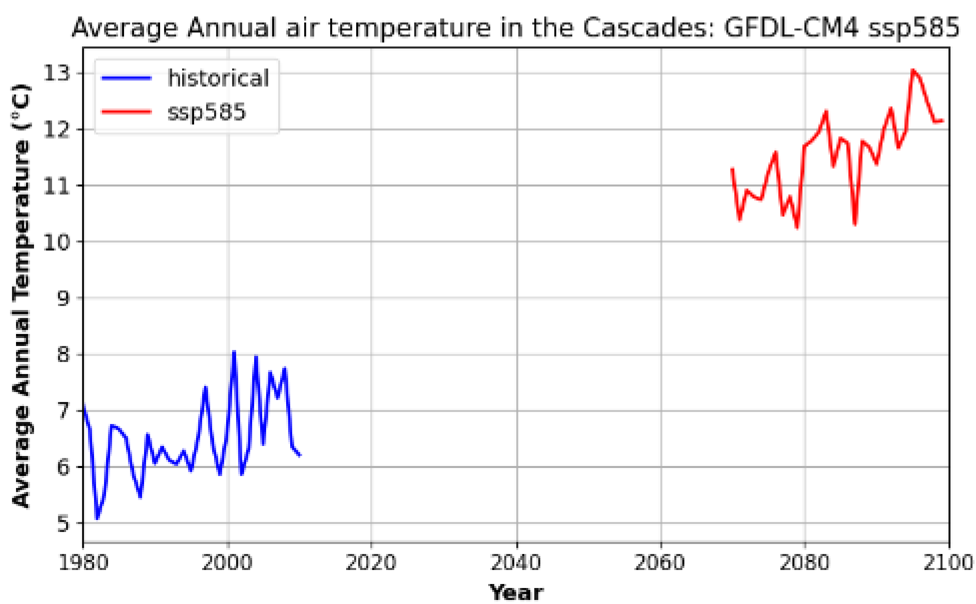

Fig 3.9: Average Annual air temperature in the Cascades: GFDL-CM4 ssp585

6. Results

We created a simple climate change data analysis program written in Python that allows analyzing multidimensional datasets that have data corresponding to specific latitudes, longitudes, and time using python.. The program uses various datasets such as Daily gridded observation data (NOAA CPC & GridMET), Monthly climate forecast data (NMME), Climate change projection data (CMIP6).

In this project we analyze gridded precipitation data (from CPC) to extract daily data, find monthly totals, find spatial average of precipitation in a given domain, plot the results, and we save the outputs as netcdf files. The project discusses different data formats commonly used in climate science such as NetCDF. Grib, HDF, Tiff, CSV or JSON and shp We have worked with xarray functionalities which is a powerful python library for analyzing geospatial data. We open data for two years and concatenate them into one file.

We extracted the monthly precipitation sum and made a simple plot of the months. Using libraries we develop a more personalized plot for all the 12 months. We have done the project in three different parts. The first one is the basic climate change analysis, then we have intermediate climate change analysis and last but not the least advanced climate change analysis.

7. Conclusion

Climate change data analysis in big data analytics is an essential field that aids in understanding, mitigating, and adapting to the effects of climate change. The growing volume and complexity of climate data necessitate the use of advanced technologies and analytical methods to make informed decisions and take meaningful actions to combat this global crisis.The importance of this area cannot be overstated, as it equips us with the knowledge and tools necessary to address one of the most pressing challenges of our time. By harnessing the power of big data technologies, climate change data analysis enables us to gain valuable insights about climate patterns, trends, and potential consequences, leading to a better understanding of the complex dynamics of our planet's climate system. Decision-makers at various levels, from governments and organizations to individuals, can make informed choices about climate mitigation and adaptation strategies based on data-driven evidence. Climate models and real-time data analysis allow us to predict the potential impacts of climate change, helping communities and nations prepare for extreme weather events, sea-level rise, and other challenges. It aids in the development of sustainable policies and practices, from renewable energy strategies to conservation efforts aimed at preserving ecosystems.

However, the field of climate change data analysis is not without its challenges. Data quality issues, integration complexities, and the need for interdisciplinary expertise make it a multifaceted endeavor. But, with continuous advancements in technology, machine learning, and data visualization, these challenges can be addressed.

In a world where the consequences of climate change are increasingly evident and pressing, the role of big data analytics in climate change data analysis cannot be overstated. It provides us with the means to take action, make informed decisions, and strive for a more sustainable future. As we move forward, it is crucial to continue investing in research, technology, and collaboration to ensure that climate data analysis remains a central pillar in our collective efforts to address climate change effectively and safeguard our planet for future generations.

References

[1] " *Climate Data Guide Climate Data Analysis Tools & Methods" | Climate Data Guide*. Available at: https://climatedataguide.ucar.edu/climate-tools (Accessed: 31 October 2023).

[2] "*Natural Resources Conservation Service Climatic Data Elements"*. Available at: https://www.nrcs.usda.gov/wps/portal/wcc/home/climateSupport/fieldOfficeGuide/climaticDataElements (Accessed: 31 October 2023).

[3] "Earth, B. (2017) *Climate change: Earth surface temperature data*, *Kaggle"*. Available at: https://www.kaggle.com/datasets/berkeleyearth/climate-change-earth-surface-temperature-data (Accessed: 31 October 2023).

[4] "*Climate Data Guide Statistical & Diagnostic Methods Overview" | Climate Data Guide*. Available at: https://climatedataguide.ucar.edu/climate-tools/statistical-diagnostic-methods-overview (Accessed: 31 October 2023).

[5] "*Community Data Analysis Tools CDAT"*. Available at: https://cdat.llnl.gov/ (Accessed: 31 October 2023).

[6]"*Climate change analysis, tools, and software RTI"*. Available at: https://www.rti.org/focus-area/climate-change-analysis-tools-and-software (Accessed: 31 October 2023).