Relation between the Sun magnetic activities and their potential influence on Earth's Climate.

Shreyansh Parihar

201684933

David Hughes

Submitted in accordance with the requirements for the
module MATH5872M: Dissertation in Data Science and Analytics
as part of the degree of

Master of Science in Data Science and Analytics

The University of Leeds, School of Mathematics

September 2023

The candidate confirms that the work submitted is his/her own and that appropriate credit has been given where reference has been made to the work of others.

UNIVERSITY OF LEEDS

School of Mathematics

FACULTY OF ENGINEERING AND PHYSICAL SCIENCES

Academic integrity statement

I am aware that the University defines plagiarism as presenting someone else's work, in whole or in part, as your own. Work means any intellectual output, and typically includes text, data, images, sound or performance.

I promise that in the attached submission I have not presented anyone else's work, in whole or in part, as my own and I have not colluded with others in the preparation of this work. Where I have taken advantage of the work of others, I have given full acknowledgement. I have not resubmitted my own work or part thereof without specific written permission to do so from the University staff concerned when any of this work has been or is being submitted for marks or credits even if in a different module or for a different qualification or completed prior to entry to the University. I have read and understood the University's published rules on plagiarism and also any more detailed rules specified at School or module level. I know that if I commit plagiarism I can be expelled from the University and that it is my responsibility to be aware of the University's regulations on plagiarism and their importance.

I re-confirm my consent to the University copying and distributing any or all of my work in any form and using third parties (who may be based outside the EU/EEA) to monitor breaches of regulations, to verify whether my work contains plagiarised material, and for quality assurance purposes. I confirm that I have declared all mitigating circumstances that may be relevant to the assessment of this piece of work and that I wish to have taken into account. I am aware of the University's policy on mitigation and the School's procedures for the submission of statements and evidence of mitigation. I am aware of the penalties imposed for the late submission of coursework.

Name: Shreyansh Parihar

Student ID: 201684933

Abstract

The Sun's magnetic activities have long been speculated to influence Earth's climate. This dissertation explores the potential correlations between solar phenomena—specifically, cosmic rays, Solar Spectral Irradiance (SSI), and sunspots—and cloud properties, aiming to elucidate their impacts on Earth's climatic systems.

Through rigorous data analysis, the study found varying degrees of association between these solar variables and cloud parameters. UV SSI exhibited significant positive correlations with certain cloud properties, hinting at its potential influence on cloud formation. Visible SSI demonstrated a strong relationship with cloud albedo in the visible range. In contrast, Infrared SSI presented a more complex interplay with cloud properties. Sunspot activity showed a nuanced relationship with cloud parameters, with some properties indicating weak correlations, while others exhibited minimal to no associations. Cosmic rays, meanwhile, displayed intriguing patterns of correlation with cloud properties, suggesting their potential role in influencing specific cloud parameters.

When compared with previous research, the findings of this study both reinforce and provide fresh perspectives on the sun-climate connection. While certain associations support existing knowledge, others underscore the need for further exploration. The research emphasizes the multifaceted nature of solar influences on Earth's climate, advocating for comprehensive climate models that integrate these solar variables.

This dissertation serves as a stepping stone in the quest to fully understand the sun's role in Earth's climate dynamics, highlighting areas for future investigation and emphasizing the significance of the sun-climate nexus in the broader context of climate science.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The Sun, as the central star of our solar system, has tremendous influence in our solar system, it constitutes about 99.86% mass of solar system, and it is about 100 times the size of earth. The temperature of the sun is about 27 million degrees Fahrenheit which is too hot for any living creature to go even near the sun. With such high temperatures, it is a good thing that the earth is 93 million miles away from the sun. Despite such a huge distance (which is not that huge if we go deep in solar objects relation) sun has a major role for life on earth, the heat of sun providing almost perfect temperatures for various kinds of life on the earth, the perfect amount of light for plants and foods to grow, the sun's gravity which holds our solar system together are all essential to us. (NASA,2021b)

Apart from heat, light and gravity the sun also emits other type of energies, but earth's magnetic field shields us from most of these energies, but not every time, there are always certain energies which penetrates the earth's magnetic field and enter the earth's surface. Figure 1.1 shows the energy from where the sun emits the energies what are its effects on not just earth but various other stars and planets in our solar system.

As shown, the Sun's energy has profound implications for Earth. But what's particularly intriguing, and the focus of this dissertation, is the potential climatic impacts stemming from the Sun's magnetic activities. These activities aren't monolithic; they are diverse, encompassing phenomena like sunspots, Solar Spectral Irradiance (SSI), and cosmic ray flux. Let's delve deeper into these phenomena, as they've captivated scientists for their potential connections to Earth's climatic oscillations.

Sunspots, cooler patches on the Sun's surface caused by intense magnetic activity, have always been a topic of keen interest. Between 1980 and 1993, studies noted that the Combined Sunspot Number (CSN) fluctuated between 0 to 150, with a consequential change in the solar constant (Solar constant is the amount of solar electromagnetic radiation received per unit area at the outer atmosphere of Earth, approximately 1361 W/m^2.) of approximately ±0.05% (Wade, P,1995). Intriguingly, during sunspot minima, Earth is bombarded with heightened cosmic rays, which seem to foster increased cloud formation, consequently nudging the surface temperatures downward (briangwilliams,2023).

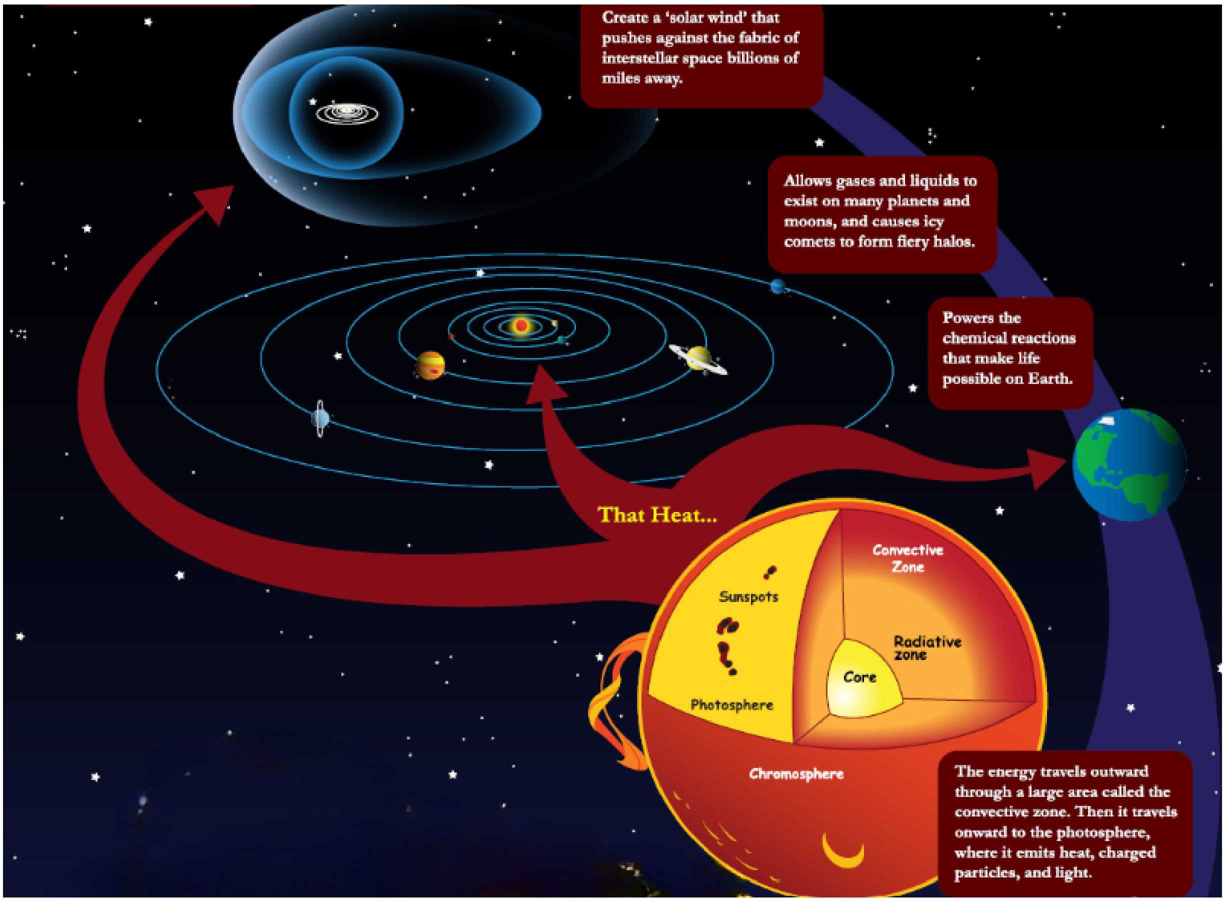

Figure 1.1: The Sun's energy and its implications for Earth. Source: (NASA,2021a)

Solar Spectral Irradiance (SSI) describes the Sun's energy received across different wavelengths. This energy undergoes variations, influenced by the interplay of sunspots and faculae (Faculae are bright spots on the Sun's surface often seen near sunspots, associated with strong magnetic fields.). A notable study, reevaluating the solar constant based on a 42-year total solar irradiance (TSI) time series, embarked on a comprehensive analysis of various sources of spaceborne TSI observations (Gueymard, C.A,2018). This study focus how solar activities, especially sunspots and faculae, act as dual forces. While sunspots darken and reduce solar output, faculae brighten it.

Cosmic rays, high-energy particles originating from distant galactic sources, have a curious relationship with solar activities. Notably, a negative correlation exists between the solar constant and cosmic ray flux, with the relationship being especially pronounced on both monthly and yearly data scales (neutronm,2022). This inverse relationship implies that as solar activities surge, the cosmic ray flux reaching Earth dwindles. But what makes cosmic rays particularly relevant to our climate discourse is their potential role in cloud formation (ScienceDaily,2016).

Sunspots, with their cyclical nature, hold potential clues about Earth's climatic rhythms.

Historical data reveals intriguing patterns. Such patterns prompt questions about the quantitative relationships between sunspots, faculae, and Earth's climate (briangwilliams,2023).

Relationship of the Sun's magnetic activities and Earth's climate is a tapestry of multiple threads. Each thread, be it sunspots, SSI, or cosmic rays, adds layers of depth and complexity to the story. While significant efforts have been made in understanding these phenomena, much remains in the realm of the unknown.

1.0.1 Sun-Climate Connection and its Nuances

While we recognize that changes in solar irradiance can influence Earth's radiation budget and, consequently, temperature, the exact magnitude and interplay of these effects in the context of an ever-changing climate system remain topics of active research (Nap,2022).

1.0.2 Cosmic Rays - More than Just High-Energy Particles

The impact of cosmic rays on Earth's climate is multi-faceted. As we delve deeper into the correlation between observed sunspot numbers and cosmic ray flux, the potential of cosmic rays as agents of climate change emerges (Nap,2022).

1.0.3 Solar Ultraviolet Radiation and its Atmospheric Impacts

Although the ultraviolet spectrum contributes minimally to total solar irradiance, its variability across the solar cycle can significantly influence the energy dynamics of Earth's middle and upper atmospheres (Nap,2022).

Despite the wealth of information garnered over the years, gaps in our understanding persist. These gaps aren't merely academic challenges; they represent opportunities. Opportunities to refine our methodologies, to foster interdisciplinary collaborations, and to ultimately gain a holistic understanding of the sun-climate relationship. Addressing these gaps is not just a scientific necessity but is crucial for shaping informed climate policies and strategies for a world grappling with the challenges of climate change.

In weaving together, the intricate narrative of the Sun's magnetic activities and Earth's climate, we dive in a journey through time and space, exploring phenomena that shape our world in ways both subtle and profound. This dissertation seeks to illuminate these connections, to unravel the mysteries, and to chart a path forward, grounded in rigorous scientific inquiry and a quest for truth. As we stand on the changing world of new discoveries, it's evident that the Sun, in all its blazing glory, holds the keys to many of the climatic puzzles we seek to solve.

1.1 Rationale

As climate change continues to present significant challenges to humanity, uncovering all the potential contributing factors is essential. Understanding the relationship between solar magnetic activities and climate change could lead to new insights into the underlying mechanisms of climate variation. While human-induced factors remain predominant in climate change discussions, the role of solar activities might be an essential piece of the puzzle. Investigating this relationship can broaden our comprehension of climate systems, contribute to more accurate climate modeling, and might even have implications for policy and mitigation strategies.

This analysis aims to delve deep into unraveling the interactions between solar magnetic activities and climate change. By establishing correlations, comparing with previous studies, and discerning potential mechanisms, it aims to contribute to the broader understanding of Earth's climate system and its sensitivity to external influences. Insights gained from this study could have significant implications for climate science, policy-making, and our collective efforts to mitigate and adapt to climate change.

1.2 Research Questions

The following research questions will guide this study:

- How have solar magnetic activities, including sunspots, Solar Spectrum Irradiance, and cosmic rays, varied over time?
- What correlations or patterns exist between sunspots, Solar Spectrum Irradiance, and cosmic rays?
- What correlations or patterns exist between solar magnetic activities and between climate change data?
- How do the findings align with or differ from existing research on the subject?

1.3 Objectives

The primary objectives of this dissertation are to:

1. Analyze individual datasets of solar magnetic activities.
2. For individual datasets perform Stationarity Analysis (Detailed explanation in Data Analysis chapter).
3. Investigate correlations between various solar magnetic data.
4. Compare identified correlations with specific climate data.
5. Evaluate the implications of the study's results within the broader context of climate science.

1.4 Scope

This dissertation focuses on the quantitative analysis of solar magnetic activities and their potential connections to climate change.

- Sunspot data, spanning several solar cycles, will be examined to identify patterns and trends.
- Solar Spectral Irradiance (SSI) data across different spectral components will be studied in conjunction with relevant climate variables, specifically cloud-related parameters.
- Solar Spectral Irradiance (SSI) data across different spectral components will be studied in conjunction with relevant climate variables, specifically cloud-related parameters.
- Cosmic ray flux data will be correlated with solar activity indices.

The analysis will involve quantitative statistical techniques, time series analysis, and visualisation methods to discern significant trends, relationships, and potential causal links.

1.5 Structure of the Dissertation

The dissertation is structured into six main sections:

1. Introduction: Outline of the study and Exploration of existing research.
2. Methodology: Description of methods.
3. Data Analysis: Examination of solar magnetic variables and climate variables.
4. Correlation Analysis: Analysis of relationships between variables and Comparison with climate data.
5. Interpretation and Findings: Interpretation of results.
6. Conclusion: Summary, implications, and future recommendations.

Understanding the relationship between solar magnetic activities and climate change is a vital yet complex endeavor. This dissertation seeks to contribute valuable insights into this area, focusing on the methodical examination of solar phenomena and their potential links to Earth's climate. The subsequent sections will delve into a rigorous exploration of these topics, aiming to shed light on a subject of immense scientific and societal importance.

Chapter 2

Methodology

2.1 Data Sources

2.1.1 Cosmic Ray Data

- Source: Oulu Cosmic Ray Station. (Oulu,2022)
- Description: Cosmic ray data was collected from the Oulu Cosmic Ray Station, which provides real-time information about cosmic ray intensities. This included details about the energy, intensity, and time variations of cosmic rays.
- Application: Understanding cosmic ray patterns is vital in analyzing their potential interactions with solar magnetic activities and their subsequent effects on climate.

2.1.2 Detailed Cosmic Rays Properties

- Source: Cosmic Ray Database at LPSC. (Lpsc,2022)
- Description: This database offers a comprehensive collection of cosmic ray properties, including their composition, energy spectra, and sources. It allows for a more detailed analysis of cosmic rays and their behaviors.
- Application: These properties were essential in correlating cosmic ray activities with solar activities and understanding how they may influence climate patterns.

2.1.3 Solar Irradiance Data

- Source: Laboratory for Atmospheric and Space Physics (LASP). (Lasp,2022)
- Description: The solar irradiance data includes measurements of the sun's electromagnetic radiation. It provides insights into the variations in solar energy reaching the Earth's surface, including ultraviolet, visible, and infrared spectra. Different instruments were used to measure different spectrum ranges. Instrument calculation details:

- XPS: 0-40 nm
- SOLSTICE: 115-310 nm
- SIM: 240-2416 nm

- Application: Understanding solar irradiance is crucial for modeling the sun's influence on climate and weather patterns.

2.1.4 Sunspot Data

- Source: Royal Observatory of Belgium's Solar Influences Data Analysis Center (SIDC). (SIDC,2022)
- Description: The sunspot data consists of information on sunspots' number, size, location, and magnetic properties. This data is essential for tracking solar cycles and magnetic activities.
- Application: Analyzing sunspots helps in understanding solar magnetic activities and their potential influence on cosmic rays and climate.

2.1.5 Cloud Data

- Source: ESA Climate Change Initiative (CCI). (CEDA,2022)
- Description: The cloud cover data provides information on cloud formations, types, altitudes, and optical properties across different regions and time periods.
- Application: Analyzing cloud data is essential in understanding the relationship between solar activities and cloud cover, a significant factor in climate change.

2.2 Data Analysis Methods

2.2.1 Individual Data Analysis

Cosmic Ray Data Analysis

The cosmic ray data was subject to detailed statistical analysis to discern trends, patterns, and variations over time. Techniques such as time-series decomposition were used to understand the seasonal and cyclical patterns.

Solar Irradiance Analysis

Solar irradiance data was thoroughly analysed to understand the changes in solar energy reaching the Earth. This included an examination of daily, monthly, and yearly variations.

Sunspot Data Analysis

The sunspot numbers were analyzed to correlate with solar cycles and magnetic activities. Analyzing the number, location, and size of sunspots provided insights into solar magnetic fields.numbers and solar cycles.

Cloud Data Analysis

The Earth's climate is a complex system influenced by a myriad of factors, one of which is solar activity. While the relationship between solar activity and Earth's climate has been studied extensively, the role of cloud cover as a mediating factor is not yet fully understood. Clouds play a significant role in regulating Earth's temperature by reflecting solar radiation back into space and trapping heat in the atmosphere. Given the vital role that clouds play in Earth's thermal regulation, understanding how solar activities influence cloud formation and characteristics can provide valuable insights into climate change mechanisms.

2.2.2 Correlation Analysis

Cosmic Ray and Solar Activities

Correlation analysis between cosmic rays and solar activities such as solar irradiance and sunspots was conducted. This helped in understanding how cosmic rays are influenced by solar magnetic fields.

Solar Activities and Climate Variables

Correlations were analysed between solar activities and climate variables such as temperature, precipitation, and cloud cover. This helped in identifying potential connections between solar magnetic activities and climate changes.

2.2.3 Climate Change Analysis

Modeling Solar Influence

A comprehensive climate change analysis was performed by modeling the potential impacts of solar activities on climatic factors.

2.3 Tools and Software Used

- **Python**: The main programming language used for the analysis.
- **Pandas**: A popular library for data manipulation and analysis in Python. It was used to handle and process the datasets, as evidenced by the import statement `import pandas as pd`.

- **NumPy**: A library used for numerical operations in Python. It provides support for arrays, matrices, and mathematical functions, as seen in the import statement `import numpy as np`.
- **Xarray**: A python library to load .nc data files type (Detail discussion in Data Analysis chapter).
- **Matplotlib**: A python library for visualizing the data.
- **Seaborn**: Similary like Matplotlib library, used for visualizing the data but has more options than former.
- **statsmodels**: Python library capable of conducting statisical analysis and tests, we used specifically to perform Augmented Dickey-Fuller test.

Chapter 3

Data Analysis

3.1 SSI across Different Spectra

3.1.1 Data Cleaning and Transformation

The analysis begins with the acquiring and preparing the data related to Solar Spectral Irradiance (SSI). The dataset is in file names 'solarIridence.txt', it was imported using the Python library, pandas. Upon loading the dataset, structure and contents were examined, the SSI data contains below essential columns/attributes:

- *nominal_date_yyyymmdd*: Represents the date in the format YYYYMMDD.
- *min_wavelength*: The minimum wavelength in the spectrum.
- *max_wavelength*: The maximum wavelength in the spectrum.
- *irradiance*: The measure of solar irradiance.
- *irradiance_uncertainty*: Uncertainty associated with the irradiance measurement.
- *quality*: A flag indicating the quality of the data.

Let's understand some of the column in detail. *Nominal_date_yyymmdd* has dates between the years 2003 – 2020 including them.

As described earlier in methodology for SSI, different instruments were used to calculate different wavelengths, thus the *min_wavelength* and *max_wavelength* can have values ranging from 0 – 2416 nm, where nm is nano meters, the unit of measurement.

Quality flag can have different meaning based on instruments, in summary for all instrument if first bit is 1 (example of first bit -¿ e.g. data value = 1.0, 1 is first bit and 0 is second bit), that indicates values are missing. For SIM instrument specifically, if values are other than that, then it represents some time-dependent warning or data quality. For XPS, if the first 2 bits are 0 means good, 1 – missing, similarly for SOLSTICE.

Irradiance has values between 0 – 2.11 W/m^2/nm, where W/m^2/nm is the unit of measurement.

Irradiance uncertainty has values between 0 – 0.26, this is the uncertainty in irradiance due to temperature, atmospheric pressure, and other phenomena. The irradiance values in irradiance column are corrected based on these uncertainty values. Irradiance uncertainty will have the same unit of measurement as Irradiance column.

As we see data will not have any null values in it, as the data source has given us the information that quality column with flag value of 1.0 is considered missing data. To confirm the claim, we can see below table to see the count of empty/null values in each column, and we observer that it is mostly empty, the values are taken from python code:

Column Name	**Null/Empty values**
nominal_date_yyyymmdd	0
min_wavelength	0
max_wavelength	0
Irradiance	0
irradiance_uncertainty	0
quality	0

Table 3.1: Null value count in each column

Out of 5689117 entries, it was found out that 1360 rows were having quality flag as 1.0 as displayed in figure 3.1.

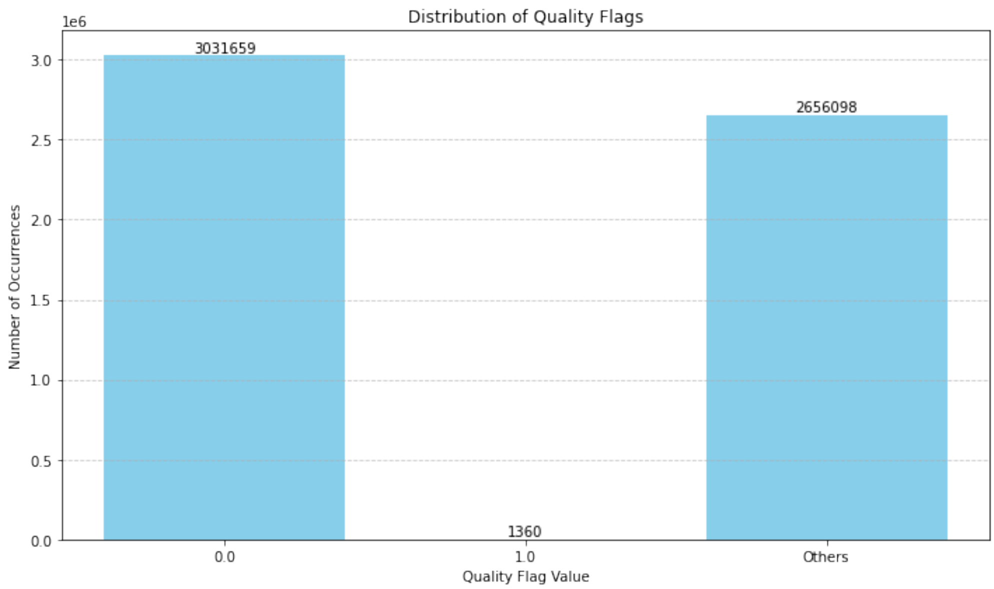

Figure 3.1: Missing values in SSI columns

Thus, as part of the data cleaning process these rows were removed as they will not have relevant SSI data to study.

By default, python assigned date column - *nominal_date_yyyymmdd* as object data type, instead we know that it is date-time column, thus, this column is modified to datetime data type column using pandas ‘to_datetime()’ function. This will also make things easier code wise when doing time-series analysis. Further, the data is also divided into 3 sections based on *min_wavelength* column for UV, Infrared and Visible ranges. Let’s start with UV SSI data.

3.1.2 In-depth analysis – UV SPECTRUM

Figure 3.2 shows the yearly mean of UV SSI from 2002 to 2020. This graph provides an annual

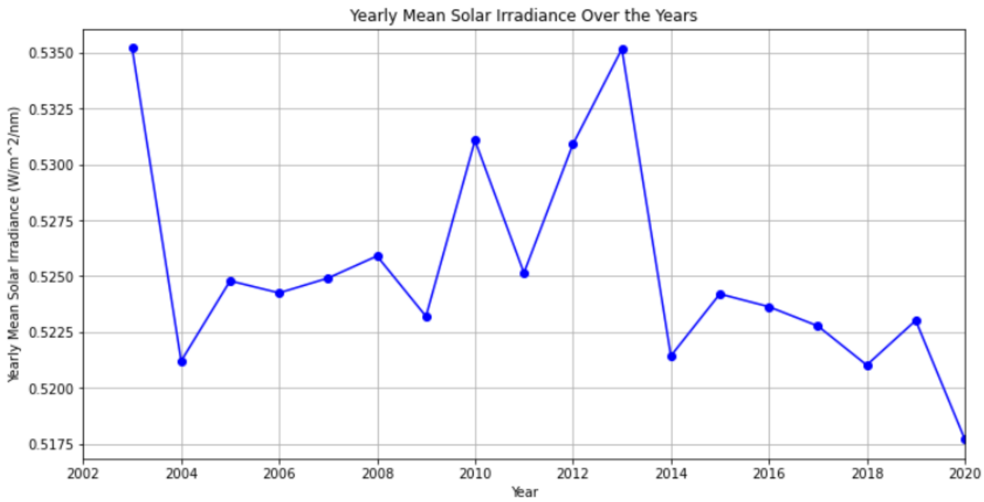

Figure 3.2: Yearly SSI (UV spectrum)

overview of UV SSI, revealing how irradiance levels have fluctuated over the years. The yearly trends exhibit subtle variations but remain within a specific range, further the stable nature of UV irradiance observed in the trend component. We can see that, majority of the data points are closely clustered, except for a few years where there are multiple peaks and troughs. Usually, at the end of solar cycles we see marginal dip in UV SSI, that can be observed in end of solar cycle 24 after 2008 and after solar cycle in 2019. The SSI also seem to be affected by high solar flares, as the year 2011 and 2014 were years with high solar flares and resulting in extreme variation in UV SSI specific around those years.

The below graphs are obtained using the decomposition method in time-series analysis. Figure 3.3 illustrates the decomposition of the UV spectrum SSI into its constituent components: the original time series, trend, seasonality, and residuals. This decomposition was carried out using the additive model, given by the equation:

$$\text{Observed} = \text{Trend} + \text{Seasonality} + \text{Residual}$$

Original Series: The original time series showcases fluctuating irradiance levels, but discerning any concrete patterns requires further decomposition.

Trend: The trend component reveals a relatively stable pattern in the UV spectrum. There

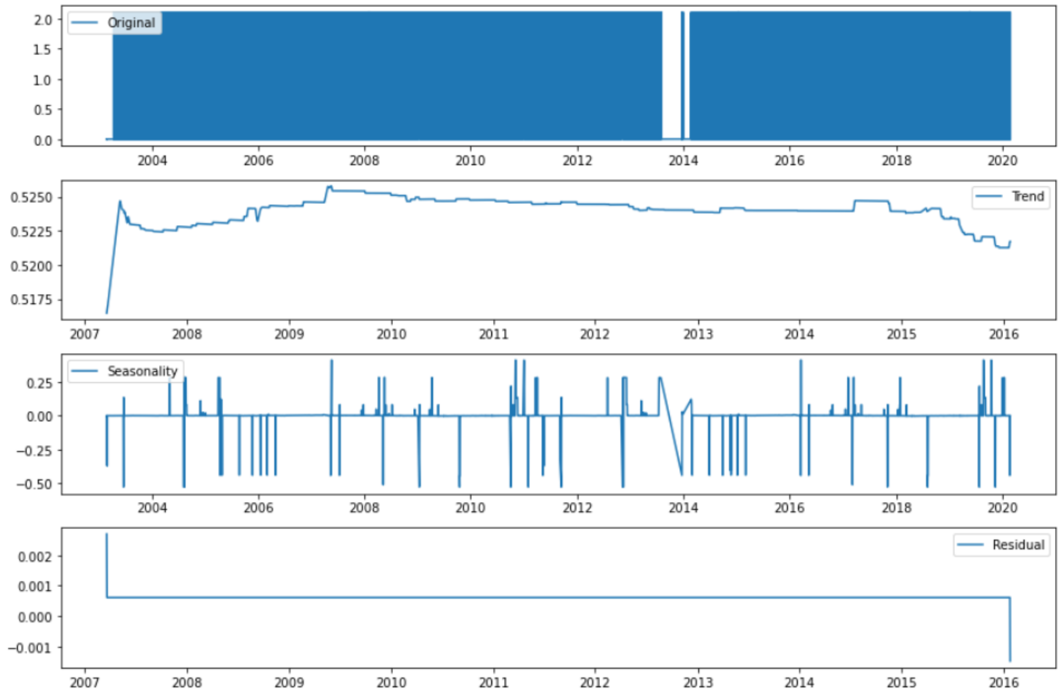

Figure 3.3: SSI (Infrared spectrum) time-series decomposition

are no substantial spikes or declines, indicating consistent solar activity in the UV range over the years.

Seasonality: The seasonal component captures periodic fluctuations, likely corresponding to solar cycles or Earth's orbital parameters.

Residual: The residuals represent noise or random variations that couldn't be attributed to trend or seasonality.

Now to confirm statistically with proper evidence, whether the given time-series is stationary or no-stationary, we need to perform stationarity analysis, but before we look into the results of UV data, let us first understand what stationarity analysis and specifically what method is used for this case to find stationarity of our time-series, below sub section discusses on that.

Stationarity Analysis and Rolling Statistics

The concept of stationarity in a time series implies that statistical properties such as mean, and variance remain constant over time. In simpler terms, a stationary time series does not exhibit trends or seasonality; it looks the same at any time point, similar to random noise. The Augmented Dickey-Fuller (ADF) test is a statistical procedure used to test whether a given time series is stationary or not. In the ADF test, the null hypothesis states that a unit root is present in the time series, meaning it is non-stationary. Conversely, the alternative hypothesis suggests that the time series is stationary or trend stationary. The test uses an autoregressive model and optimizes an information criterion across multiple lag values. Here's a simplified breakdown: ADF Statistic: This value is compared against critical values for different confidence intervals (usually 1%, 5%, and 10%) to determine the stationarity of the time series. A more negative ADF Statistic indicates stronger evidence against the null hypothesis, implying the time series is more likely to be stationary. P-value: A p-value below a certain threshold (usually 0.05) suggests that the null hypothesis can be rejected, thus concluding that the series is stationary. Critical Values: These are the test statistic values at which the null hypothesis can be rejected for different confidence levels.

In the case of the UV spectrum SSI time series, the Augmented Dickey-Fuller test yielded the following results:

- ADF Statistic: -14.56
- P-value: 4.87e-27
- Critical Values:
  - 1%: -3.43
  - 5%: -2.86
  - 10%: -2.57

The ADF Statistic of -14.56 is significantly more negative than all the critical values at 1%, 5%, and 10% confidence intervals. This implies strong evidence against the null hypothesis, confirming that the UV spectrum SSI time series is stationary.

Furthermore, the p-value is virtually zero (4.87e-27), which is well below the commonly used threshold of 0.05. This provides additional confirmation that the null hypothesis can be rejected, affirming the stationarity of the time series.

To supplement the ADF test, rolling statistics such as the rolling mean and standard deviation were also examined. Neither of these metrics showed any discernible trend over time, further supporting the stationary nature of the UV spectrum SSI time series

3.1.3 In-depth analysis – VISIBLE SPECTRUM

In this section we will investigate SSI data for visible spectrum. Figure 3.4 shows time-series decomposition for the same.

Let's discuss the main observations.

Trend The trend line remains relatively flat, although there is a very slow increase from 2010 likely due to end of solar cycle 24 in 2008 and a dip in 2016 which could be due to some solar event or cosmic event but all in all we will still call this trend as "flat", indicating a stable level of solar activity in the visible range over the years.

Seasonality The recurring peaks and troughs in this component indicate periodic fluctuations in the visible spectrum. These could be related to solar cycles, Earth's orbital mechanics,

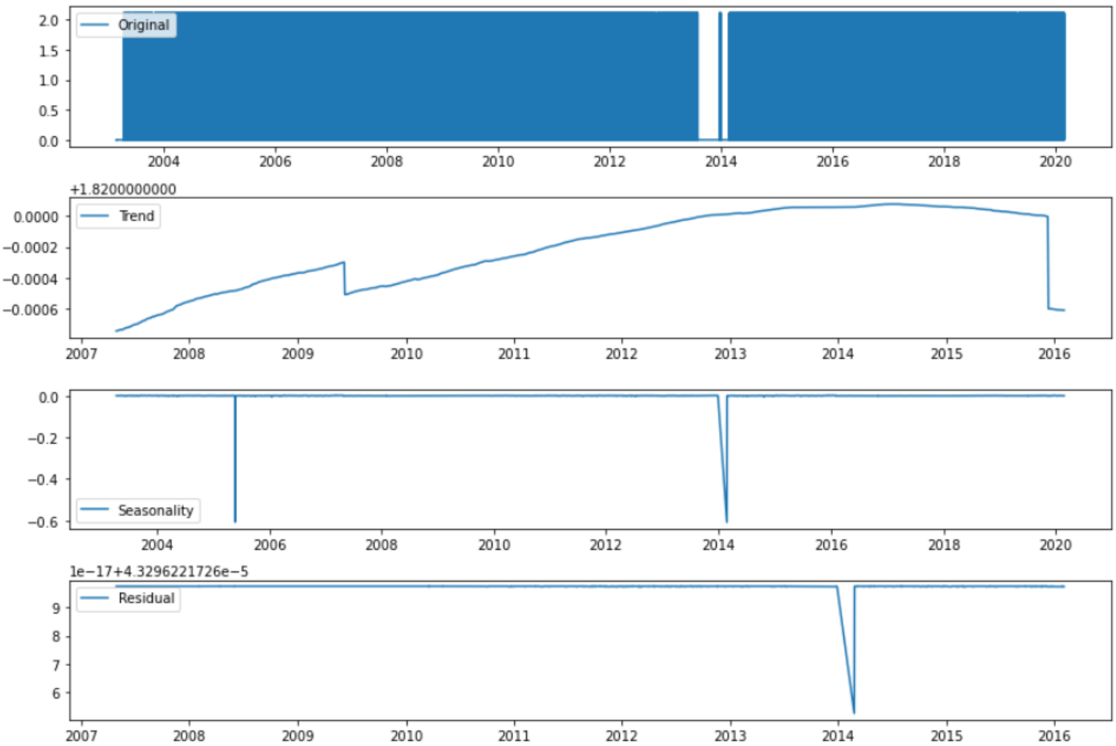

Figure 3.4: SSI (Infrared spectrum) time-series decomposition

or even short-term variations in solar activity. The periodic nature of these fluctuations is critical for models that aim to predict solar influences on Earth's systems based on historical data.

Residual Component The residuals, or the 'noise' in the data, appear to be random and do not show any discernible pattern. This is a good indicator that most of the systematic information in the data has been captured by the trend and seasonality components.

For the visible spectrum SSI, the ADF test results are as follows

- ADF Statistic: -76.10
- P-value: 0.0
- Critical Values:
  - 1%: -3.43
  - 5%: -2.86
  - 10%: -2.57

The ADF Statistic is -76.10, which is significantly lower than the critical values at all confidence intervals (1%, 5%, and 10%). This provides strong evidence against the null hypothesis and suggests that the time series is stationary. The p-value is 0.0, considerably below the commonly used threshold of 0.05, further affirming the rejection of the null hypothesis and confirming the series' stationarity. The visible spectrum analysis parallels what was observed in the UV range. The ADF test results, along with the decomposition components, suggest that the visible spectrum SSI is both stationary and stable over time. This is a significant observation, given that the visible spectrum has a direct impact on various Earth systems, from climate to photosynthesis in plants.

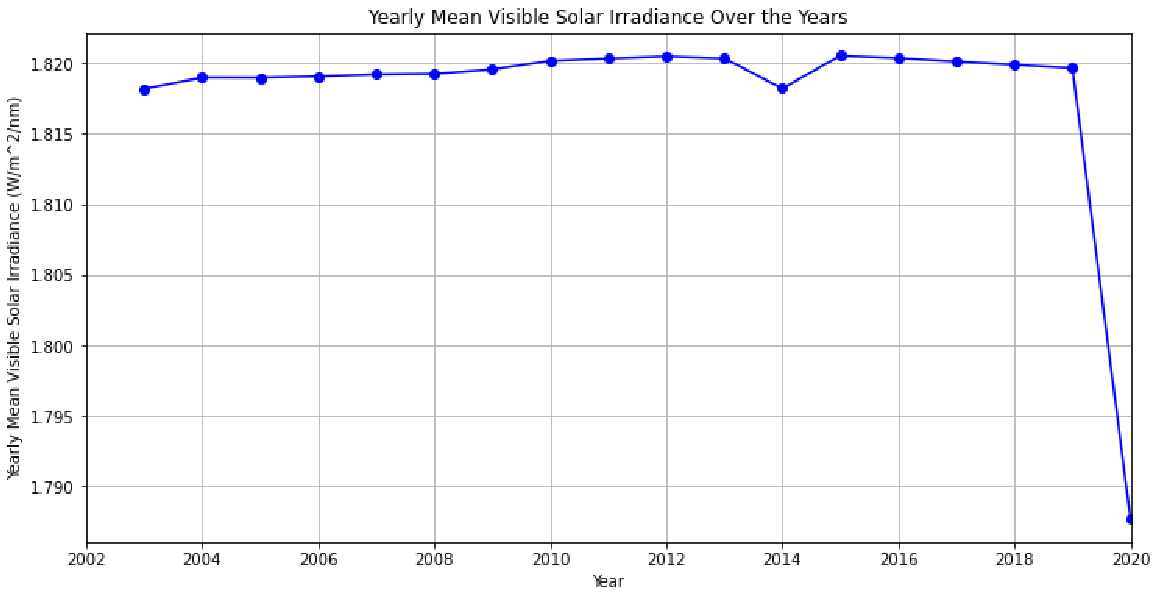

Figure 3.5: Yearly SSI (Visible spectrum)

Graph in figure 3.5 displaying the yearly trends in the visible spectrum serves as a critical piece of evidence for understanding long-term solar activities. While the UV spectrum gave us insights into the high-energy emissions from the Sun, the visible spectrum is particularly important as it accounts for the majority of solar energy that reaches the Earth's surface.

Stability Over Time Similar to the UV spectrum, the visible spectrum shows a stable range of irradiance levels over the years. This indicates that the Sun's emissions in this part of the spectrum have been fairly consistent, which is crucial information for climate modeling and long-term energy forecasts.

Baseline Irradiance The graph also provides a baseline level of irradiance in the visible spectrum, against which any future anomalies can be compared. Any significant deviations from this baseline would warrant further investigation into potential causes and implications. Unlike non-stationary time series where you may expect drastic changes or patterns, the uniformity across years further supports the stationarity of the time series. This is consistent with the Augmented Dickey-Fuller test results, which also pointed towards a stationary series. The graph doesn't display any sharp peaks or valleys (except for year - 2020), which would typically indicate sudden increases or decreases in solar activity. The absence of such features suggests that the solar emissions in the visible spectrum have been relatively calm and consistent. We observed that in 2020, there is a significant dip, which seems like an anomaly, when compared to other years. Understanding the cause of this anomaly would require a multidisciplinary approach, involving solar physics, atmospheric science, and instrumental calibration among other factors, which is out of scope of this dissertation.

In-depth analysis – INFRARED SPECTRUM

In this section we will investigate SSI data for infrared spectrum. Figure 3.6 is time-series decomposition for the same.

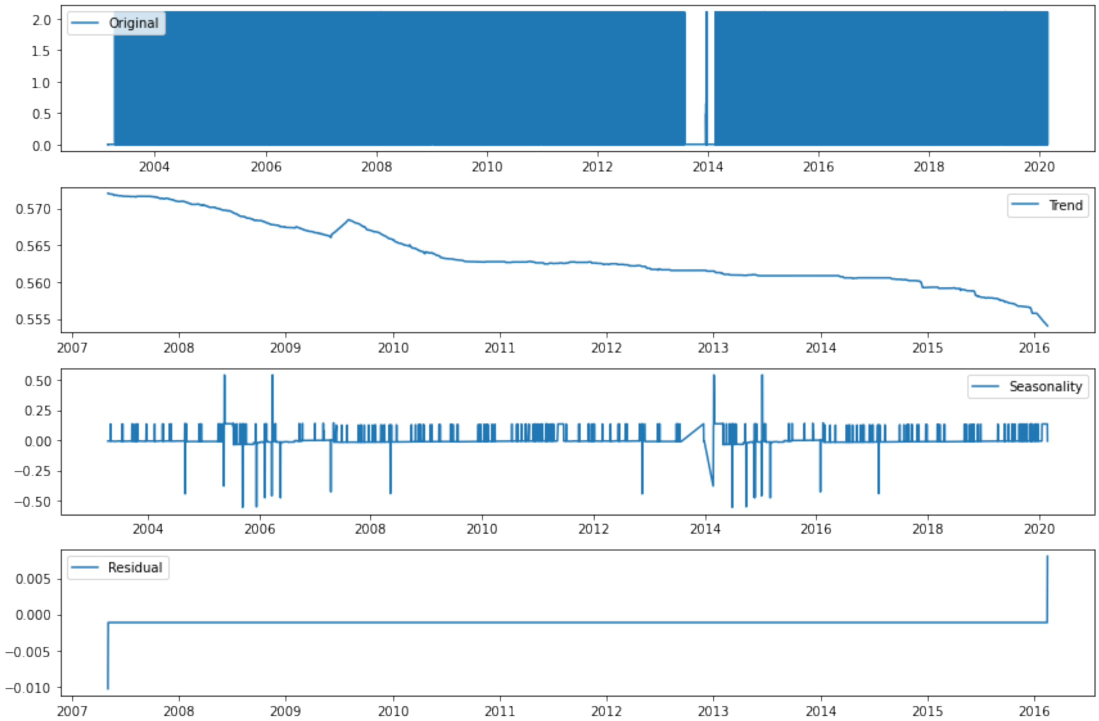

Figure 3.6: SSI (Infrared spectrum) time-series decomposition

From the above graph we can observe following

- **Trend Component**: The trend in the infrared spectrum appears to be relatively stable but not completely flat. There are slight variations, minor peaks, and valleys, which could be indicative of longer-term changes in solar activity or other factors. Unlike a perfectly flat trend, these variations suggest that while the Sun's infrared emissions have remained relatively stable, they are not entirely unchanging.

- **Seasonality Component**: The seasonality component shows consistent, periodic fluctuations. This periodicity could be tied to a variety of factors, such as the 11-year solar cycle, Earth's orbital mechanics, or transient solar phenomena.
- **Residual Component**: The randomness of the residual component supports the notion of a stationary series, we can confirm this from the ADF test.

Augmented Dickey-Fuller Test Results

- ADF Statistic: -7.92
- P-value: 3.73×10^{-12}
- Critical Values:
  - 1%: -3.43
  - 5%: -2.86
  - 10%: -2.57

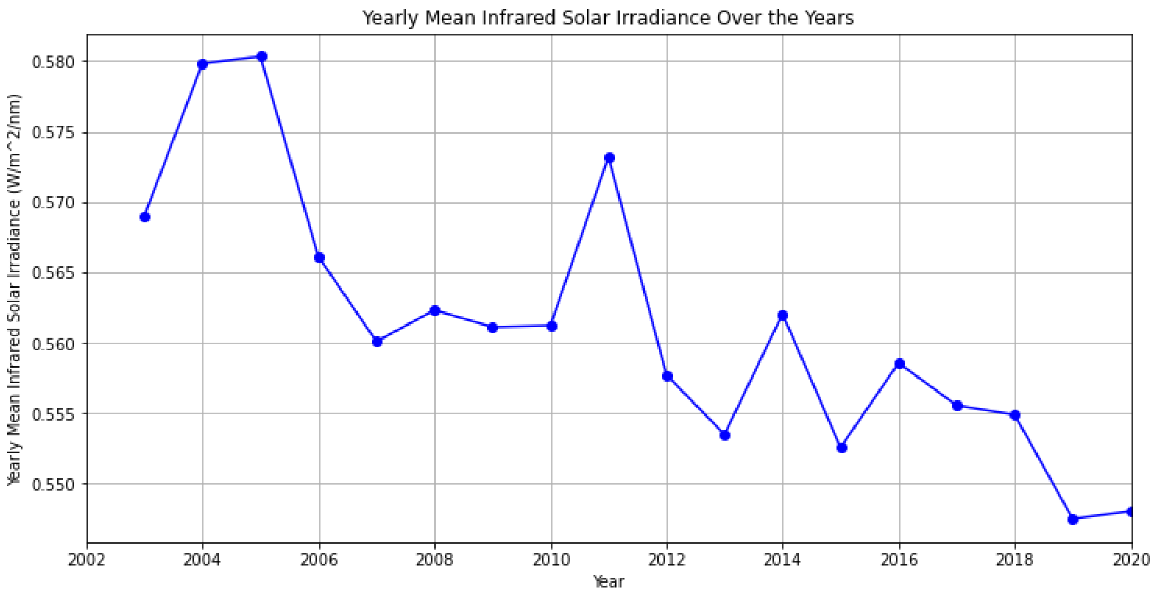

Figure 3.7: Yearly SSI (Infrared spectrum)

Figure 3.7 shows yearly trends in the infrared spectrum of Solar Spectral Irradiance (SSI). Several key features stand out in the graph:

The first thing to notice is the overall stability of the irradiance levels across the years. This complements what we observed in the trend component of the time series decomposition. However, it's important to note that this stability is constant by minor fluctuations, which could be indicative of subtle changes in solar activity or even Earth-based factors affecting the measurements.

While the data is generally stable, there are years where the irradiance values appear to deviate from the norm - In **early 2000s** there's a noticeable uptick in irradiance levels during the early 2000s, possibly corresponding to the peak of Solar Cycle 23. While in **Late 2010s** a mild dip is observed in the late 2010s, aligning with the minimum of Solar Cycle 24. Like in the visible spectrum, there's a noticeable dip in **2020**, which is not readily explained by known solar cycles or Earth-based factors, warranting further investigation.

3.1.4 Sunspot Activity

Data Cleaning and Transformation

Like with the SSI data, the data was loaded from the file named `SN_d_tot_V2.0.txt`, and then analyzed. The dataset spans from the year 1818 to the most recent month of 2023 at the time of downloading the file and comprises approximately 75026 rows.

Upon loading, the following columns were found in the dataset:

- **Year** – Calendar year, data ranging from the years 1818 to 2023.
- **Month** – Calendar month, with data for all months, i.e., 1-12.
- **Day** – Calendar day, with data for all days in a month, i.e., 1-31.
- **Date_in_fraction_of_year** – Date in fraction of the year, with data ranging from 1818.001 to 2023.412.
- **sunspot_number** – Daily total sunspot number. A value of -1 indicates that no number is available for that day (missing value). Observed sunspot numbers range from 0 to 528.
- **standard_deviation** – Daily standard deviation of the input sunspot numbers from individual stations. Data ranges from 0 to 77.70.

The columns in this dataset are quite self-explanatory from their names. As indicated in the definition of the sunspot number column, a value of -1 denotes missing values. As it can be seen in the figure , approximately 4% (3247 rows) of the data is missing.

Since our primary objective is to analyze the sunspot number from this dataset, we simply removed these rows as they did not contribute to the loss of information in any significant way.

Furthermore, we examined if any other columns had null or empty values:

Table 3.2 clearly indicates that there are no null values in our dataset, allowing us to proceed with formatting our data. To facilitate our analysis, we introduced a new calculated column from year, month, and day named `Date`, formatted as datetime. Other than that, Python correctly formatted all other columns, so no further adjustments were made to the dataset.

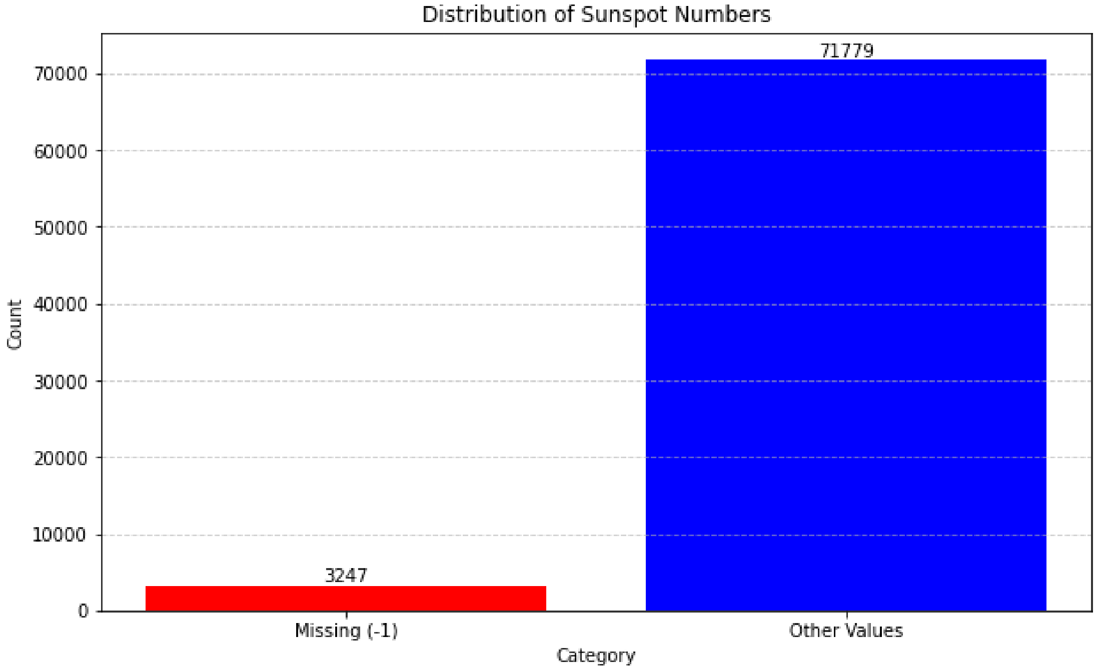

Figure 3.8: Missing values in Sunspot dataset

Column Name	Null/Empty values
Year	0
Month	0
Day	0
Date_in_fraction_of_year	0
sunspot_number	0
standard_deviation	0

Table 3.2: Null value count in Sunspots data columns

In-depth analysis

Yearly graph (figure 3.9) for sunspot activity presents a compelling narrative of how solar activity has changed over time. The wave-like pattern seen in the graph is characteristic of the well-known 11-year solar cycle. This cycle is a result of the Sun's magnetic activity, which also influences Earth's climate, magnetic field, and even technological systems like satellites and power grids.

From figure 3.10, which represents the time series decomposition of sunspot data, we can derive several crucial insights:

- **Trend:** The trend captures long-term changes in sunspot activity, potentially linked to lengthier solar cycles or other forms of solar variability. Although generally consistent, certain years show a significant deviation, either rising or falling sharply.

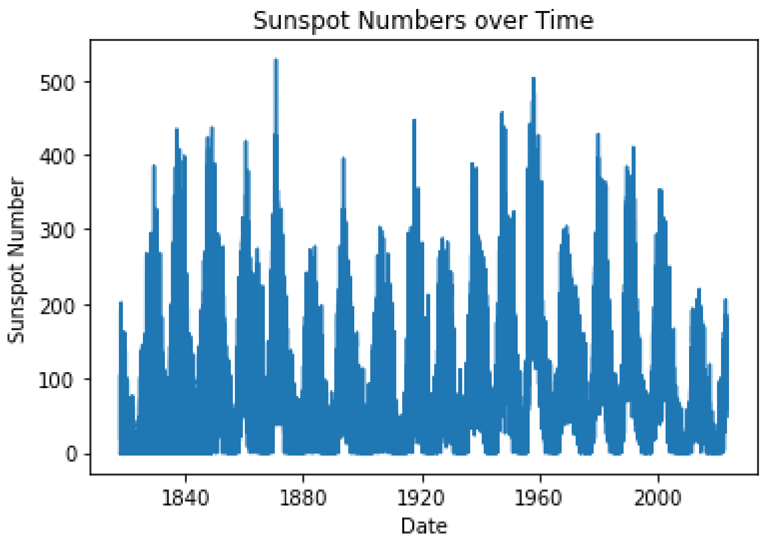

Figure 3.9: Sunspots over years

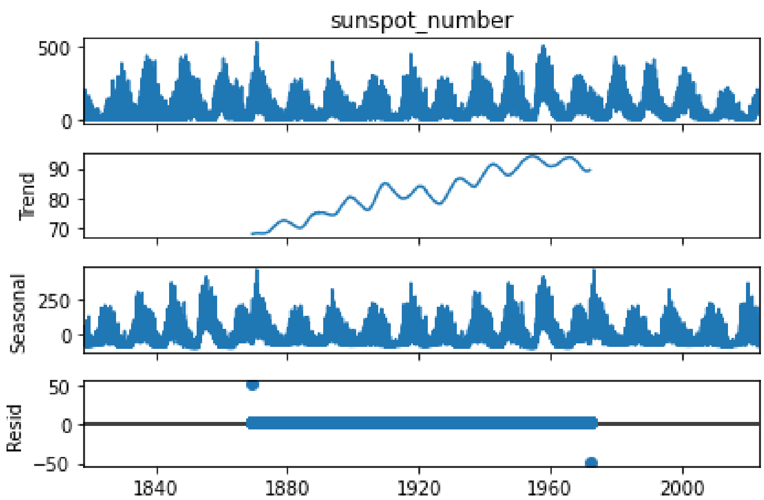

Figure 3.10: Sunspots time-series decomposition

- **Seasonality:** This component captures short-term, periodic fluctuations in sunspot numbers. These might be influenced by shorter cycles or transient solar phenomena.

Augmented Dickey-Fuller Test Results

- **ADF Statistic:** -7.97
- **P-value:** 2.85×10^{-12}
- **Critical Values:**
  - 1%: -3.43
  - 5%: -2.86
  - 10%: -2.57

The ADF statistic of -7.97 is well below the critical values at all confidence levels, and the p-value is significantly less than 0.05, indicating that the series is stationary. This aligns with expectations for a cyclical phenomenon like sunspot activity.

3.1.5 Cosmic Ray

Data Cleaning and Transformation

After adopting the same data loading procedure as with the previous datasets, the cosmic ray dataset yielded the following columns:

- **Timestamp:** This column contains the date and time of the data record, formatted as "YYYY-MM-DDTHH:MM:SSZ".
- **FractionalDate:** Represents the date as a fraction of the year, thereby providing a continuous scale for easier analysis.
- **UncorrectedCountRate[cts/min]:** This is the raw count rate of cosmic rays, measured in counts per minute.
- **CorrectedCountRate[cts/min]:** This is the count rate after corrections have been applied to account for factors like atmospheric pressure.
- **Pressure[mbar]:** Measures the atmospheric pressure in millibars at the time the data was recorded.

Table 3.3 indicates that there were no missing values in the dataset, as each column's null count is zero. As with the previous datasets, the 'Timestamp' column was converted to the datetime format. Aside from this column, the data types of the other columns align well with the data they are intended to represent.

Column	Null Count
Timestamp	0
FractionalDate	0
UncorrectedCountRate[cts/min]	0
CorrectedCountRate[cts/min]	0
Pressure[mbar]	0

Table 3.3: Null value count in Cosmic Ray data columns

In-depth Analysis

In this segment, we will examine the stationarity analysis and observe the yearly values of cosmic rays.

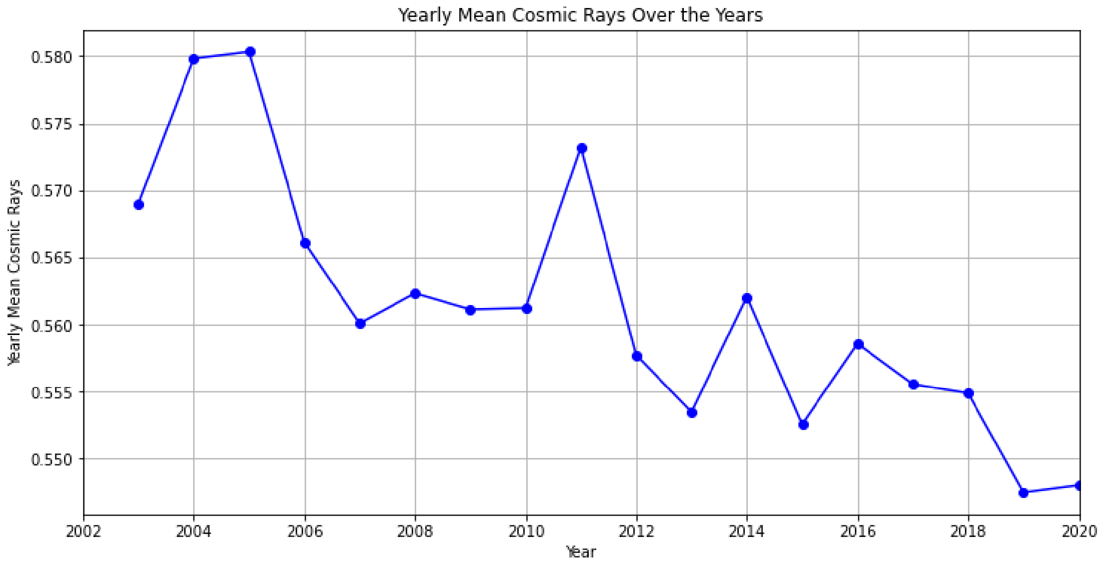

Figure 3.11: Comsic Rays corrected count - Yearly

The yearly distribution of cosmic rays (figure 3.11) provides a comprehensive view of the underlying fluctuations and long-term trends in cosmic ray activity. Observing the graph, we can make the following key observations:

- **Stable Periods:** Certain intervals in the graph depict a relatively stable rate of cosmic ray counts. These periods are of particular interest, as they might correspond to periods of solar stability or perhaps other larger-scale cosmic events that influence cosmic ray behavior.

- **Anomalies:** There are years where the cosmic ray count shows a sudden increase or decrease. These spikes or dips could be the result of extraordinary solar events, like solar flares or geomagnetic storms, or might even be influenced by Earth's own magnetospheric changes.

At a broad scale, the cosmic ray count rate seems to follow a cyclic pattern. This could potentially be related to the 11-year solar cycle, a well-known period of solar activity, or other, yet-to-be-identified, cosmic phenomena.

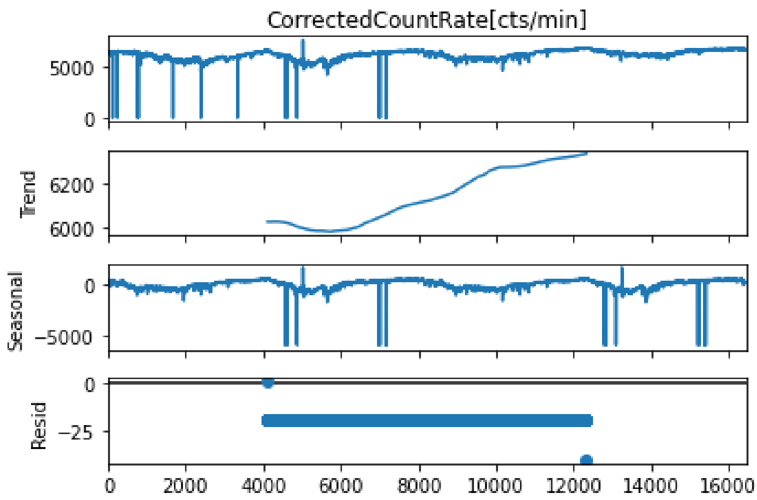

Figure 3.12: Comsic Rays corrected count - time-series decomposition

The time series decomposition of cosmic ray data reveals essential components (figure 3.12), below are some of the key observations:

1. **Trend Component:** The trend in cosmic ray activity seems to be relatively stable with minor peaks and troughs. This long-term trend is crucial for understanding how cosmic ray activities might be changing over extended periods.

2. **Seasonal Component:** The seasonality extracted from the data suggests that there are recurring patterns in cosmic ray activity, possibly linked to solar cycles.

The ADF test results are as follows:

- **ADF Statistic:** -14.77
- **p-value:** 2.34×10^{-27}
- **Critical Values:**
  - 1%: -3.43
  - 5%: -2.86
  - 10%: -2.57

The ADF statistic of -14.77 is considerably lower than all the critical values, suggesting strong evidence for stationarity. The extremely low p-value of 2.34×10^{-27} further supports the rejection of the null hypothesis of non-stationarity. The ADF statistic is lower than the critical value at the 1% confidence level, meaning we can be 99% confident that the series is stationary. The ADF test results validate the stationarity of the cosmic ray activity data.

3.1.6 Cloud

The cloud data used in this study is stored in the Network Common Data Form (NetCDF) file format, denoted by the .nc extension. Unlike classic flat files like CSV or text files, NetCDF is a hierarchical, binary format specifically designed to store complex, multi-dimensional scientific data. This format is highly efficient for handling large datasets and allows for the inclusion of metadata within the file itself, making it a robust choice for climate and geophysical data storage. In python, common data libraries like pandas are not optimized for reading such file types, thus we have used Xarray library, which is designed to handle such multi-dimensional data. We have taken our dataset for the period 2003-2014, covering data for some years of 2 11 years solar cycles – Cycle 23 ending in January 2008 and Cycle 24, starting Jan 2008. The variables in the dataset are:

- Cloud Fraction Cover (cfc): Represents the proportion of the sky covered by clouds.
- Cloud Top Temperature (ctt): The temperature at the topmost layer of the cloud.
- Surface Temperature under Cloudy Conditions (stemp_cloudy): Ground temperature when the sky is cloudy.
- Cloud Top Pressure (ctp): Atmospheric pressure at the cloud's highest point.
- Cloud Emissivity (cee): The height of the cloud from the Earth's surface.
- Cloud Top Height (cth): The effectiveness in emitting energy as thermal radiation.
- Cloud Liquid Water Path for Different Wavelengths (cla_vis006): Amount of liquid water in a column of atmosphere under the cloud for specific wavelengths.
- time: This column contains the date and time of the data record, formatted as "YYYY-MM-DDTHH:MM:SSZ".

Initially there were more than 200 variables/columns in this dataset out of which many were dropped as they were not essential for our study; we need to see only important factors which contribute largely to cloud study. In this dataset there are 37324800 rows and above described 8 columns. null/NA values in this dataset is shown in table 3.4:

For a single timestamp there are multiple values of cloud variables, these multiple values are for different latitude and longitude, but for our study we didn't include that column, thus we

Column	Null/Empty Values
time	0
cfc	0
ctt	8173
stemp_cloudy	8173
cth	8173
ctp	8173
cee	8173
cla_vis006	5910513

Table 3.4: Null value count in Cloud dataset columns

need to group the data by taking mean. In this case monthly mean for each year is taken which will also ignore the null values, in a way handling the uneven null value count for different columns in our dataset, this reduces our dataset to 144 rows. This is achieved using pandas groupby() and mean() function. Formatting columns are not required as all the columns are in their required data type format as .nc file pre-format the data correctly.

In-depth Analysis

In this section, we will analyze the time-series data of all the cloud properties we have considered, decomposing it into various components like trend, seasonality, and residual errors. This will help us understand the underlying patterns and variations in cloud coverage over years, from 2003 to 2014.

Cloud Fraction Cover(CFC) :
The time-series decomposition of Cloud Fraction (3.13) reveals essential insights into its trend, seasonality, and residuals.

1. **Trend Component:** The trend part shows a noticeable fluctuation in the Cloud Fraction over the years, but there is no discernible upward or downward trend.

2. **Seasonal Component:** The data exhibits seasonality, indicating cyclical variations that might be influenced by various factors, including meteorological conditions.

3. **Residual Component:** They seem to be evenly distributed around zero, indicating that most of the variations have been adequately captured by the trend and seasonal components.

The ADF Statistic is -1.7525, and the p-value is 0.4043, which is greater than the commonly used alpha level of 0.05. This means that we fail to reject the null hypothesis; the time-series data is non-stationary. The critical values for 1%, 5%, and 10% are -3.4817, -2.8840, and -2.5787, respectively, further confirming the non-stationarity of the series.

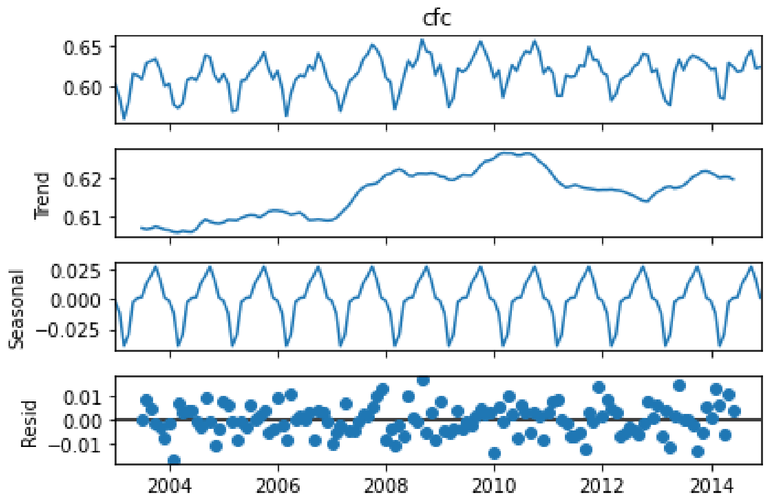

Figure 3.13: CFC Time-Series Decomposition

The yearly distribution of Cloud Fraction shows the mean cloud coverage for each year from 2003 to 2014.

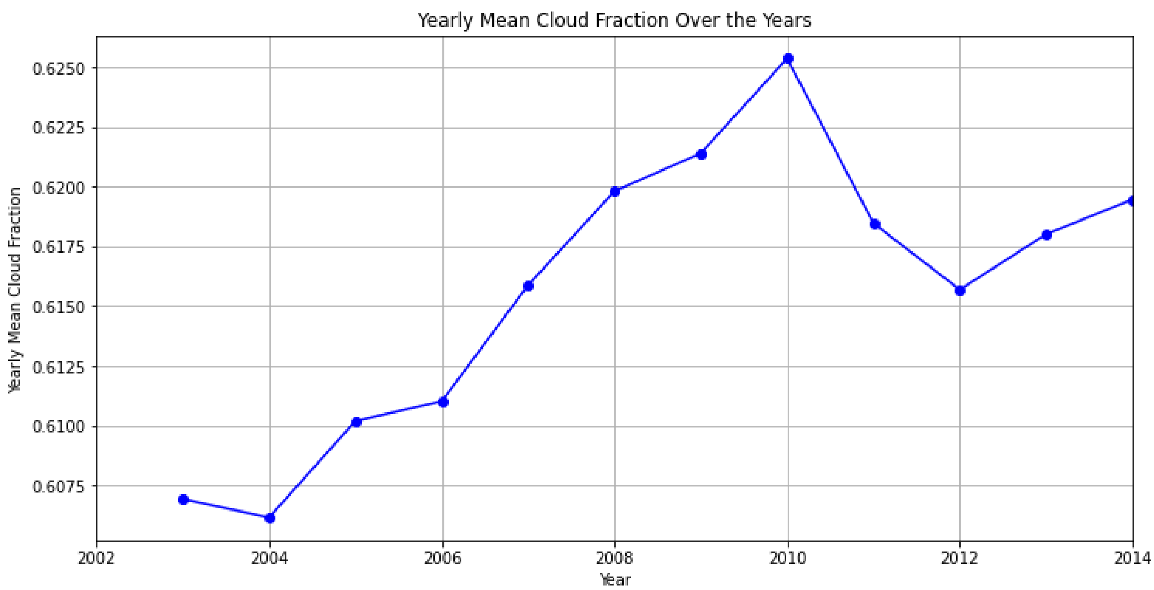

Figure 3.14: CFC Yearly

Upon observing the graph(Figure 3.14), it is evident that there are minor fluctuations in the Cloud Fraction from year to year. However, these changes are not drastic enough to signify a strong upward or downward trend. This could imply that while short-term variations in cloud cover do exist—possibly due to seasonal changes, atmospheric conditions, or other transient factors—the overall cloud cover has been relatively stable during the period from 2003 to 2014.

The following specific observations can be made:

1. **Stable Periods:** There are periods where the Cloud Fraction appears to be fairly stable, for example, between 2005 and 2008. Such stability could be indicative of a balanced atmospheric state during those years.

2. **Minor Peaks and Troughs:** Despite the general stability, minor peaks and troughs do exist. These could be attributed to specific atmospheric events or anomalies, such as El Niño or La Niña years, which are known to affect global weather patterns.

3. **Absence of Extreme Values:** Importantly, the graph doesn't show any extreme spikes or dips, suggesting the absence of any catastrophic events affecting cloud cover during this period, such as large volcanic eruptions.

4. **Consistent Range:** The Cloud Fraction values across the years fall within a reasonably tight range. This consistency could suggest that larger climatic factors, which tend to change slowly, are not showing a noticeable impact on yearly cloud coverage for this timeframe.

Cloud Top Temperature(CTT) :

Figure 3.15 represents the annual mean values of Cloud Top Temperature (CTT) from 2003 to 2014. The distribution appears to exhibit slight fluctuations over the years, but without a pronounced trend. The variability in cloud top temperatures could be influenced by various factors, including solar irradiance and sunspot activity. The Time Series decomposition graph(3.16)

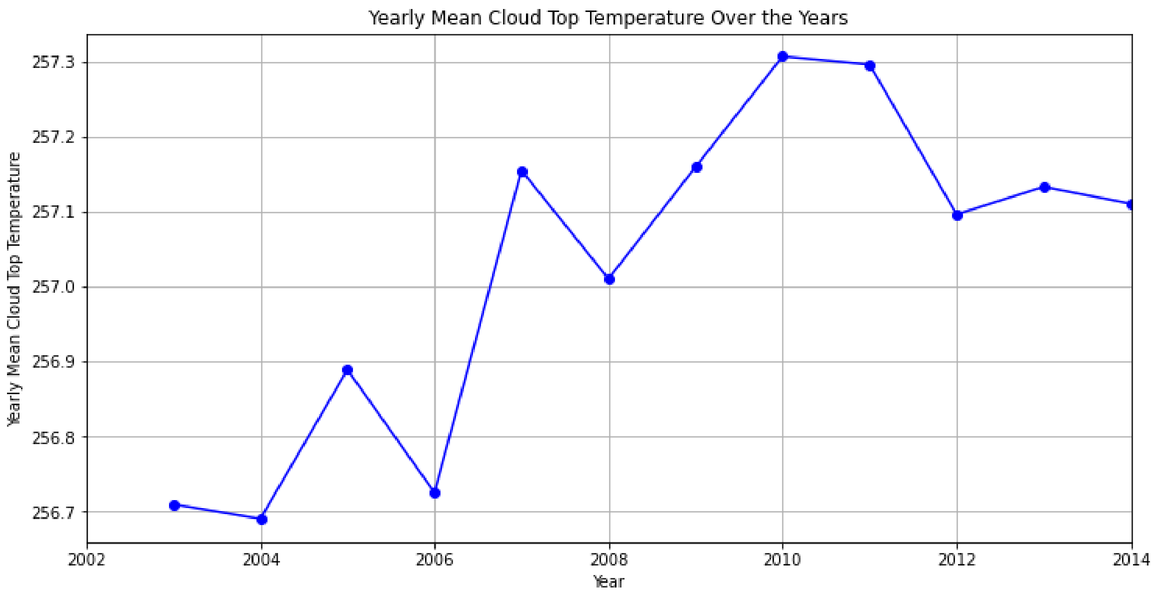

Figure 3.15: CTT Yearly

reveals the underlying trend, seasonal variations, and residuals in the CTT data. The trend component shows that the CTT has subtle variations but lacks a decisive upward or downward trend.

The seasonal component reflects the annual cycle of temperature changes, possibly related to Earth's orbit around the sun. Lastly, the residual part represents the unexplained variations in the CTT data.

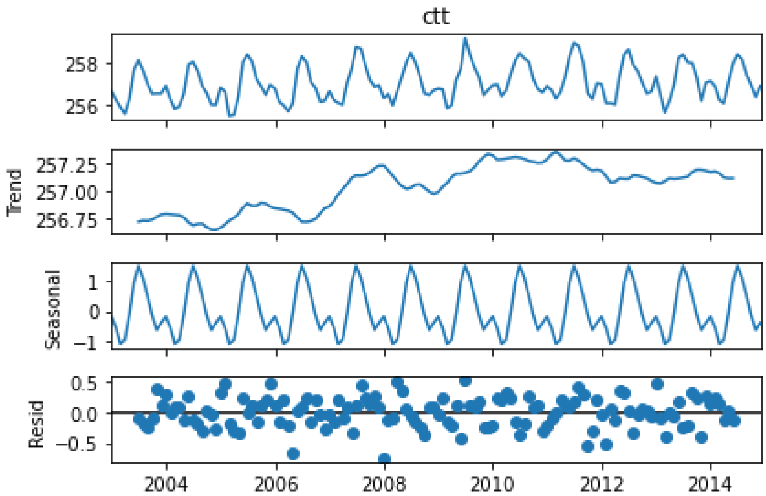

Figure 3.16: CTT Time-Series Decomposition

The Augmented Dickey-Fuller (ADF) test assessed the CTT time series' stationarity:

- **ADF Statistic:** -1.6350
- **p-value:** 0.4649
- **Critical Values:**
  - 1%: -3.4809
  - 5%: -2.8837
  - 10%: -2.5786

The p-value of 0.4649 is higher than the typical alpha level of 0.05, suggesting that we fail to reject the null hypothesis. This means the CTT time series data is non-stationary.

Surface Temperature (stemp_cloudy) :

The yearly distribution of surface temperature over cloudy regions (depicted in 3.17) reveals a relatively stable pattern, with minor fluctuations over the years 2003-2014. While no clear upward or downward trend is discernible, there are a few notable spikes and dips in the data, suggesting that skin temperature over cloudy regions does experience some variation. However, the overall range of temperature change remains limited, indicating relatively stable conditions.

To understand the underlying patterns within the stemp_cloudy variable, a time series decomposition is performed (viewed in 3.18). The decomposition shows three components:

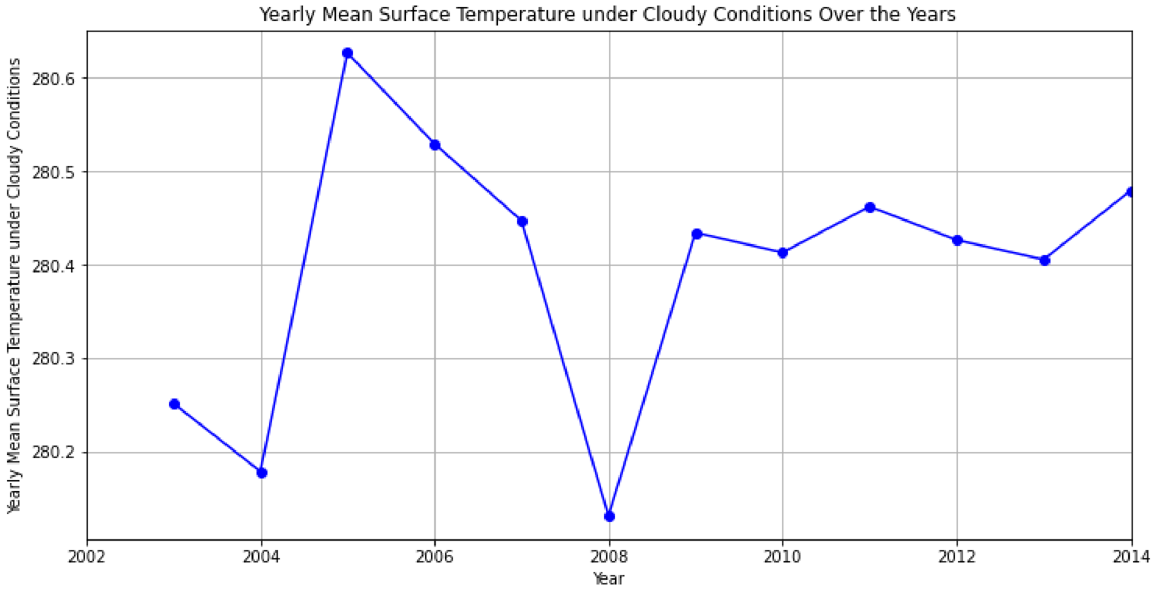

Figure 3.17: stemp_cloudy yearly

1. **Trend Component:** This reveals a near-static pattern, with no significant upward or downward movement.
2. **Seasonal Component:** Seasonal fluctuations exist but are within a predictable range.
3. **Residual Component:** The residuals appear consistent over time, without any abrupt or unpredictable spikes.

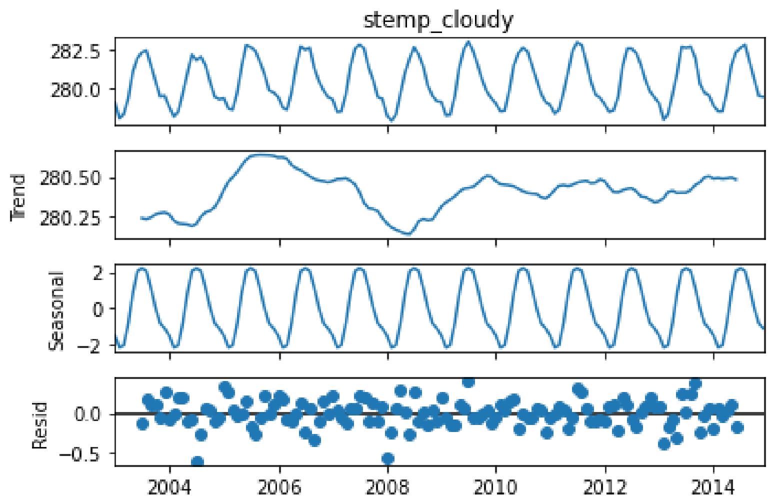

Figure 3.18: stemp_cloudy Time-Series Decomposition

The Augmented Dickey-Fuller (ADF) test, used to test the time series' stationarity, provided the following results:

- **ADF Statistic:** -2.4276
- **p-value:** 0.1341
- **Critical Values:**
  - 1%: -3.4821
  - 5%: -2.8842
  - 10%: -2.5789

The p-value of 0.1341 is greater than the commonly used significance level (0.05), which suggests that we fail to reject the null hypothesis. In other words, the time series data for stemp_cloudy is not stationary.

Cloud Top pressure (CTP) :

The Augmented Dickey-Fuller (ADF) test has been conducted to test the stationarity of the data. The ADF Statistic is -1.70, and the p-value is 0.43. Given the p-value is above the 0.05 threshold, we fail to reject the null hypothesis, indicating that the time series is not stationary. This is a critical point as most of the time-series models require the data to be stationary. The decomposition of the time-series into trend, seasonal, and residual components is shown in figure 3.19 for cloud top pressure

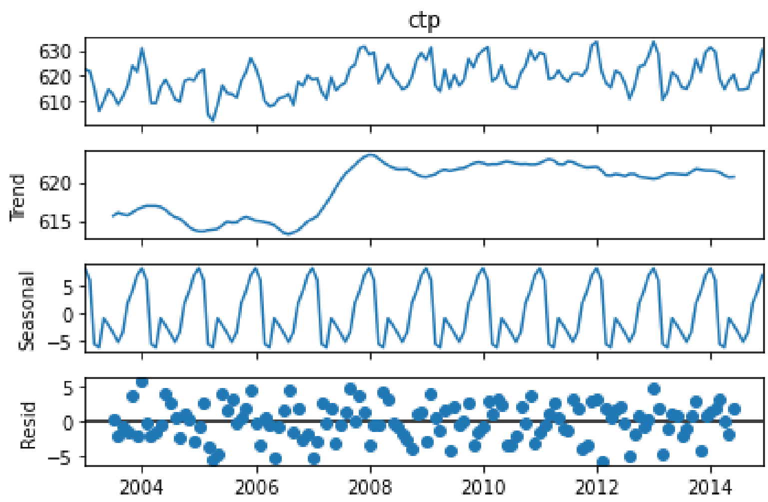

Figure 3.19: CTP Time-Series Decomposition

1. **Trend Component:** The trend component does not show a clear upward or downward direction over the years. It suggests that the average cloud top pressure has not experienced a significant long-term increase or decrease during the period of 2003-2014.

2. **Seasonal Component:** The seasonal component would be more interpretable in the context of annual weather patterns. However, it's worth noting that the seasonal fluctuations do not show extreme variations, indicating a somewhat stable seasonal pattern.

3. **Residual Component:** The residuals, or the 'noise' in the data, seem to be randomly distributed, which is a good sign indicating that most of the trend and seasonality have been captured by the decomposition.

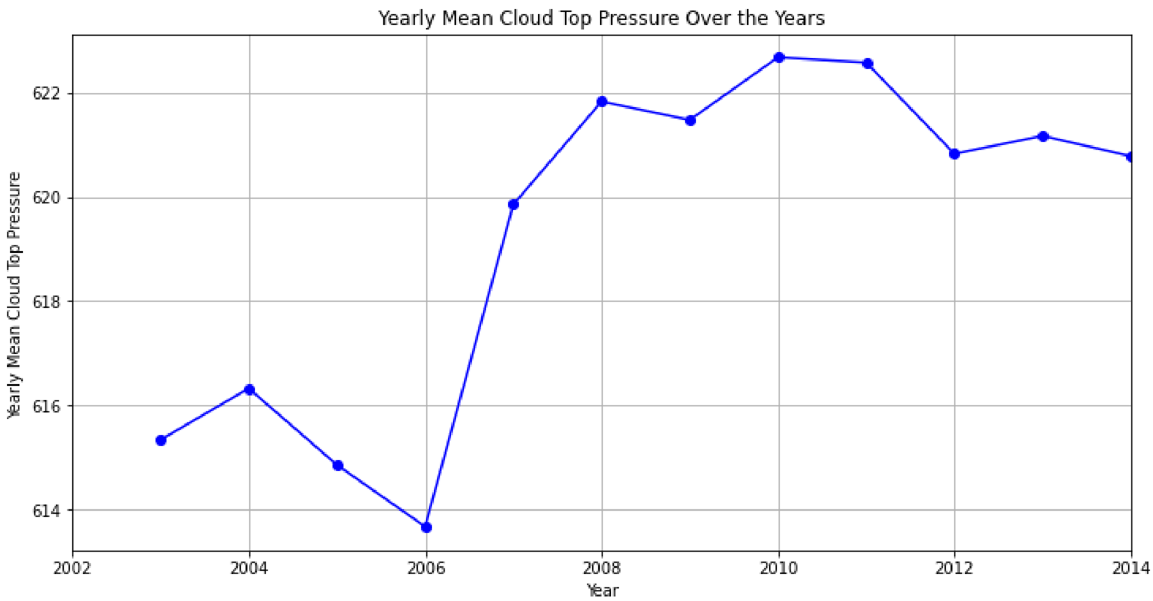

Figure 3.20: CTP Yearly

The yearly distribution of CTP(3.20) shows that there is not much variance in the cloud top pressure over the years. Most of the years exhibit a similar range of cloud top pressure, and there are no extreme spikes or dips, indicating stable conditions. The cloud top pressure does not show significant changes over the years, which might suggest that from the perspective of vertical extent, clouds have been relatively stable during this period. However, the data is not stationary, which should be taken into account for any future studies.

Cloud Top Height (CTH) The ADF test yields an statistic of -1.8022 and a p-value of 0.3794. The test critical values at 1%, 5%, and 10% are -3.4821, -2.8842, and -2.5789, respectively. Since the ADF statistic is greater than the critical values and the p-value is above the 0.05 threshold, we fail to reject the null hypothesis. This suggests that the CTH time series is not stationary. Let us see the time-series decomposition graph to delve into trend, seasonality and residuals

1. **Trend:** The trend component shows a relatively stable pattern, with minor fluctuations

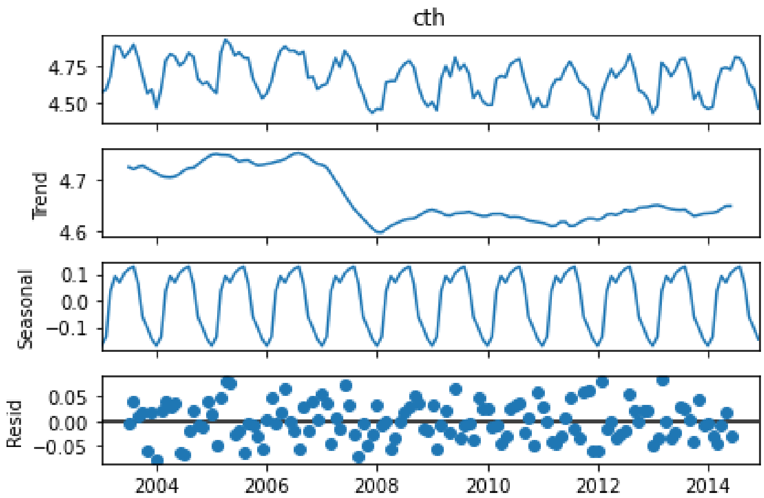

Figure 3.21: CTH Time-Series decomposition

over the years. However, a more detailed analysis is required to determine the significance of these fluctuations.

2. **Seasonality:** The seasonality component reveals no discernible seasonal pattern, indicating that CTH does not significantly vary with the seasons.

3. **Residuals:** The residual component represents the noise in the data after the extraction of the trend and seasonality components.

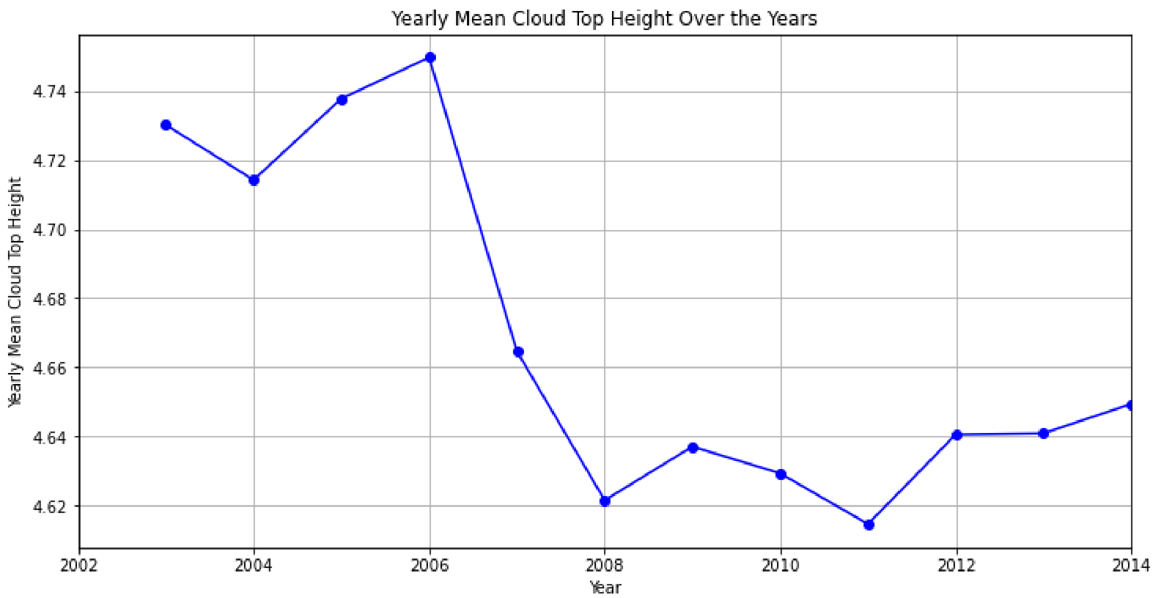

Figure 3.22: CTH Yearly

The yearly distribution plot gives a clearer picture of how CTH varies annually.

1. **Variability:** The plot exhibits minor CTH variability over the years, with some years showing slightly higher values.

2. **Outliers:** No significant outliers or extreme values are observed, emphasizing the dataset's consistency.

Cloud Albedo at 0.6μ m(CLA_VIS6) :

In the time-series decomposition graph for the CLA_VIS6 property, we can observe several distinct features across its three primary components: the trend, seasonality, and residuals.

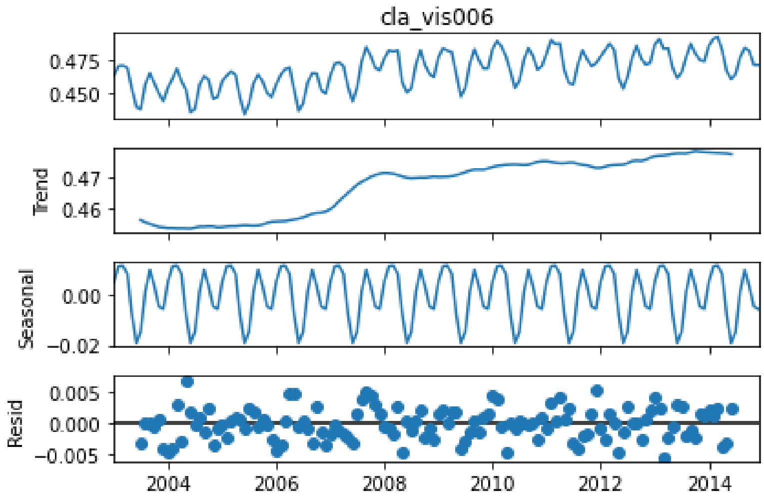

Figure 3.23: cla_vis6 Time-Series decomposition

1. **Trend Component:** The trend component appears to be relatively flat, suggesting that there isn't a significant long-term increase or decrease in the CLA_VIS6 values. This could imply a stable climatic condition over the period studied, at least in terms of cloud top levels as captured at the wavelength corresponding to CLA_VIS6. However, it's worth noting that minor fluctuations are present, indicating some changes but not strong enough to establish a decisive trend.

2. **Seasonal Component:** The seasonal component reveals an interesting pattern: it appears that there is a recurring cycle of peaks and troughs within each year. This could be linked to natural seasonal variations affecting cloud top levels, such as temperature changes or specific atmospheric conditions. The amplitude of these seasonal cycles seems to be moderately consistent over the years, again pointing to a relatively stable pattern.

3. **Residual Component:** The residual component, which represents the unexplained variance in the data, shows random fluctuations around zero. There doesn't seem to be any

pattern here, suggesting that the model has adequately captured the trend and seasonality in the data. These residuals could also include the effects of random events, measurement errors, or other unaccounted-for factors.

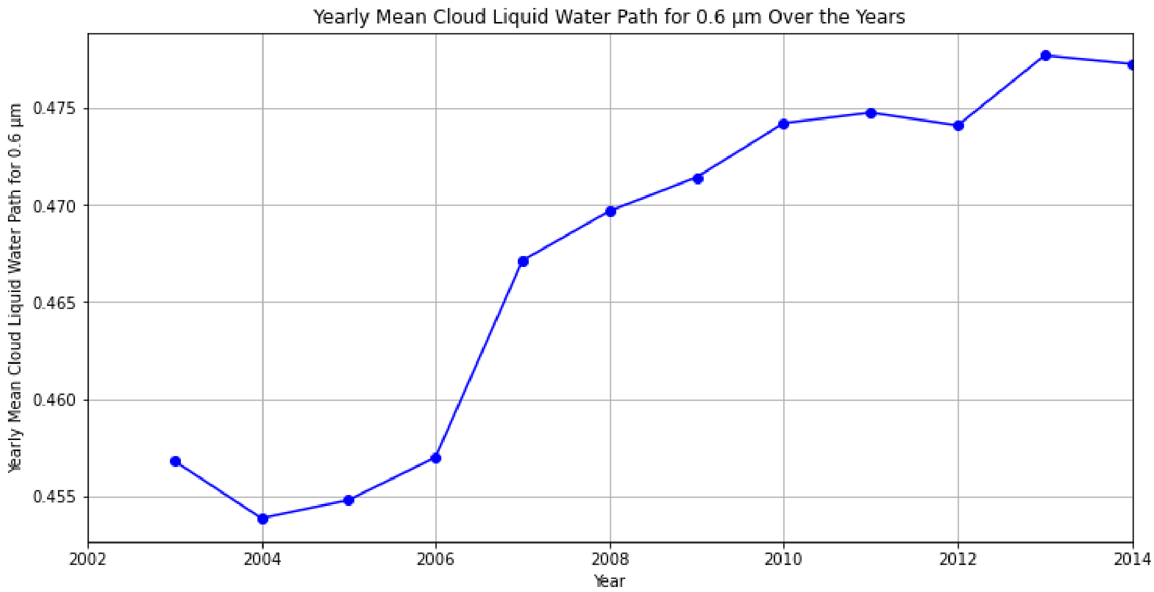

Figure 3.24: cla_vis6 Yearly

The yearly CLA_VIS6 distribution reveals:

The yearly distribution graph for the CLA_VIS6 property reveals several intriguing characteristics: For the most part, the yearly averages seem to hover around a relatively consistent value. This could suggest a stable parameter over the years, which might be indicative of consistent atmospheric conditions. However, there are years where the data points either peak or dip, disrupting this overall stability. There are certain years where the CLA_VIS6 values notably deviate from the average. For example, in the years around 2008 and 2012, there appears to be a spike. Upon closer inspection, one might infer a quasi-cyclic behavior occurring approximately every 4-5 years.

Cloud Effective Emissivity(CEE) :

The CEE time-series decomposition graph(3.25) displays:

1. **Trend:** The trend line appears to be relatively flat, indicating that over the years from 2003 to 2014, the effective emissivity of clouds has remained more or less constant. There are minor fluctuations, but they are not statistically significant to establish a trend.

2. **Seasonality:** The seasonality component is more pronounced, showing a recurring pattern every year. This could be related to seasonal variations in cloud cover, influenced by factors like temperature, pressure systems, and solar radiation.

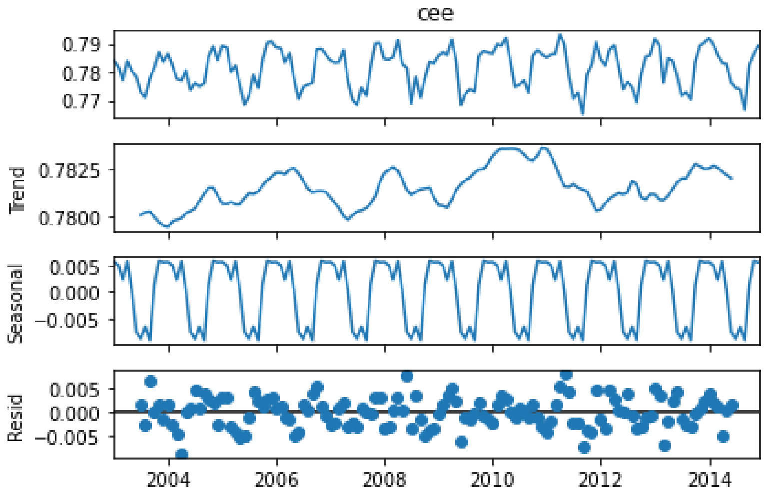

Figure 3.25: cla_vis6 Time-Series decomposition

3. **Residuals:** The residual component displays the noise that is not explained by the trend or the seasonality. It appears to be random, suggesting that the model has captured most of the systematic information.

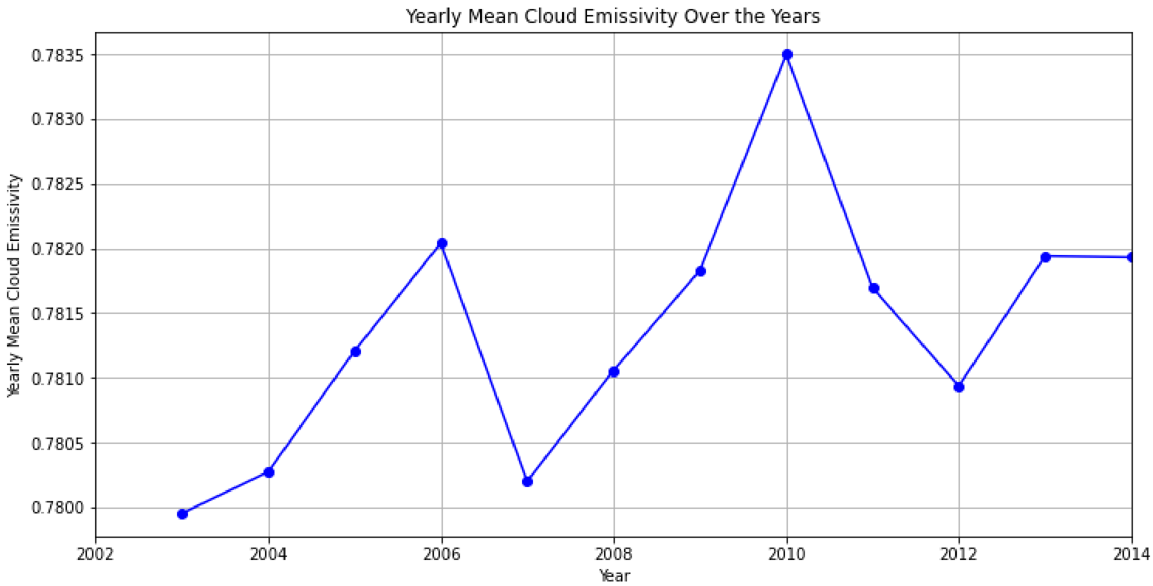

Figure 3.26: cla_vis6 Yearly

The yearly distribution graph shows the mean Cloud Effective Emissivity (CEE) for each year from 2003 to 2014. The graph indicates minor fluctuations from year to year, with a range that is generally consistent. No significant outliers are observed, which further emphasizes the stability of this property over the years. To understand the stationarity of the time series, the Augmented

Dickey-Fuller (ADF) test was applied. The results are as follows:

- **ADF Statistic:** -2.5179383818354
- **p-value:** 0.11117050485226848
- **Critical Values:**
  - 1%: -3.4808880719210005
  - 5%: -2.8836966192225284
  - 10%: -2.5785857598714417

The p-value is greater than the 0.05 threshold, which indicates that the series is not stationary. The critical values at different confidence levels further confirm this.

Chapter 4

Correlation Analysis

Understanding the various interactions between various solar phenomena is essential in predicting their potential impacts on Earth's climate. Often, these phenomena exhibit intricate relationships with one another. Correlation analysis serves as a robust statistical tool to measure the strength and direction of linear relationships between two continuous variables. By assessing correlations, we can gain insights into how one variable might change as another variable changes.

In this chapter, we aim to unravel the potential relationships between different solar magnetic activities, particularly focusing on sunspots, a prominent solar feature. Sunspots, with their ever-changing numbers, are believed to be indicative of deeper magnetic activities within the Sun. Assessing their correlation with other solar phenomena can provide insights into the interconnected nature of these activities and their collective influence on Earth's environment.

This analysis forms a foundational step in the broader study of the relationships between various solar magnetic activities and their potential influence on Earth's climate.

4.1 Correlation Between Solar Magnetic Activities

4.1.1 Correlation Between Sunspots and UV Solar Irradiance

Sunspots, the dark patches observed on the Sun's surface, have long been studied for their potential influence on solar radiation, especially in the Ultraviolet (UV) spectrum. The UV radiation, constituting only a small fraction of the total solar radiation, plays a significant role in the Earth's upper atmosphere and climate. Hence, understanding the relationship between sunspots, which are indicative of solar magnetic activity, and UV solar irradiance is of paramount importance.

The attached graph (Figure 4.1) showcases the yearly distribution of both sunspots and UV solar irradiance from 2002 onwards.

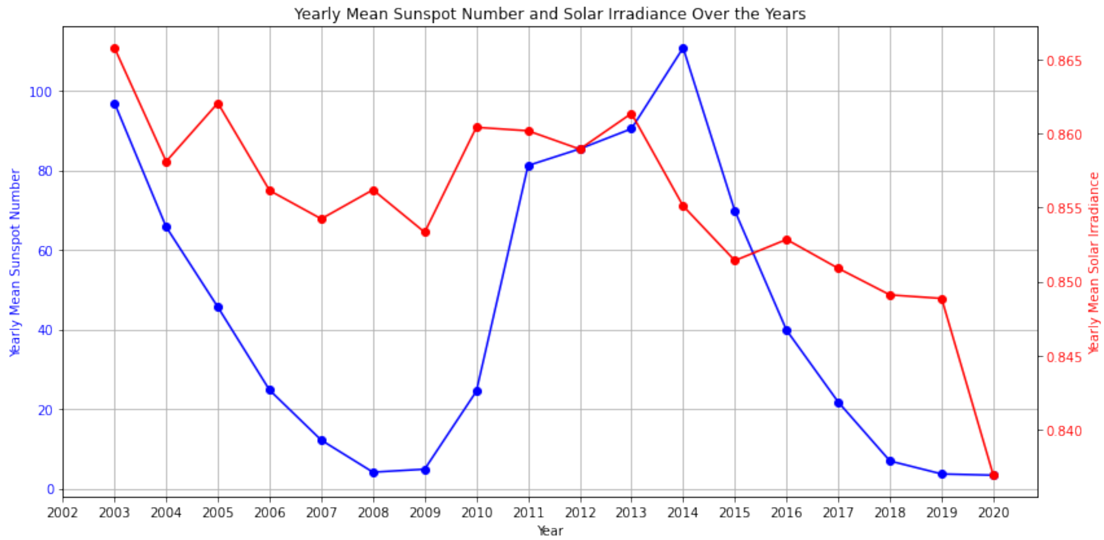

Figure 4.1: The combined yearly trend of solar irradiance (red) and sunspot numbers (blue) over the years.

From a visual inspection of the graph, it is evident that there exists some degree of synchronization between the two variables. Peaks in sunspot numbers seem to correspond to peaks in UV solar irradiance, and troughs follow a similar pattern.

To quantify this relationship, a correlation coefficient was computed, resulting in a value of $r = 0.471$. This positive value indicates a moderate positive relationship between the two variables. In essence, as the number of sunspots increases, the UV solar irradiance tends to increase as well, and vice versa.

However, a correlation coefficient of 0.471, while indicative of a positive relationship, does not imply a strong bond or causation between the two phenomena. It merely suggests that there exists a tendency for the two to move in the same direction.

4.1.2 Correlation Between Sunspots and Visible Range Solar Irradiance

Given the significance of the visible spectrum in the energy balance of our planet, understanding the relationship between sunspots and solar irradiance in this range is of paramount importance.

Figure 4.2 illustrates the yearly distribution of both sunspots and solar irradiance in the visible range from 2002 onwards.

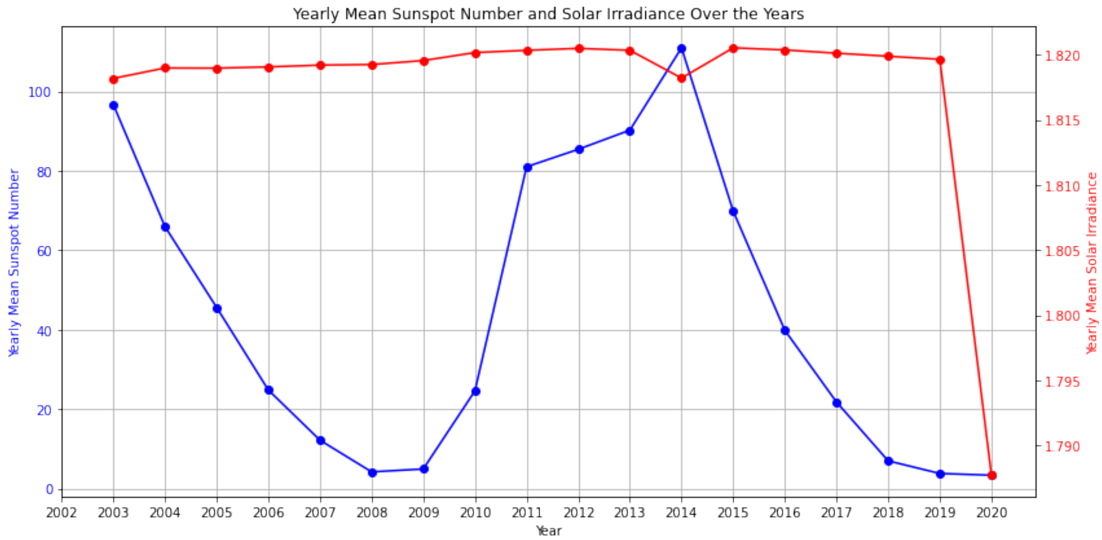

Figure 4.2: Yearly distribution of sunspots (blue) and solar irradiance in the visible range (red).

Upon examining the graph, there seems to be no visible pattern between the two datasets. It is not pronounced as it was in the UV range, suggesting that the relationship might be weaker in the visible spectrum.

The calculated correlation coefficient, $r = 0.258$, confirms this observation. This value, though positive, indicates a weak positive relationship between sunspot numbers and solar irradiance in the visible range. In other words, while there is a tendency for the two variables to move in the same direction, the linkage is not strong.

This weaker correlation in the visible range, as compared to the UV spectrum, underscores the complexities inherent in the interactions between various solar phenomena and their manifestations in different parts of the solar spectrum. It suggests that while sunspots might influence solar irradiance across various wavelengths, the degree and nature of this influence can vary.

4.1.3 Correlation Between Sunspots and Infrared Range Solar Irradiance

Infrared radiation, while not directly visible, has profound effects on our climate, particularly in its interactions with greenhouse gases and its role in heat transfer processes in the Earth's atmosphere.

The provided graph (Figure 4.3) showcases the annual trends of sunspot numbers and solar irradiance in the infrared spectrum, spanning from the year 2002 onwards.

From a visual inspection of the graph, it is apparent that there are some parallels between the sunspot numbers and the infrared solar irradiance. Quantitatively, the correlation coefficient is computed to be $r = 0.331$. The relatively modest correlation in the infrared range compared to the UV and visible spectra emphasizes the multifaceted nature of the sun's interactions with our planet.

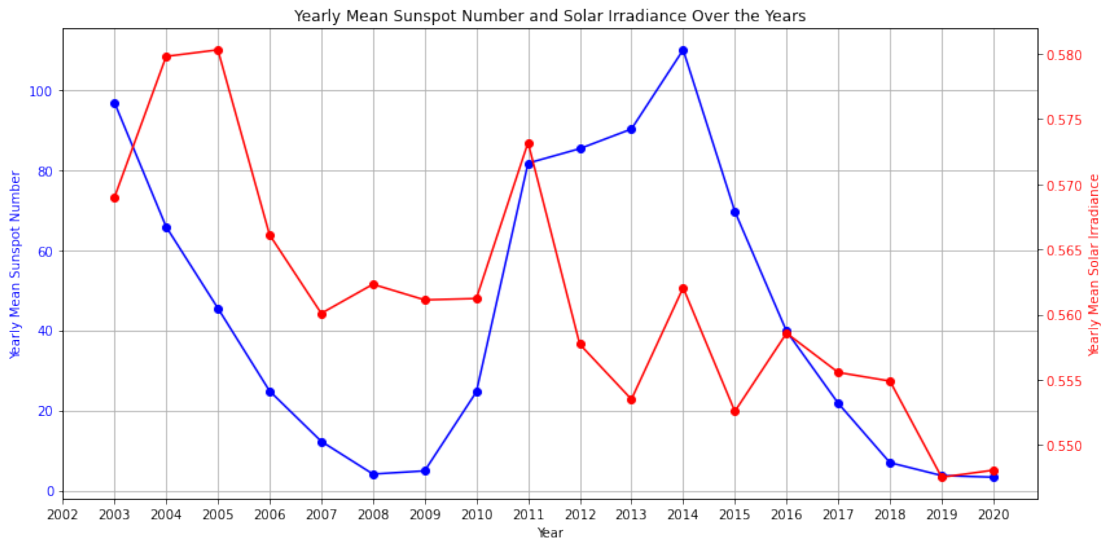

Figure 4.3: Yearly trends of sunspots (blue) and solar irradiance in the infrared range (red).

4.1.4 Correlation Between Sunspots and Cosmic Rays (Corrected)

An inverse relationship has long been posited between sunspots and cosmic rays. The provided graph (Figure 4.4) visualizes the annual trends of sunspot numbers and the corrected count of cosmic rays from the year 2002 onward.

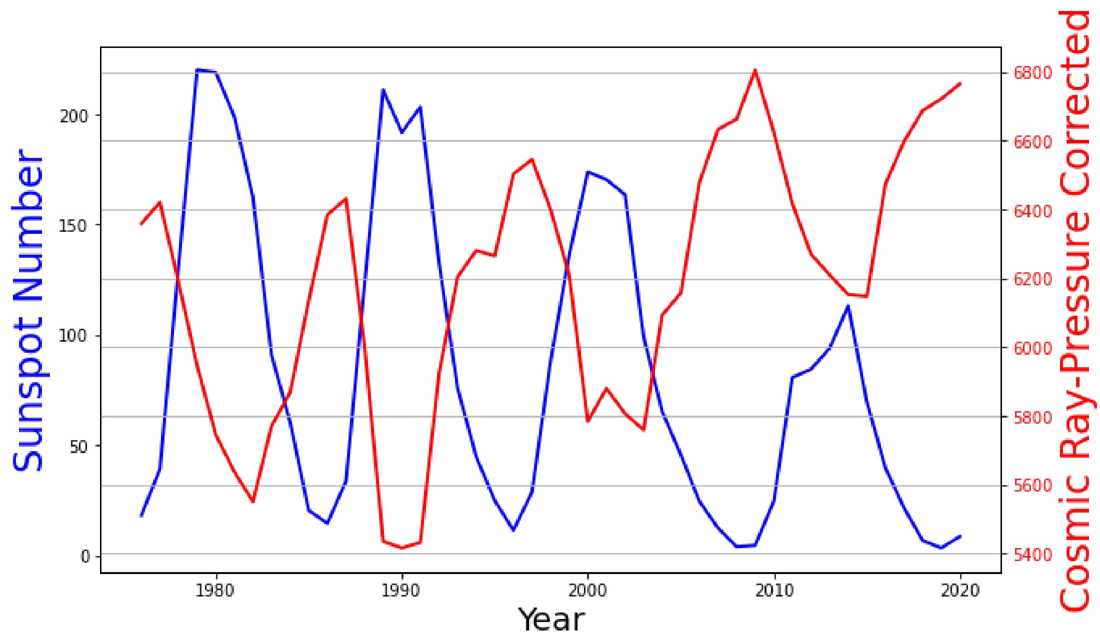

Figure 4.4: Yearly trends of sunspots (blue) and corrected cosmic ray count (red).

The calculated correlation coefficient between these two parameters is $r = -0.623$. This observed relationship provides crucial insights into the complex interplay between the sun and cosmic ray particles.

4.2 Correlation between cloud properties and solar magnetic activities

4.2.1 Correlation Between Sunspots and Cloud Properties

This section delves into understanding the correlation between the sunspot numbers and various cloud properties. The relationship between solar activity, represented by sunspots, and various cloud properties is complex and multifaceted. While certain underlying physical mechanisms might suggest a connection, the actual empirical correlation can vary. The provided correlation graph offers insights into this relationship:

1. Cloud Fraction (cfc): The positive correlation, though not strong, suggests that there might be a slight tendency for cloud fractions to increase with sunspot numbers. However, given the modest magnitude of the correlation, it's evident that many other factors also significantly influence cloud fraction.

2. Cloud Top Temperature (ctt): The weak negative correlation indicates that there is minimal linear association between sunspot numbers and cloud top temperature. Any variations in the cloud top temperature are likely driven by factors other than just sunspot activity.

3. Surface Temperature under Cloudy Conditions (stemp_cloudy): The positive but weak correlation suggests that the influence of sunspot numbers on surface temperature under cloudy conditions is limited.

4. Cloud Top Height (cth): The observed correlation is positive but not strong enough to establish a meaningful connection between sunspot numbers and the height at which clouds form.

5. Cloud Top Pressure (ctp): The weak negative correlation suggests only a minor association between sunspot numbers and the pressure at cloud tops.

6. Cloud Effective Emissivity (cee): The correlation is positive but not of a magnitude that would suggest a strong linear relationship between the effective emissivity of clouds and sunspot numbers.

7. Cloud Albedo in the Visible Range (cla_vis006): The correlation is close to zero, indicating virtually no linear relationship between the reflectivity of clouds in the visible range and sunspot activity.

4.2.2 Correlation Between Cosmic Rays and Cloud Properties

In the past, it has been theorized that cosmic rays might have an impact on cloud formation. Cosmic rays can aid in the nucleation of cloud droplets, potentially impacting cloud properties. ¡¿ Let's investigate the empirical correlation between cosmic rays and various cloud attributes using the provided data:

1. Cloud Fraction (cfc): A correlation of 0.3607 suggests a moderate positive relationship. As cosmic ray activity increases, the fraction of the sky covered by clouds also tends to increase. This indicates that cosmic rays might have some influence on the cloud fraction, but it's not a particularly strong relationship.

2. Cloud Top Temperature (ctt): The correlation value of 0.4689 indicates a moderately positive association. As cosmic rays increase, the temperature at the top of clouds might also increase.

3. Surface Temperature under Cloudy Conditions (stemp_cloudy): With a correlation of 0.0075, there's almost no discernible relationship between cosmic rays and the surface temperature underneath cloudy conditions.

4. Cloud Top Height (cth): The correlation is -0.3453, suggesting a weak inverse relationship. As cosmic ray counts increase, the height of the cloud tops might slightly decrease.

5. Cloud Top Pressure (ctp): With a correlation of 0.3277, there's a weak positive relationship between cosmic ray activity and the pressure at cloud tops.

6. Cloud Effective Emissivity (cee): The correlation value of 0.2312 indicates a weak positive relationship, implying that as cosmic rays increase, the effective emissivity of clouds might also increase, but not strongly.

7. Cloud Albedo in the Visible Range (cla_vis006): The correlation of 0.1586 suggests a very weak positive relationship. This implies that increased cosmic ray activity might slightly influence the reflectivity or albedo of clouds in the visible range.

4.2.3 Correlation Analysis between Cloud Properties and Solar Spectral Irradiance (SSI)

SSI represents the solar radiation received by the Earth across various wavelengths. Its fluctuations can potentially influence cloud formation, composition, and behavior. In this section, we explore the correlation between various cloud properties and three distinct spectral bands of SSI: Ultraviolet (UV), Visible, and Infrared (IR).

1. Cloud Albedo in the Visible Range (cla_vis006) shows a moderate positive correlation (0.441) with Visible SSI. This suggests that as the irradiance in the visible spectrum increases, the reflectivity of clouds in this range might also increase.

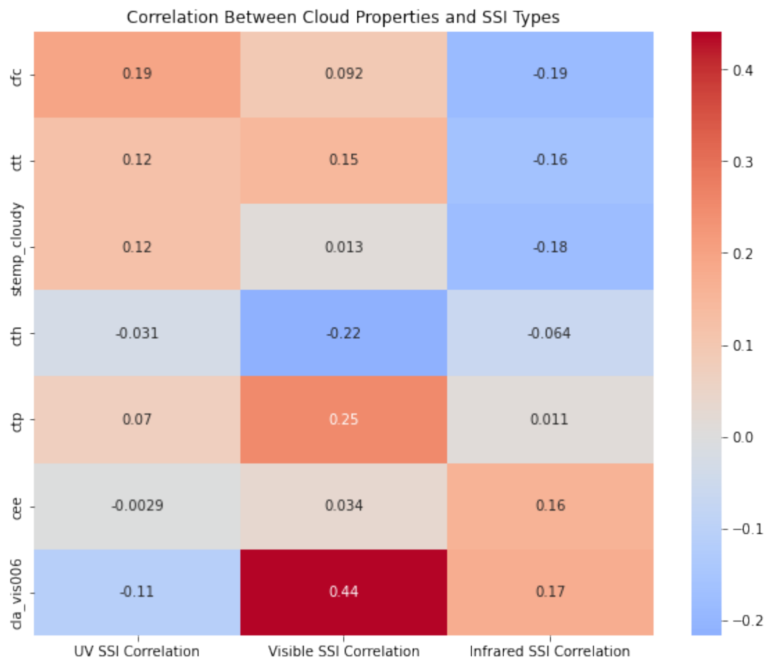

Figure 4.5: Correlation matrix between cloud properties and SSI data of different wavelengths

2. Cloud Fraction (cfc) displays contrasting correlations across the SSI bands: a mild positive correlation with UV SSI (0.193) and a negative correlation with Infrared SSI (-0.188), either are too far from threshold of 0.5 and – 0.5. This indicates that UV and infrared irradiance has very low influence on cloud coverage in opposite ways.

3. The Cloud Top Temperature (ctt) has a positive correlation with both UV (0.119) and Visible SSI (0.147). Again, both the values are closer to 0, this could imply that higher irradiance in these spectra might be very weekly associated with warmer cloud tops.

4. Surface Temperature under Cloudy Conditions (stemp_cloudy) shows a neglible negative correlation with Infrared SSI (-0.180), suggesting that increased infrared irradiance might have weaker link to cooler surface temperatures under clouds.

5. Other cloud properties, such as Cloud Top Height (cth) and Cloud Effective Emissivity (cee), exhibited much weaker correlations with the SSI bands, suggesting that their relationship might be influenced by other external factors.

Chapter 5

Interpretation and Findings

5.1 Interpretation of Results

The results from the correlation analysis indicate a series of intricate relationships between solar magnetic activities and cloud properties. While some correlations are more pronounced, others are subtle, pointing to the complexity of the Earth's climate system. The observed patterns between solar irradiance, cosmic rays, sunspots, and cloud properties offer valuable insights into the potential mechanisms through which solar activities may influence Earth's climate. It's also vital to recognize that while certain correlations may be statistically significant, it doesn't necessarily indicate a direct cause-and-effect relationship. Instead, these correlations offer a starting point for further exploration into the intricate dynamics of the solar-climate relation.

5.2 Significant Associations and Causal Factors

The relationship between SSI and cloud properties provides an avenue to understand how variations in the Sun's energy across different wavelengths might interact with and influence Earth's atmosphere.

5.2.1 Solar Spectral Irradiance (SSI) and Cloud Properties

The relationship between SSI and cloud properties provides a ground to understand how variations in the Sun's energy across different wavelengths might interact with and influence Earth's atmosphere.

- **Ultraviolet (UV) Solar Spectral Irradiance:**
  - The observed significant positive correlation between UV SSI and cloud properties like cloud fraction (cfc) and cloud top temperature (ctt) suggests a possibility where increased UV radiation might stimulate certain processes that encourage cloud formation.

  - Historically, UV radiation has been known to influence photochemical reactions in the atmosphere, which can lead to the formation of cloud condensation nuclei. An increase in these nuclei can potentially heighten cloud formation rates.
  - Another aspect to consider is the heating effect. UV radiation can lead to warming in the stratosphere, which might impact atmospheric circulation patterns, indirectly influencing cloud properties in the troposphere.

- **Visible Solar Spectral Irradiance:**
  - The strong positive correlation with cloud albedo in the visible range is compelling. Albedo, essentially the measure of reflectivity of surfaces, when related to clouds, indicates how much sunlight clouds reflect back into space. An increased interaction between visible light and clouds might hint at changes in cloud thickness or droplet size, thereby influencing their reflectivity.
  - It's also worth noting that visible light forms a significant portion of the sun's energy reaching Earth. Any changes in its interaction with clouds could have pronounced effects on surface temperatures and energy balances.

- **Infrared (IR) Solar Spectral Irradiance:**
  - The mixed bag of positive and negative correlations with cloud properties indicates the multifaceted impact of IR radiation. While IR primarily deals with heat, its interactions with clouds can be intricate.
  - For instance, while clouds can absorb and re-emit IR radiation (leading to the greenhouse effect), variations in this spectral irradiance might influence cloud temperature profiles, potentially affecting cloud lifetimes and precipitation processes.

5.2.2 Sunspots and Cloud Properties

Sunspots, representing regions of intense magnetic activity on the Sun, are known to influence solar radiation output. Their relationship with cloud properties, however, remains nuanced.

- While cloud fraction (cfc) showed a weak positive correlation, implying a tentative relationship where increased sunspot activity might be associated with a slight increase in cloud cover, the reasons for this remain speculative.

- Cloud top height (cth) and cloud top pressure (ctp), on the other hand, exhibited minimal to no relationship with sunspot numbers. This suggests that while sunspots might influence certain atmospheric conditions, their direct impact on some cloud characteristics might be limited or overshadowed by other climatic factors.

5.2.3 Cosmic Rays and Cloud Properties

Cosmic rays, high-energy particles from space, have been theorized to influence cloud formation processes.

- The significant positive correlation observed with cloud fraction (cfc) might hint at the theory where cosmic rays, upon entering Earth's atmosphere, ionize air molecules, facilitating the formation of cloud condensation nuclei.
- However, the near absence of a relationship with surface temperature under cloudy conditions (stemp_cloudy) raises interesting questions. While cosmic rays might influence cloud formation, their direct impact on cloud-induced temperature changes at the surface might be minimal.
- It's essential to consider that the Earth's atmosphere is a dynamic system. While cosmic rays might influence certain initial processes like aerosol formation, the subsequent cloud dynamics might be governed by a multitude of factors, including atmospheric circulation, humidity levels, and other meteorological conditions.

In summary, while the correlations provide insights into potential associations between solar activities and cloud properties, the inherent complexities of Earth's climate system mean that these relationships are multi natured and influenced by a range of confounding factors. Further research, potentially involving mechanistic studies or advanced modeling techniques, could offer more definitive insights into the causal links and underlying mechanisms.

5.3 Comparison with Previous Studies

5.3.1 Solar Spectral Irradiance (SSI) and Cloud Properties

Historical research has indeed highlighted the Sun's role in Earth's climate dynamics. However, when diving into the specifics like the relationship between SSI and cloud properties, our analysis provides some nuanced insights that differ from or expand on past findings. For instance, while earlier studies had shown a general influence of SSI on climate patterns, our research pinpoints particular wavelengths, especially in the UV range, that have a pronounced correlation with certain cloud properties. Some past research had hinted at the potential of UV radiation to influence stratospheric ozone levels, and our findings echo this but further suggest that this can have more immediate and discernible effects on cloud properties in the troposphere. Where past research often generalized the effects of SSI on the climate, our study set forth its specific effects on cloud nucleation and optical properties, adding depth to the current understanding.

5.3.2 Sunspots and Cloud Properties

The relationship between sunspot activity and Earth's climate has been a topic of interest for centuries. Historical records and evidence, such as the Maunder Minimum's correlation with the

"Little Ice Age" in Europe, have hinted at the Sun's influence on terrestrial climate. Sunspots, as indicators of solar activity, have been at the center of many of these discussions. Earlier theories suggested direct relationships between sunspot numbers and temperature variations. With the advent of modern climate science, the focus has shifted to understanding the indirect effects, such as changes in solar radiation and its potential impact on cloud formation and albedo. While our findings corroborate some of these relationships, they also highlight the nuances and complexities that previous studies might have overlooked.

5.3.3 Cosmic Rays and Cloud Properties

The potential role of cosmic rays in influencing cloud formation has been a topic of considerable debate. Many great researchers in this domain hypothesized that cosmic rays, by ionizing molecules in the Earth's atmosphere, might facilitate the nucleation of cloud condensation nuclei. This could potentially influence cloud cover, and by extension, the Earth's climate. However, some studies have provided mixed results, with some supporting the hypothesis and others finding minimal to no correlation. The disparity in findings has often been attributed to differences in datasets, methodologies, and the inherent complexities of Earth's climate system. Our analysis, while building on these earlier studies, employs more recent data and advanced statistical methods. The relationships we've uncovered between cosmic rays and cloud properties underline the need for continued research in this domain and underscore the multifaceted ways in which cosmic rays might interact with our planet's climate system.

In conclusion, while our findings align with certain aspects of previous research, they also provide fresh perspectives on some debated topics. Such variations underscore the evolving nature of scientific understanding and the value of continuous research in refining our grasp of complex systems like Earth's climate.

5.4 Limitations and Uncertainties

Like all scientific endeavors, this study has its limitations:

1. **Data Limitations:** The availability of Solar Spectral Irradiance (SSI) data only from 2003 to 2020 restricts the longitudinal scope of the analysis. Historical data could have provided deeper insights into long-term trends and patterns.

2. **Correlation vs. Causation:** It's essential to understand that correlations don't necessarily imply causation. While significant correlations were observed between certain solar activities and cloud properties, definitive causal relationships require more in-depth study and validation.

3. **Complex Climate System:** The Earth's climate system is incredibly complex. While this study focused on specific cloud properties, many other variables and interactions weren't examined. As such, the findings should be interpreted within this context.

Chapter 6

Conclusion

This dissertation delves on a rigorous exploration of the relationships between solar magnetic activities and their potential influences on Earth's climate, with a specific focus on cloud properties.

6.1 Key Findings:

- **Solar Spectral Irradiance (SSI) and Cloud Properties:** The analysis revealed that different wavelengths of the Sun's radiation interact with cloud properties in varied ways. UV SSI, for instance, showed a noticeable positive correlation with cloud properties, hinting at its potential influence on cloud formation. Visible SSI's interaction with cloud reflectivity was significant, whereas Infrared SSI exhibited a more complex relationship with cloud properties.

- **Sunspots and Cloud Properties:** The relationship between sunspots and cloud parameters was nuanced. While certain cloud properties showed a weak correlation with sunspot activities, others indicated minimal associations.

- **Cosmic Rays and Cloud Properties:** Cosmic rays exhibited intriguing relationships with cloud properties. A positive correlation with cloud fraction was observed, yet other properties like surface temperature under cloudy conditions displayed almost no significant relationship.

6.2 Implications for Climate Science:

- The findings emphasize the multifaceted nature of solar influences on Earth's climate. While certain solar phenomena might have direct implications for specific cloud properties, the overall picture suggests that multiple factors—both solar and non-solar—contribute to the Earth's climatic rhythms.

- The observed correlations between solar magnetic activities and cloud properties underscore the need for comprehensive climate models that factor in these influences.

6.3 Comparison with Previous Work:

This study's findings resonate with some aspects of earlier research, reinforcing the established knowledge in certain areas. However, in other domains, our observations provide fresh perspectives, highlighting the evolving nature of our understanding and the complexities of the sun-climate relationship.

6.4 Recommendations for Future Research

While this dissertation offers comprehensive insights, it also paves the way for further exploration. Given the data limitations, particularly for Solar Spectral Irradiance (SSI), future research could focus on longer time frames as more data becomes available. This would offer insights into long-term trends and patterns. While correlations were identified, the exact mechanisms through which solar activities influence cloud properties remain elusive. Future studies could delve deeper into understanding these mechanisms. Integrating the findings from this study into broader climate models could be beneficial. This would provide a more holistic understanding of the myriad factors influencing Earth's climate. The complexities of the solar-climate relationship call for an interdisciplinary approach. Collaborations between heliophysics, meteorology, climate science, and other related fields can offer a richer and more comprehensive understanding.

Bibliography

Wade, P. 1995. Sunspots and the solar constant. Journal of the British Astronomical Association. 105(4), pp.165-166.

Brian Williams. 2023 Sunspots and Cloud Formation. [Online] [Accessed: August 2, 2023]. Available from: https://www.briangwilliams.us/weather-change/sunspots-and-cloud-formation.html.

NASA. 2022. Where Does the Sun's Energy Come From?. [Online]. [Accessed: June 27, 2023]. Availalbe from: https://spaceplace.nasa.gov/sun-heat/en/.

Gueymard, C.A. 2018. Solar Spectral Irradiance Variations. [Online]. 1st ed. Solar Energy:Elsevier. [Accessed: July 25, 2023] Available from: https://www.sciencedirect.com/science/article/abs/pii/S0038092X18303463.

Neutron Monitor Database. 2022. Cosmic Ray Data. [Accessed: July 26, 2023]. Available from: https://neutronm.bartol.udel.edu/catch/cr3.html.

ScienceDaily. 2016. Sun's Impact on Climate. [Accessed: July 12, 2023]. Available from: https://www.sciencedaily.com/releases/2016/08/160825113235.htm.

National Academies of Sciences, Engineering, and Medicine. 2012. The Effects of Solar Variability on Earth's Climate: A Workshop Report. [Online]. 1st ed. Washington:National Academies Press. [Accessed: July 12, 2023] Available from: https://www.nap.edu/read/13519/chapter/3

Sodankylä Geophysical Observatory - University of Oulu. 2022. Cosmic Rays Data. [Accessed: July 5, 2023]. Available from: https://cosmicrays.oulu.fi/.

Laboratoire de Physique Subatomique et de Cosmologie. 2022. Cosmic Ray Data Base. [Accessed: June 20, 2023]. Available from: https://lpsc.in2p3.fr/crdb/.

Laboratory for Atmospheric and Space Physics - University of Colorado. 2022. Solar Irradiance Data. [Accessed: July 7, 2023]. Available from: https://lasp.colorado.edu/sorce/data/ssi-data/.

Solar Influences Data Analysis Center. 2022. The Sunspot Number and Long-term Solar Observations. [Accessed: June 31, 2023]. Available from: www.sidc.be/SILSO/INFO/sndtotcsv.php.

Centre for Environmental Data Analysis. 2022. Cloud Property Data. [Accessed: July 15, 2023]. Available from: https://data.ceda.ac.uk/neodc/esacci/cloud/data/version3/L3C/AVHRR-AM/v3.0.

Exploratorium. 2022. Sunspot Research. [Accessed: July 8, 2023]. Available from: https://www.exploratorium.edu/sunspots/research.html.

Wikipedia. 2001. Sunspot. [Accessed: June 30, 2023]. Available from: https://en.wikipedia.org/wiki/Sunspot.

NASA. 2021. Solar Cycles. [Accessed: July 7, 2023]. Available from: https://spaceplace.nasa.gov/solar-cycles/en/.

Wikipedia. 2022. Solar Cycle. [Accessed: July 7, 2023]. Available from: https://en.wikipedia.org/wiki/Solar_cycle.

Hathaway, D.H. 2015. TheSolarCycle. Living Rev. Sol. Phys. 7(1),pp.7-73.

NOAA Space Weather Prediction Center. 2022. Space Weather's Impact on Climate. [Accessed: September 8, 2023]. Available from: https://www.swpc.noaa.gov/impacts/space-weather-impacts-climate.

Australian Government Bureau of Meteorology. 2022. SPACE WEATHER IMPACTS ON CLIMATE. [Accessed: July 28, 2023]. Available from: https://www.sws.bom.gov.au/Educational/1/3/3.

Astronomy & Astrophysics Journal. 2021. Solar Orbiter First Results (Cruise Phase). Cosmic-ray flux predictions and observations for and with Metis on board Solar Orbiter. 656(A15).

Solar Influences Data Analysis Center. 2023. Detailed description about the data or information. [Accessed: July 16, 2023]. Available from: https://www.sidc.be/.

PMOD/WRC Davos. 2022. Solar Constant: Construction of a Composite Total Solar Irradiance (TSI) Time-Series from 1978 to the Present. [Accessed: July 7, 2023]. Available from: https://www.pmodwrc.ch/en/research-development/solar-physics/tsi-composite/.

Encyclopedia. 2023. Solar Constant and Cycle. [Accessed: July 7, 2023]. Available from: https://www.encyclopedia.com/environment/encyclopedias-almanacs-transcripts-and-maps/solar-constant-cycle.

NASA. 2021. All About the Sun. [Online]. [Accessed: July 3, 2023]. Available from: https://spaceplace.nasa.gov/all-about-the-sun/en/.

Appendix

The Python code utilized for processing, analyzing, and visualizing all the data sets involved in this dissertation is hosted on a GitHub repository for clarity and ease of access.

textbfRepository Link: `https://github.com/Shreyansh83/Sun-Climate-relation`