**I. Introduction**

**A. Background**

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.

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. Below image 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.

A diagram of the solar system

Description automatically generated

Figure 1 https://spaceplace.nasa.gov/sun-heat/en/

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 of approximately ±0.05% [source](https://adsabs.harvard.edu/full/1995JBAA..105..165W). Intriguingly, during sunspot minima, Earth is bombarded with heightened cosmic rays, which seem to foster increased cloud formation, consequently nudging the surface temperatures downward [source](https://www.briangwilliams.us/weather-change/sunspots-and-cloud-formation.html).

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. 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. This included data from ACRIM, PMOD composites, as well as recent measurements from SORCE-TIM, TCTE-TIM, and PICARD-PREMOS instruments. The results of this in-depth analysis led to a revised solar constant value of 1361.1 W/m^2, albeit with a standard uncertainty of 0.5 W/m^2 [source](https://www.sciencedirect.com/science/article/abs/pii/S0038092X18303463). This study accentuated how solar activities, especially sunspots and faculae, act as dual forces. While sunspots darken and reduce solar output, faculae brighten it. The study further emphasized the sheer magnitude of the Sun's influence on Earth, affecting a multitude of geobiological processes, from weather patterns to overarching climatic conditions.

Venturing beyond the immediate realm of the Sun introduces us to the world of cosmic rays. These 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 [source](https://neutronm.bartol.udel.edu/catch/cr3.html). 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. By producing electrically charged molecules or ions in Earth's atmosphere, cosmic rays can potentially enhance aerosol formation, which can act as nuclei for cloud droplets. This intricate dance between solar eruptions, cosmic rays, and cloud cover further underscores the Sun's indirect influence on Earth's climate, especially through phenomena known as 'Forbush decreases' [source](https://www.sciencedaily.com/releases/2016/08/160825113235.htm).

Sunspots, with their cyclical nature, hold potential clues about Earth's climatic rhythms. Historical data reveals intriguing patterns. For instance, during the 17th century, a prolonged period known as the Maunder Minimum saw the Sun almost devoid of sunspots. This coincided with a mini-ice age on Earth. Such patterns prompt questions about the quantitative relationships between sunspots, faculae, and Earth's climate. The puzzle becomes even more complex when considering the dual effects of sunspots: while they directly reduce the solar output, their associated solar wind can amplify cosmic ray interactions, thereby influencing cloud cover and Earth's albedo [source](https://www.briangwilliams.us/weather-change/sunspots-and-cloud-formation.html).

The tangled narrative 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 strides have been made in understanding these phenomena, much remains in the realm of the unknown.

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. The Sun's cycles, especially the 11-year sunspot cycle, have discernible imprints on our climate records. Yet, extrapolating these patterns to understand future climatic shifts demands rigorous scientific scrutiny and advanced modeling techniques [source](https://www.nap.edu/read/13519/chapter/3#6).
2. **Cosmic Rays - More than Just High-Energy Particles**: The impact of cosmic rays on Earth's climate is multi-faceted. Beyond influencing cloud formation, their potential role in affecting Earth's albedo (reflectivity) is a topic warranting deeper exploration. 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. Their role in modulating cloudiness and the subsequent climatic implications are areas ripe for further investigation [source](https://www.nap.edu/read/13519/chapter/3#6).
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. How these variations drive atmospheric processes and potentially cascade down to influence the lower atmosphere and surface climates remains an exciting frontier in sun-climate research [source](https://www.nap.edu/read/13519/chapter/3#6).

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 imperative 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 embark on 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 cusp of further 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.

**B. 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 profound implications for climate science, policymaking, and our collective efforts to mitigate and adapt to climate change.

**C. 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?

**D. Objectives**

The primary objectives of this dissertation are to:

* Analyze individual datasets of solar magnetic activities.
* Investigate correlations between various solar magnetic data.
* Compare identified correlations with specific climate data.
* Interpret the findings and explore potential causal mechanisms.
* Evaluate the implications of the study's results within the broader context of climate science.

**E. Scope and Limitations**

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. 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.

While the analysis offers comprehensive insights into the subject, it is crucial to recognize and understand the limitations that define the boundaries of this study. These limitations include:

* Solar Spectrum Irradiance (SSI) data is available only for the period from 2003 to 2020, as no measurements of SSI were taken before 2003. This restriction narrows the scope of the study, limiting the examination of correlations between SSI, solar magnetic activities, and cloud factors to these specific years.
* Additionally, it is vital to emphasize that correlation does not equate to causation. The Earth's climate system is intricate, involving various confounding factors that may contribute to the observed patterns. The complexity of analyzing these interrelated elements falls beyond the scope of this analysis.

**F. Structure of the Dissertation**

The dissertation is structured into eight 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.

**II. Methodology**

**A. Data Sources**

***1. Cosmic Ray Data***

a. Source: Oulu Cosmic Ray Station

b. Website: cosmicrays.oulu.fi

c. 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.

d. Application: Understanding cosmic ray patterns is vital in analyzing their potential interactions with solar magnetic activities and their subsequent effects on climate.

***2. Detailed Cosmic Rays Properties***

a. Source: Cosmic Ray Database at LPSC

b. Website: lpsc.in2p3.fr/crdb/

c. 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.

d. Application: These properties were essential in correlating cosmic ray activities with solar activities and understanding how they may influence climate patterns.

***3. Solar Irradiance Data***

a. Source: Laboratory for Atmospheric and Space Physics (LASP)

b. Website: lasp.colorado.edu/sorce/data/ssi-data/

c. 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:

|  |  |
| --- | --- |
| INSTRUMENT NAME | SPECTRUM RANGE |
| XPS | 0-40 nm |
| SOLSTICE | 115-310 nm |
| SIM | 240-2416 nm |

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

***4. Sunspot Data***

a. Source: Royal Observatory of Belgium's Solar Influences Data Analysis Center (SIDC)

b. Website: www.sidc.be/SILSO/INFO/sndtotcsv.php

c. 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.

d. Application: Analyzing sunspots helps in understanding solar magnetic activities and their potential influence on cosmic rays and climate.

***5. Cloud Data***

a. Source: ESA Climate Change Initiative (CCI)

b. Website: data.ceda.ac.uk/neodc/esacci/cloud/data/version3/L3C/AVHRR-AM/v3.0

c. Description: The cloud cover data provides information on cloud formations, types, altitudes, and optical properties across different regions and time periods.

d. Application: Analyzing cloud data is essential in understanding the relationship between solar activities and cloud cover, a significant factor in climate change.

**B. Data Analysis Methods**

**Individual Data Analysis**

a. 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. Potential Graph: A time-series plot showing the variation of cosmic ray intensity over time.

b. Solar Irradiance Analysis:

Solar irradiance data was thoroughly analyzed to understand the changes in solar energy reaching the Earth. This included an examination of daily, monthly, and yearly variations. Potential Graph: A line graph showing the trend of solar irradiance over different time periods.

c. 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. Potential Graph: A scatter plot showing the relationship between sunspot numbers and solar cycles.

d. 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.

**Correlation Analysis**

a. 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. Potential Graph: A correlation matrix showing the relationships between cosmic rays, solar irradiance, and sunspots.

b. Solar Activities and Climate Variables:

Correlations were analyzed 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. Potential Graph: Scatter plots showing correlations between different solar and climate variables.

**Climate Change Analysis**

a. Modeling Solar Influence:

A comprehensive climate change analysis was performed by modeling the potential impacts of solar activities on climatic factors. Various statistical and machine learning models were employed to predict potential climate changes based on solar variables. Potential Graph: Residual plots and model fit graphs for various models used.

b. Sensitivity Analysis:

Sensitivity analysis was conducted to understand how different climate variables respond to changes in solar activities. This helped in understanding the robustness and reliability of the models used. Potential Graph: Sensitivity curves showing the response of climate variables to changes in solar activities.

**C. 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.

These tools and libraries enabled the handling of large datasets and the implementation of various data analysis methods.

**III. Data Analysis**

**SSI across Different Spectra**

**A. 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 1.0, 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 |

Out of 5689117 entries, it was found out that 1360 rows were having quality flag as 1.0 as displayed in below graph.A screen shot of a graph

Description automatically generated

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 -

**B. In-depth analysis – UV SPECTRUM**

The below graph shows the yearly mean of UV SSI from 2002 to 2020,

A graph with blue lines

Description automatically generated

This graph provides an annual 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.

A screenshot of a graph

Description automatically generated

The above figure 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:

Observed=Trend+Seasonality+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 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.

Analyzing UV Spectrum Stationarity

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.

**C. In-depth analysis – VISIBLE SPECTRUM**

In this section we will investigate SSI data for visible spectrum. The below graph is time-series decomposition for the same.

A screenshot of a graph

Description automatically generated

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, 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.

A graph with a line going up

Description automatically generated

The graph 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, 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.

**D. In-depth analysis – INFRARED SPECTRUM**

In this section we will investigate SSI data for infrared spectrum. The below graph is time-series decomposition for the same.

A screenshot of a computer screen

Description automatically generated

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

The ADF statistic of -7.92 is well below the critical values at all confidence levels, indicating strong evidence against the null hypothesis of a unit root. The p-value is also considerably below the 0.05 threshold, confirming the stationarity of the series.

A graph with blue lines

Description automatically generated

The above graph shows yearly trends in the infrared spectrum of Solar Spectral Irradiance (SSI). Upon closer inspection, several key features stand out:

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.

Notable Anomalies:

While the data is generally stable, there are years where the irradiance values appear to deviate from the norm:

Early 2000s: There's a noticeable uptick in irradiance levels during the early 2000s, possibly corresponding to the peak of Solar Cycle 23.

Late 2010s: A mild dip is observed in the late 2010s, aligning with the minimum of Solar Cycle 24.

Year 2020: 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.

**Sunspot Activity**

**A. Data Cleaning and Transformation**

Like done in SSI data, the data was loaded from the file name - ‘SN\_d\_tot\_V2.0.txt’, and then analyzed. The dataset has data from Year 1818 to 2023 latest month at the time of downloading the file and has about 75026 rows.

Upon loading following columns were found in the dataset:

* Year – Calendar year, data ranges from year 1818 - 2023
* Month – Calendar month, data range has all months, i.e., 1-12
* Day – calendar day, data ranges from all days in a month, i.e., 1-31
* Date\_in\_fraction\_of\_year: Date in fraction of year, data range from 1818.001 to 2023.412.
* sunspot\_number: Daily total sunspot number. -1 value indicates that no number is available for that day (missing value). Observed sunspot number ranges 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 pretty straightforward and understandable from the name, now as mentioned in definition of sunspot number column, that -1 indicates missing values, as seen in below graph we observe that approx. 4% (3247 rows) of data is missing.

A screenshot of a graph

Description automatically generated

Since our main goal is to analyze the sunspot number itself from this dataset, we simply removed these rows as they were not contributing to loss of information in any way.

Furthermore, we can observe if any other columns has null/empty values or not:

|  |  |
| --- | --- |
| Column Name | Null/Empty values |
| Year | 0 |
| Month | 0 |
| Day | 0 |
| Date\_in\_fraction\_of\_year | 0 |
| sunspot\_number | 0 |
| standard\_deviation | 0 |

The above table clearly suggests that there are no null values in our dataset, and we can go ahead with formatting our data.

To ease our analysis, we introduced a new calculated column from year, month and day called ‘Date’, with format datetime. Other than that, all other columns are correctly formatted by python so further formatting is not done on the dataset.

In the next section we will dive into in-depth analysis of sunspot dataset

**B. In-depth analysis**

A blue lines on a white background

Description automatically generated

The trend chart 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.

A screenshot of a computer screen

Description automatically generated

The above graph representing the time series decomposition of sunspot data provides a granular view, segregating the data into various components, some of the important findings :

Trend: The trend seems to capture long-term changes in sunspot activity, which could be linked to longer solar cycles or other forms of solar variability. While generally consistent, there are years where the trend shows a significant deviation, either increasing or decreasing sharply.

Seasonality: The seasonality aspect captures short-term, periodic fluctuations in sunspot numbers. These could 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 is consistent with what one would expect for a cyclical phenomenon like sunspot activity.

**Cosmic Ray**

**A. Data cleaning and Transformation**

After following the same procedure of data loading in previous datasets, we get following columns from cosmic ray dataset:

* Timestamp: This column contains the date and time of the data record, formatted as "YYYY-MM-DDTHH:MM:SSZ".
* FractionalDate: This represents the date as a fraction of the year, providing a continuous scale that facilitates easier analysis.
* UncorrectedCountRate[cts/min]: This is the raw count rate of cosmic rays in counts per minute.
* CorrectedCountRate[cts/min]: The count rate after applying corrections for factors like atmospheric pressure.
* Pressure[mbar]: Atmospheric pressure in millibars at the time of the data record.

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

As seen from the above table, no missing values were found in the dataset, as indicated by the null count for each column being zero.

Similarly, to our previous datasets, the 'Timestamp' column was converted into datetime format. Other than this column the data types of other columns are consistent with the kind of data they are supposed to hold.

**B. In-depth Analysis**

In this section we will see the stationarity analysis and observe the yearly values of cosmic rays.

A graph with blue lines

Description automatically generated

The yearly distribution of cosmic rays 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.

A screenshot of a graph

Description automatically generated

The time series decomposition of cosmic ray data reveals essential components, 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.

Let us see the ADF test results:

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

ADF Statistic: The ADF statistic of -14.77 is much lower than all the critical values, strongly suggesting that the series is stationary.

p-value: The extremely low p-value of 2.34×10^−27, provides strong evidence against the null hypothesis of non-stationarity.

Critical Values: 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.

**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 Height (cth)**: The height of the cloud from the Earth's surface.
* **Cloud Top Pressure (ctp)**: Atmospheric pressure at the cloud's highest point.
* **Cloud Emissivity (cee)**: 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. Let us see how many of our columns have null/NA values:

|  |  |
| --- | --- |
| **Column** | **Null/Empty Values** |
| time | 0 |
| cfc | 0 |
| ctt | 8173 |
| stemp\_cloudy | 8173 |
| cth | 8173 |
| ctp | 8173 |
| cee | 8173 |
| cla\_vis006 | 5910513 |

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 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 dept 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.

**CFC**

The time-series decomposition of Cloud Fraction reveals essential insights into its trend, seasonality, and residuals.

A graph of blue lines

Description automatically generated with medium confidence

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.

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

A graph with a line going up

Description automatically generated

Upon observing the graph, 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.

**CTT**

This graph 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.

A graph with blue lines and dots

Description automatically generated

The Time Series Decomposition graph ("ctt\_trend.png") 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.

A graph of blue lines

Description automatically generated with medium confidence

The Augmented Dickey-Fuller (ADF) test was used to assess the stationarity of the CTT time series.

* 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 common alpha level of 0.05, indicating that we fail to reject the null hypothesis. This suggests that the time series data for CTT is non-stationary.

**stemp\_cloudy**

The yearly distribution of skin temperature over cloudy regions (depicted in **stp\_yearly.png**) 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.

A graph with blue lines and dots

Description automatically generated

To understand the underlying patterns within the **stemp\_cloudy** variable, a time series decomposition is performed (see **stp\_trend.png**). The decomposition shows three components:

* **Trend Component**: The trend component reveals a near-static pattern. There is no significant upward or downward movement, mirroring what was observed in the yearly distribution.
* **Seasonal Component**: Seasonal fluctuations do exist but are within a predictable range, suggesting that the variable is influenced by cyclical seasonal factors.
* **Residual Component**: The residuals, or the noise in the data, also appear to be fairly consistent over time, without any sudden or unpredictable spikes.

A diagram of a graph

Description automatically generated with medium confidence

The Augmented Dickey-Fuller (ADF) test, a statistical method used to test the stationarity of a time series, yields 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. This might imply the necessity for differencing or transformations before conducting further time series analyses.

**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 an invaluable method for understanding underlying patterns.

A line graph of different types of lines

Description automatically generated with medium confidence

**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.

**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.

**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.

A graph with a line going up

Description automatically generated

The yearly distribution of CTP 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.

**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

A diagram of a graph

Description automatically generated with medium confidence

* **Trend**: The trend component shows a relatively stable pattern, with minor fluctuations over the years. However, a more detailed analysis is required to determine the significance of these fluctuations.
* **Seasonality**: The seasonality component reveals no discernible seasonal pattern, indicating that CTH does not significantly vary with the seasons.
* **Residuals**: The residual component represents the noise in the data after the extraction of the trend and seasonality components.

A graph with a line going up

Description automatically generated

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

* **Variability**: The plot shows minor variability in CTH over the years, with some years showing slightly higher values than others.
* **Outliers**: No significant outliers or extreme values are evident, indicating a consistent dataset.

**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.

A line graph of different types of lines

Description automatically generated with medium confidence

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.

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.

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.

A graph with a line going up

Description automatically generated

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.

**CEE**

A line graph of blue lines

Description automatically generated with medium confidence

The time-series decomposition graph for CEE shows three distinct parts: trend, seasonality, and residuals.

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.

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.

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.

A graph with blue lines

Description automatically generated

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. This means that the data needs to be differenced or transformed to make it stationary if it is to be used for forecasting or other time-series analyses.

**V. 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.

**A. Correlation Between Solar Magnetic Activities**

**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 1) showcases the yearly distribution of both sunspots and UV solar irradiance from 2002 onwards.

A graph with lines and dots

Description automatically generated

*Figure 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.

**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.

The attached graph (Figure 2) illustrates the yearly distribution of both sunspots and solar irradiance in the visible range from 2002 onwards.

A graph with a line graph

Description automatically generated

*Figure 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.

**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 3) showcases the annual trends of sunspot numbers and solar irradiance in the infrared spectrum, spanning from the year 2002 onwards.

A graph with lines and dots

Description automatically generated

*Figure 3: Yearly trends of sunspots (blue) and solar irradiance in the infrared range (red).*

From a visual inspection of the graph, it is apparent that there are some parallels between the sunspot numbers and the infrared solar irradiance. However, it's also evident that the correlation isn't as robust as one might anticipate given the profound role of infrared radiation in our planet's thermal balance.

Quantitatively, the correlation coefficient is computed to be *r*=0.331. This value indicates a moderate positive relationship between sunspot numbers and solar irradiance in the infrared range. While the two variables tend to move in sync to some extent, the relationship is not as pronounced as one might expect.

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. While sunspots indeed influence solar irradiance across different wavelengths, the magnitude and characteristics of these influences can differ depending on the specific spectral range under consideration.

**4. Correlation Between Sunspots and Cosmic Rays (Corrected)**

An inverse relationship has long been posited between sunspots and cosmic rays, based on the idea that increased solar magnetic activity, evidenced by a high number of sunspots, would deflect more cosmic rays from reaching Earth.

The provided graph (Figure 4) visualizes the annual trends of sunspot numbers and the corrected count of cosmic rays from the year 2002 onward.

A graph of a sunspot

Description automatically generated

*Figure 4: Yearly trends of sunspots (blue) and corrected cosmic ray count (red).*

Observing the graph, there is a noticeable anti-correlation between the sunspot numbers and the cosmic ray count. Periods of high sunspot activity correspond with reduced cosmic ray counts and vice versa. This pattern aligns well with the theoretical understanding of the sun's magnetic influence on cosmic rays.

The calculated correlation coefficient between these two parameters is *r*=−0.623. This negative value reaffirms the inverse relationship between sunspot numbers and cosmic ray counts. The magnitude of this coefficient indicates a fairly strong negative correlation, suggesting that as sunspot numbers increase, the count of cosmic rays reaching Earth tends to decrease, and vice versa.

This observed relationship provides crucial insights into the complex interplay between the sun and cosmic ray particles. The strong inverse correlation underscores the sun's role as a modulator of cosmic ray intensity, with its magnetic field acting as a shield that deflects a significant portion of these high-energy particles.

**B. Correlation between cloud properties and solar magnetic activities**

**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. A screenshot of a computer

Description automatically generated

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.

**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.

A screenshot of a computer

Description automatically generated

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.

**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).

A screenshot of a computer

Description automatically generated

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.
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.

**VI. Interpretation and Findings**

**A. 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.

**B. 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.

**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.

**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.

**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.

**C. Comparison with Previous Studies**

**1. Solar Spectral Irradiance (SSI) and Cloud Properties**

Historical research has consistently emphasized the Sun's pivotal role in influencing Earth's climate. However, the relationship between specific wavelengths of the Sun's output (like SSI) and cloud properties has been a more recent avenue of exploration.

* **Ultraviolet (UV) Solar Spectral Irradiance**:
  + Earlier studies have highlighted the UV radiation's capacity to influence stratospheric temperatures and ozone concentrations. Our findings, which indicate a correlation between UV SSI and cloud properties, notably cloud fraction (cfc) and cloud top temperature (ctt), align with these studies, suggesting that the effects of UV radiation might cascade down to the troposphere, influencing cloud properties.
  + Some researchers have posited that UV-induced changes in the stratosphere could alter atmospheric circulation patterns, potentially influencing cloud formation and distribution. Our observed correlations provide further credence to this theory.
* **Visible Solar Spectral Irradiance**:
  + Traditional climate models have recognized the importance of visible light in Earth's energy balance. The observed strong correlation between visible SSI and cloud albedo in this study resonates with previous findings that suggest changes in cloud reflectivity can have pronounced effects on surface temperatures.
  + Earlier research emphasizing the role of cloud albedo in climate feedback mechanisms finds support in our results, underlining the significance of visible light-cloud interactions.
* **Infrared (IR) Solar Spectral Irradiance**:
  + The mixed correlations observed in our study between IR SSI and cloud properties present a more nuanced picture than some earlier research. While IR radiation's role in the greenhouse effect is well-established, its interactions with clouds, especially in influencing properties like cloud top temperatures or cloud lifetimes, are areas where our findings might challenge or augment existing knowledge.

**2. Sunspots and Cloud Properties**

The relationship between sunspot activity and Earth's climate has been a topic of interest for centuries.

* Traditional wisdom, backed by some historical data, suggests that periods of low sunspot activity (like the Maunder Minimum) were associated with cooler global temperatures. Our findings, which indicate a weak positive correlation between sunspots and certain cloud properties, offer a nuanced perspective. While there may be a relationship, the strength of this association, as observed in our study, suggests that other factors also play a substantial role in influencing cloud properties.
* Some past studies have proposed mechanisms through which sunspots might influence Earth's climate, such as changes in solar output or the modulation of cosmic ray influx. Our results, showing minimal correlation between sunspots and properties like cloud top height (cth) and cloud top pressure (ctp), provide a more intricate portrayal, suggesting that the direct impacts of sunspots on certain cloud properties might be limited.

**3. Cosmic Rays and Cloud Properties**

The potential role of cosmic rays in influencing cloud formation has been a topic of considerable debate.

* Certain studies have proposed a mechanism where cosmic rays, by ionizing air molecules, promote the formation of aerosols, which can act as cloud condensation nuclei. Our findings, which indicate a positive correlation between cosmic rays and cloud fraction (cfc), resonate with this theory.
* However, not all previous research aligns perfectly with our observations. For instance, while some studies have reported a strong link between cosmic rays and cloud cover, others have found minimal to no relationship. Our results, particularly the weak correlation between cosmic rays and surface temperature under cloudy conditions (stemp\_cloudy), suggest that while cosmic rays might play a role in cloud formation, their influence on resultant temperature changes at the Earth's surface could be more complex.

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.

**D. 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.

**VII. 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.

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.
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.
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.

**C. 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.

**IX. References**

**X. Appendices**