Predicting Geomagnetic Activity

Name : Kalpa Harindra Henadhira Arachchige

GA DSIR 8-14

Capstone project

11.06.2023

Outline

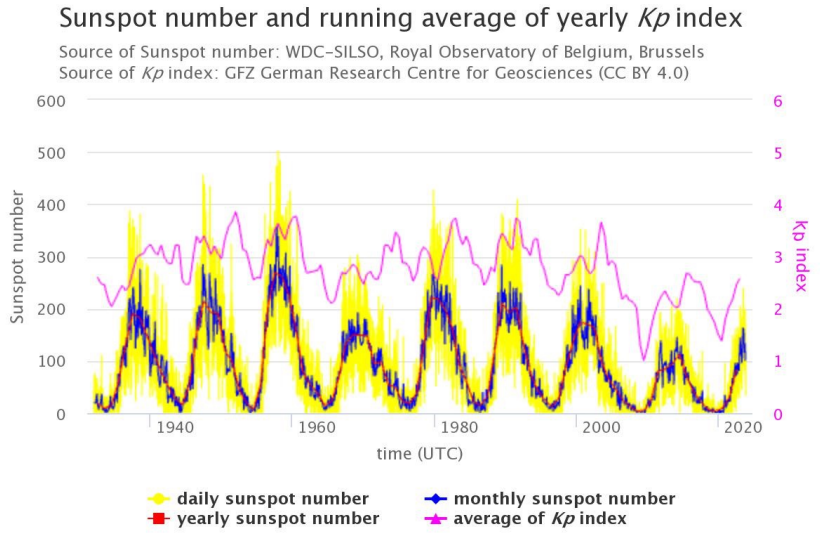
- Problem Statement
- Sunspot solar cycle
- Solar wind
- EDA and feature engineering
- Modeling
- Model evaluation
- Conclusion and future work

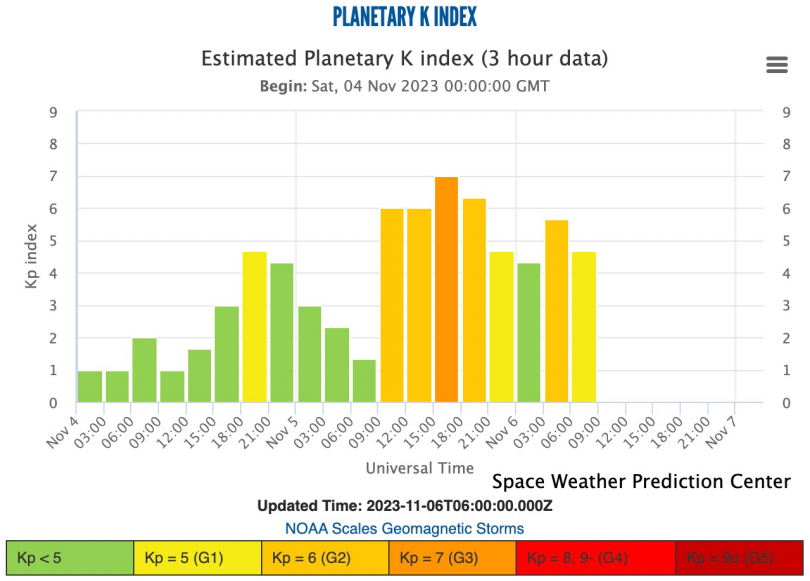
Problem Statement

- The main goal of this project is to identify the relationship between the solar indices (sunspot number, solar radio flux) and the geomagnetic indices (K index) and to track geomagnetic activity near Earth
- This study will also ensure the identification of the effects of auroras based on the solar wind properties at 1~au, derived from solar activity
- For this project, we have utilized sunspot data and solar wind data from January 1, 1992, to December 31, 2022
- To test the model, we selected data from January 1, 2023, to October 27, 2023

What is the Kp index?

- Global magnetic disturbance in near earth
- Used to indicate the possibility of viewing the aurora
- Higher values of Kp (0-9) indicate that the aurora can be viewed at mid-latitudes





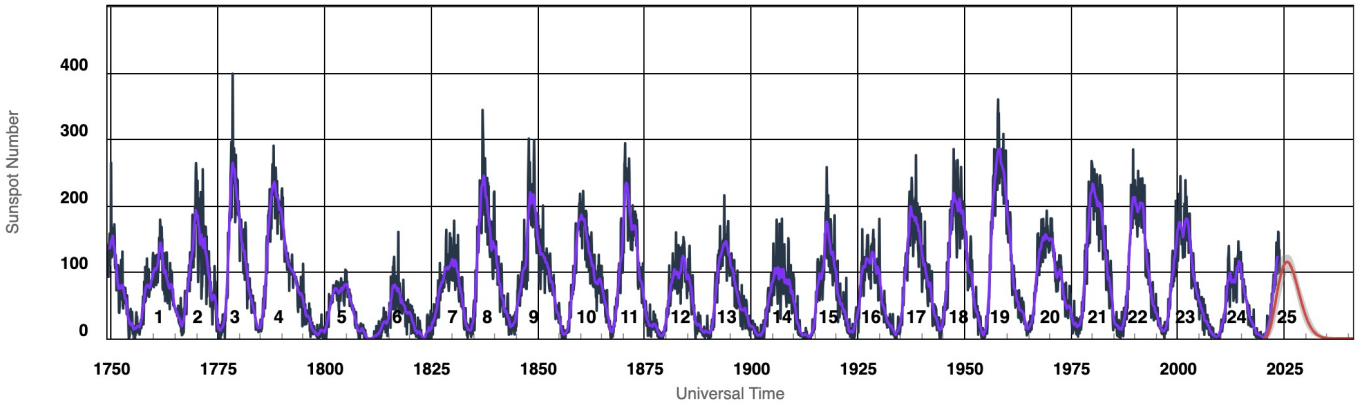
https://www.swpc.noaa.gov/communities/aurora-dashboard-experimental

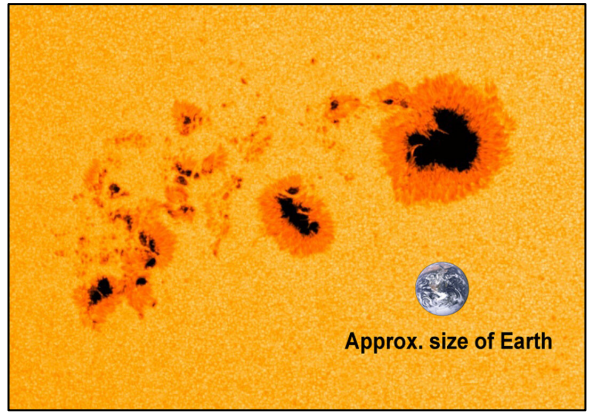
NOAA space weather scales

Scale	Description	Effect	Physical measure
G 5	Extreme	**Power systems:** Widespread voltage control problems and protective system problems can occur, some grid systems may experience complete collapse or blackouts. Transformers may experience damage. **Spacecraft operations:** May experience extensive surface charging, problems with orientation, uplink/downlink and tracking satellites. **Other systems:** Pipeline currents can reach hundreds of amps, HF (high frequency) radio propagation may be impossible in many areas for one to two days, satellite navigation may be degraded for days, low-frequency radio navigation can be out for hours, and aurora has been seen as low as Florida and southern Texas (typically 40° geomagnetic lat.).	Kp = 9
G 4	Severe	**Power systems:** Possible widespread voltage control problems and some protective systems will mistakenly trip out key assets from the grid. **Spacecraft operations:** May experience surface charging and tracking problems, corrections may be needed for orientation problems. **Other systems:** Induced pipeline currents affect preventive measures, HF radio propagation sporadic, satellite navigation degraded for hours, low-frequency radio navigation disrupted, and aurora has been seen as low as Alabama and northern California (typically 45° geomagnetic lat.).	Kp = 8, including a 9-
G 3	Strong	**Power systems:** Voltage corrections may be required, false alarms triggered on some protection devices. **Spacecraft operations:** Surface charging may occur on satellite components, drag may increase on low-Earth-orbit satellites, and corrections may be needed for orientation problems. **Other systems:** Intermittent satellite navigation and low-frequency radio navigation problems may occur, HF radio may be intermittent, and aurora has been seen as low as Illinois and Oregon (typically 50° geomagnetic lat.).	Kp = 7
G 2	Moderate	**Power systems:** High-latitude power systems may experience voltage alarms, long-duration storms may cause transformer damage. **Spacecraft operations:** Corrective actions to orientation may be required by ground control; possible changes in drag affect orbit predictions. **Other systems:** HF radio propagation can fade at higher latitudes, and aurora has been seen as low as New York and Idaho (typically 55° geomagnetic lat.).	Kp = 6
G 1	Minor	**Power systems:** Weak power grid fluctuations can occur. **Spacecraft operations:** Minor impact on satellite operations possible. **Other systems:** Migratory animals are affected at this and higher levels; aurora is commonly visible at high latitudes (northern Michigan and Maine).	Kp = 5

https://www.swpc.noaa.gov/noaa-scales-explanation

Sunspot solar cycle





- Solar cycles have an average duration of about 11 years
- Solar maximum and solar minimum refer to periods of maximum and minimum sunspot counts
- Cycles span from one minimum to the next
- **Phases of the Solar Cycle:** Solar minimum, Rising phase, Solar maximum, Declining phase

Solar wind and solar wind properties at 1~au

- Solar wind consists of particles emitted from the sun's atmosphere with a sufficiently large velocity to escape from the sun's gravitational acceleration
- Solar wind can mainly be divided into slow and fast wind. The common state of the solar wind is slow wind
- Slow wind 200 km/s - 450 km/s fast wind 450 km/s - 850 km/s

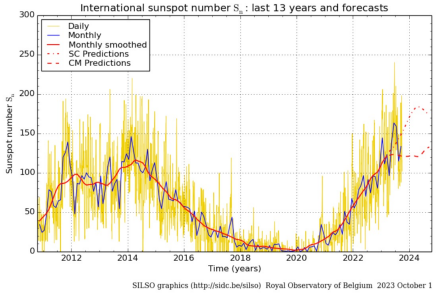
Solar wind properties :

1. Plasma density (Np)
2. Magnetic field strength (B)
3. Plasma temperature (T[K])
4. Solar Wind Speed (Ur)

Collecting data

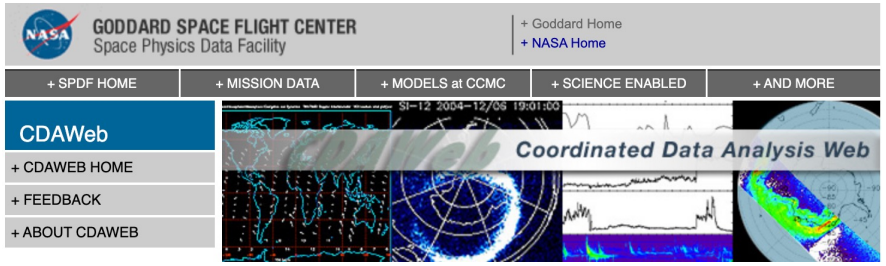
SILSO

Sunspot number series: latest update



Sunspot Index and Long-term Solar Observations

- Solar wind data from OMNI observations at 1~au (Coordinated Data Analysis Web)
- SILSO for sunspot data
- Gfz-potsdam for solar and geomagnetic indices



- https://cdaweb.gsfc.nasa.gov/index.html/
- https://www.gfz-potsdam.de/en/section/geomagnetism/data-products-services/geomagnetic-kp-index
- https://www.sidc.be/SILSO/home/

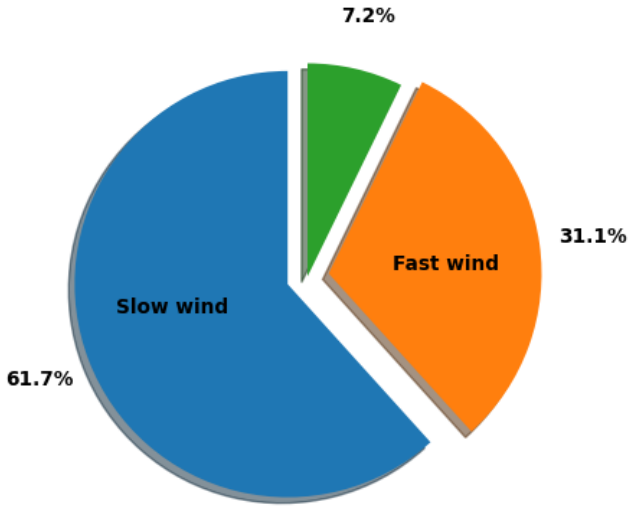
Categorizing the Solar wind data 1~au

slow wind 200 km/s - 450 km/s

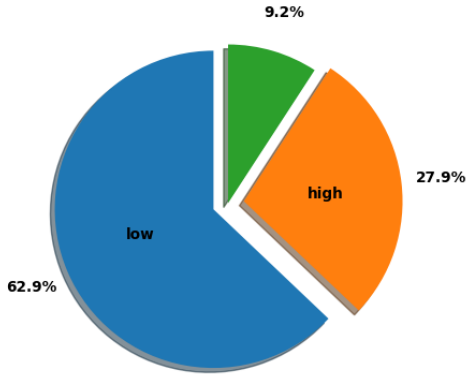
fast wind 450 km/s - 850 km/s

extreme > 850km/s

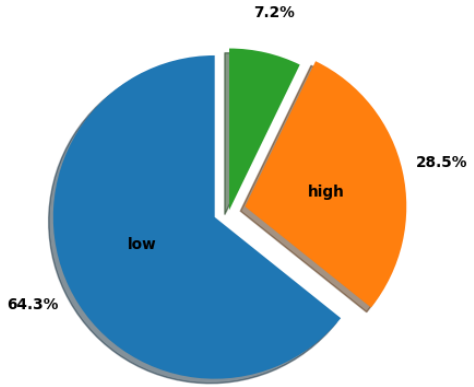
- Arranging the other solar wind parameters in order to maintain roughly an equal percentage of wind speed



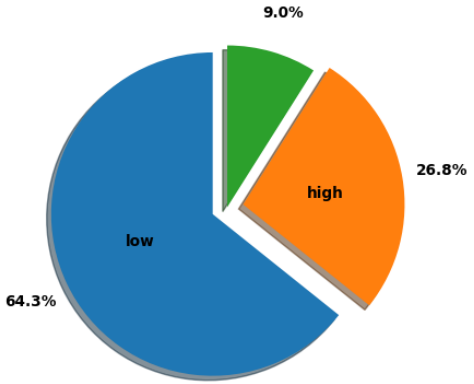
Wind speed



Plasma density

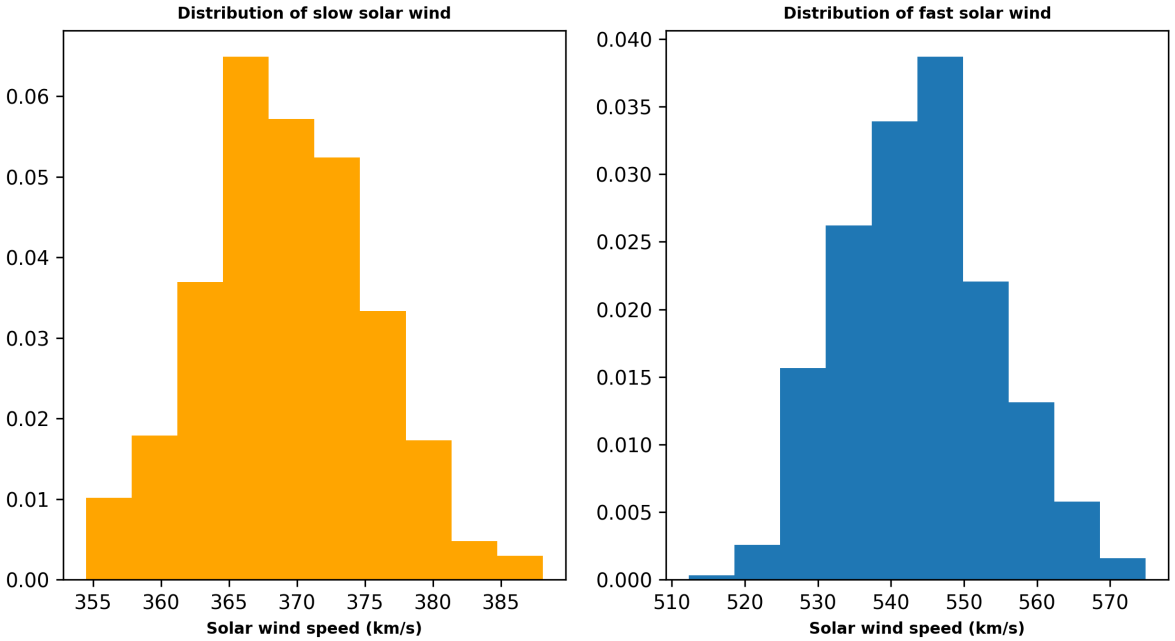


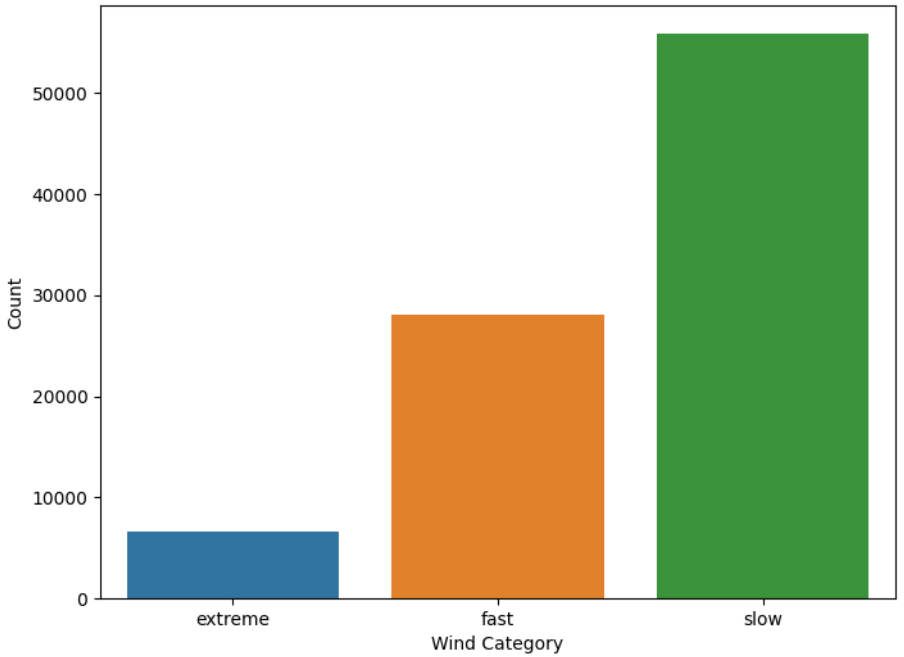
Magnetic field



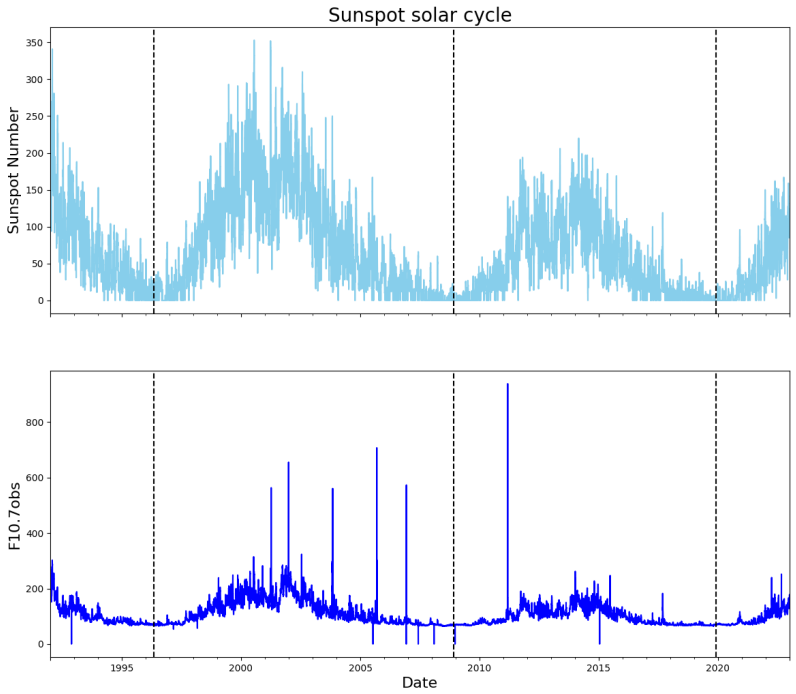
Plasma temperature

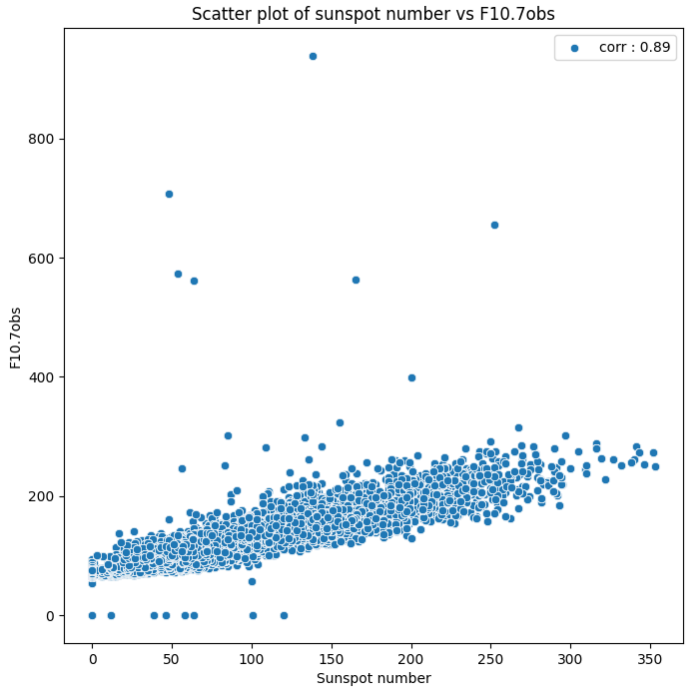
Solar wind Speed (1992.01 – 2022.12)

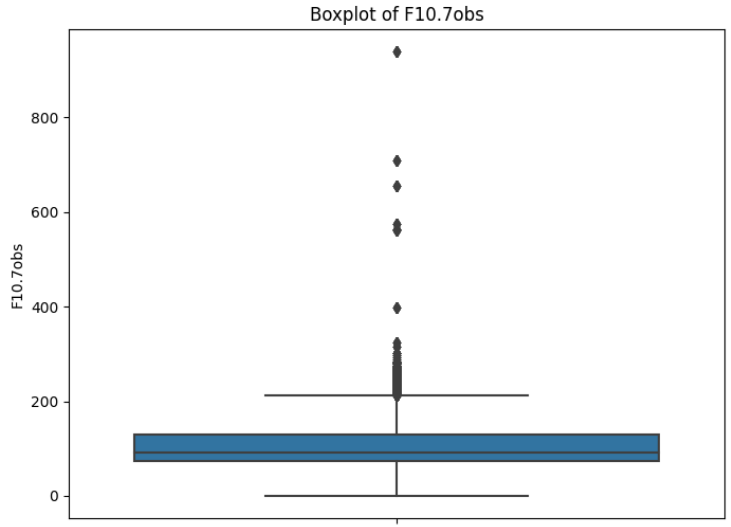




F10.7 index



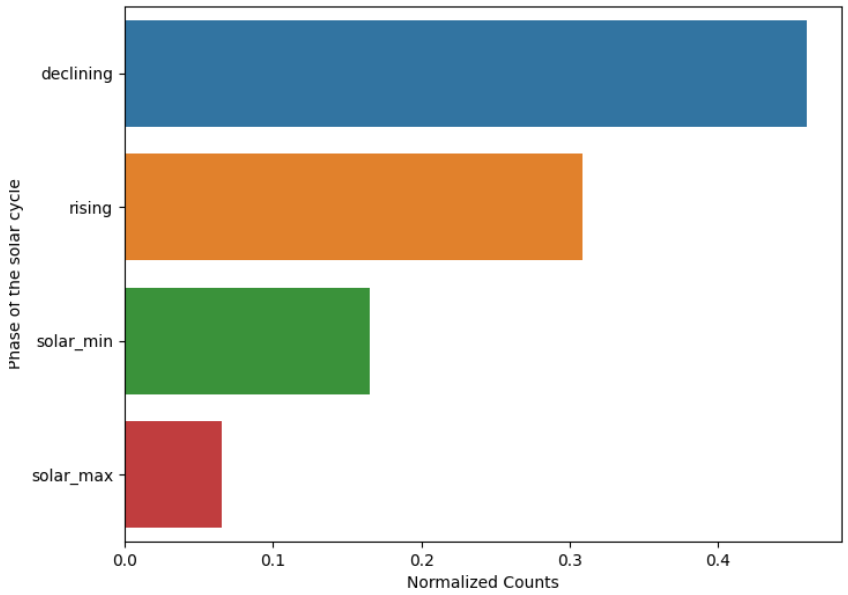


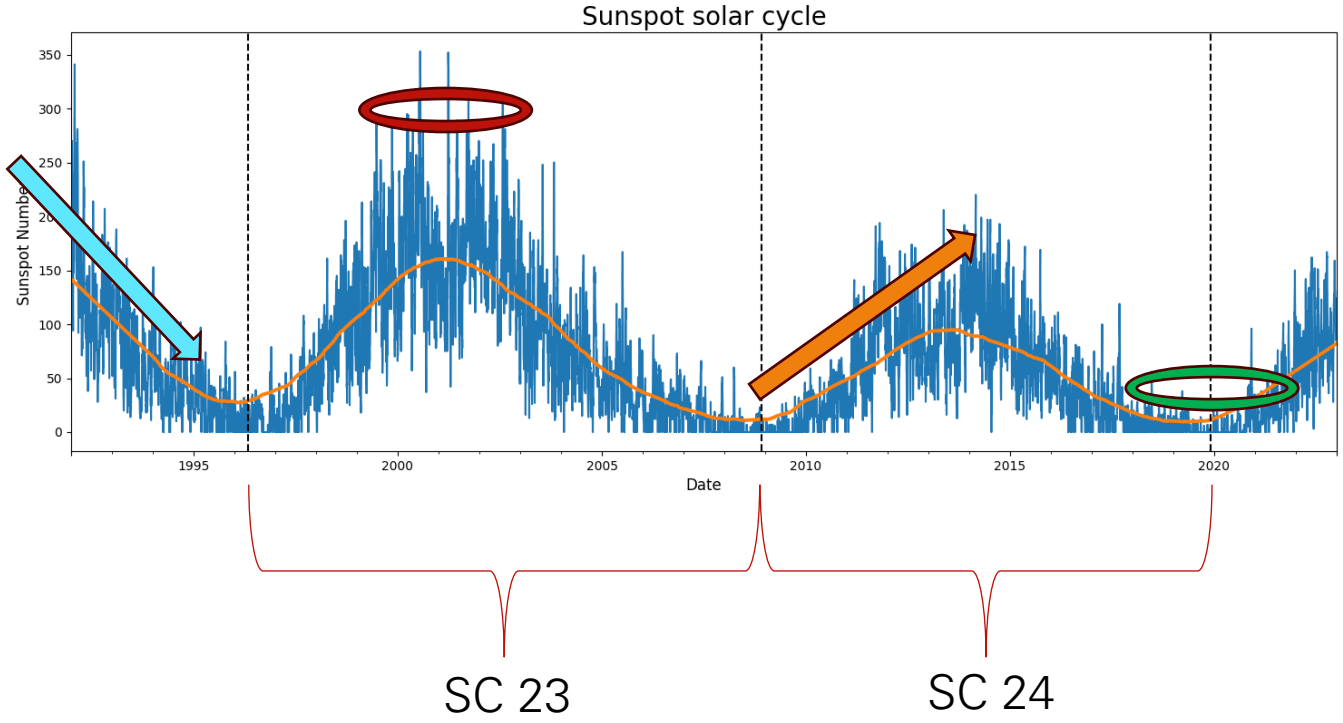


- The solar radio flux at 10.7 cm (2800 MHz) serves as an excellent indicator of solar activity. Referred to as the F10.7 index, it stands as one of the most enduring records of solar activity.

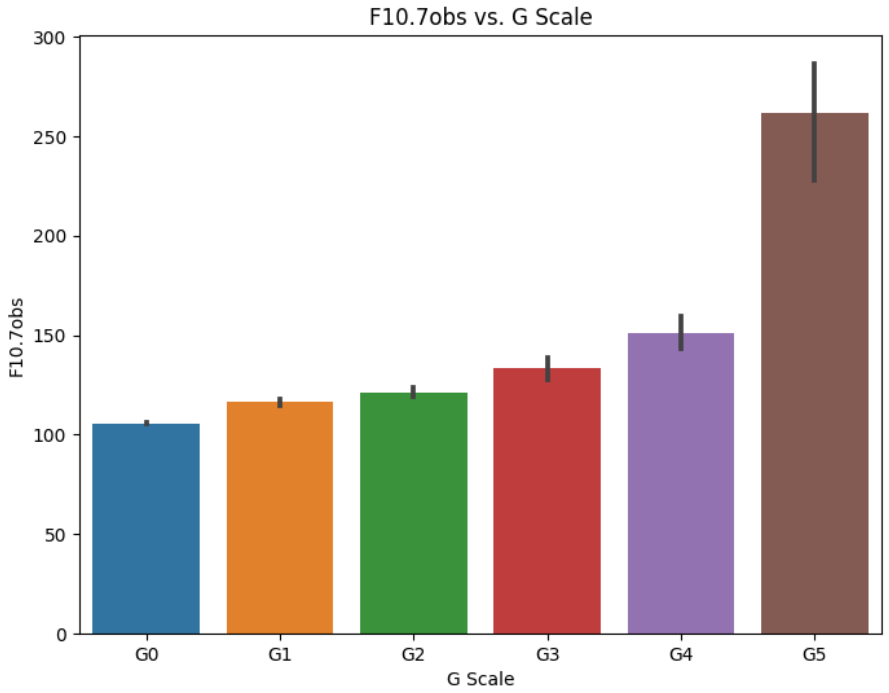
https://www.swpc.noaa.gov/phenomena/f107-cm-radio-emissions

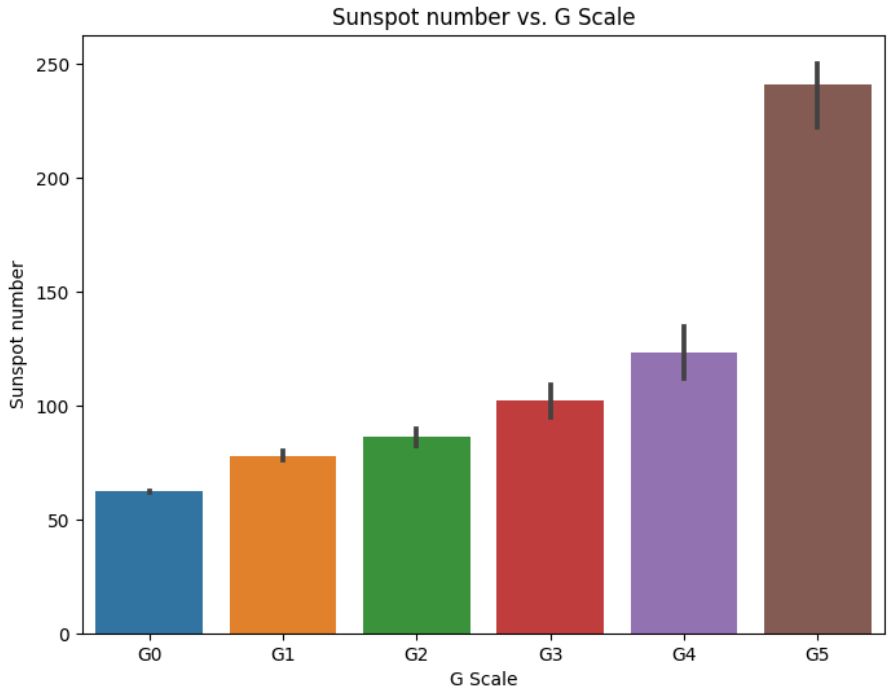
Categorizing data according to the phase of the Solar Cycle (SC)





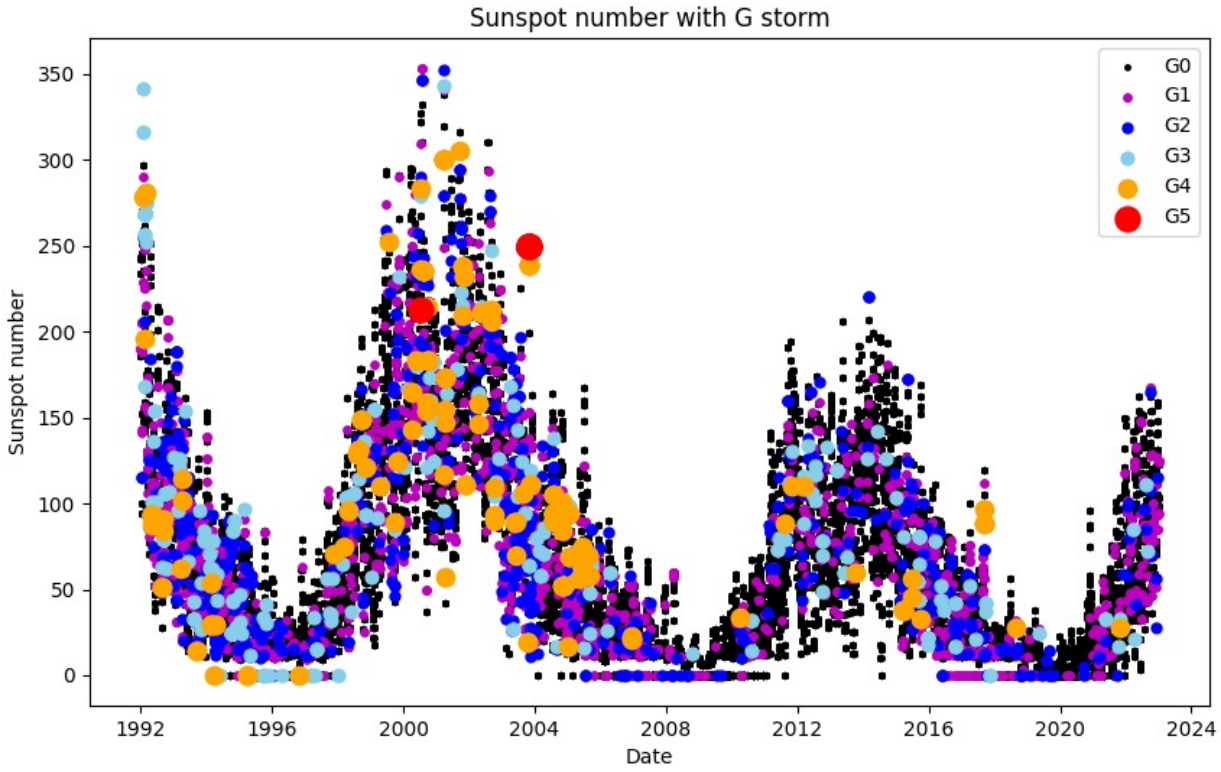
Storm activity with sunspots and F10.7cm (1992.01 – 2022.12)





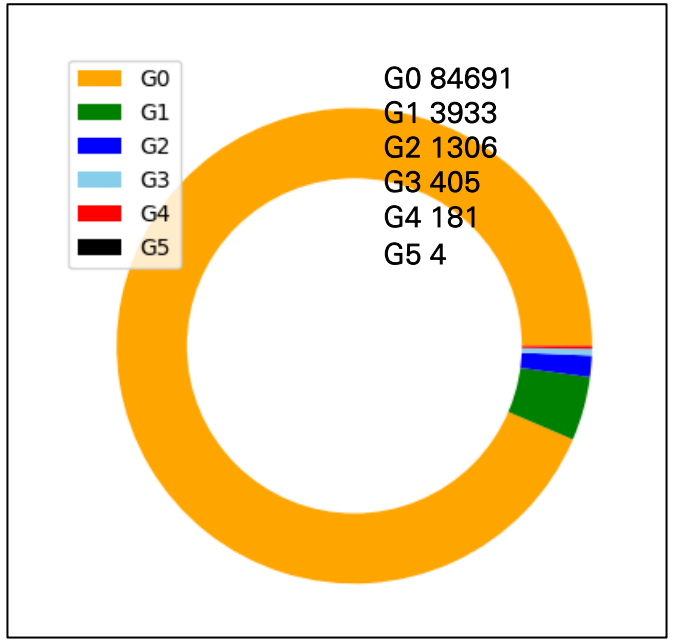
Identifying the geomagnetic activity with solar cycle (1992.01 – 2022.12)

G-scale	Kp	Auroral activity
G0	4 and lower	Below storm level
G1	5	Minor storm
G2	6	Moderate storm
G3	7	Strong storm
G4	8	Severe storm
G5	9	Extreme storm



Imbalanced classes

- Few of minority classes for a model to learn and make decision
- One approach to address this issue is oversampling the examples within the minority class
- SMOTE (Synthetic Minority Oversampling Technique) synthesizes new examples from the minority class



1992.01.01 – 2022.12.31

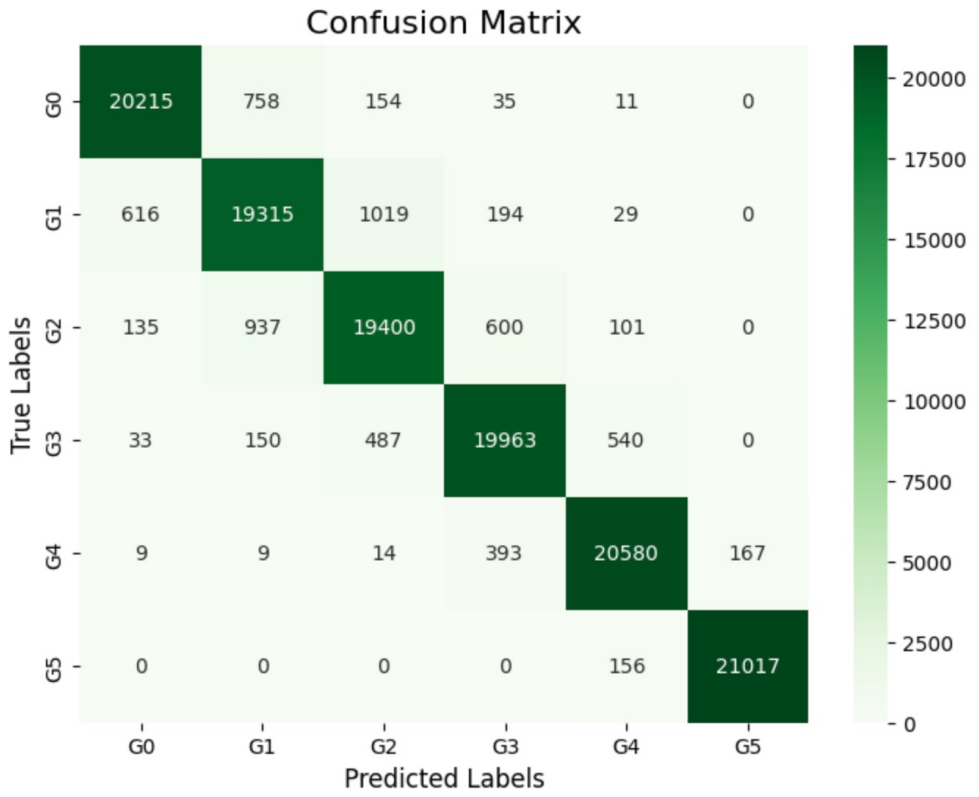
Model development

- Developed into a classification problem
- Grid searching to overcome overfitting and increasing the accuracy of the model
- Identified the RandomForest Classification model as the best model

Model	Score on train	Score on test
RandomForestc	0.982	0.951
DecisionTreec	0.968	0.935

```
 #   Column                Non-Null Count  Dtype
---  ------                --------------  -----
 0   Wind_Category_extreme 90520 non-null  object
 1   Wind_Category_fast    90520 non-null  object
 2   Wind_Category_slow    90520 non-null  object
 3   Np_Category_extreme   90520 non-null  object
 4   Np_Category_high      90520 non-null  object
 5   Np_Category_low       90520 non-null  object
 6   B_Category_extreme    90520 non-null  object
 7   B_Category_high       90520 non-null  object
 8   B_Category_low        90520 non-null  object
 9   T_Category_extreme    90520 non-null  object
 10  T_Category_high       90520 non-null  object
 11  T_Category_low        90520 non-null  object
 12  phase_declining       90520 non-null  object
 13  phase_rising          90520 non-null  object
 14  phase_solar_max       90520 non-null  object
 15  phase_solar_min       90520 non-null  object
 16  SN                    90520 non-null  float64
 17  F10.7obs              90520 non-null  float64
 18  G                     90520 non-null  object
```

Model Evaluation

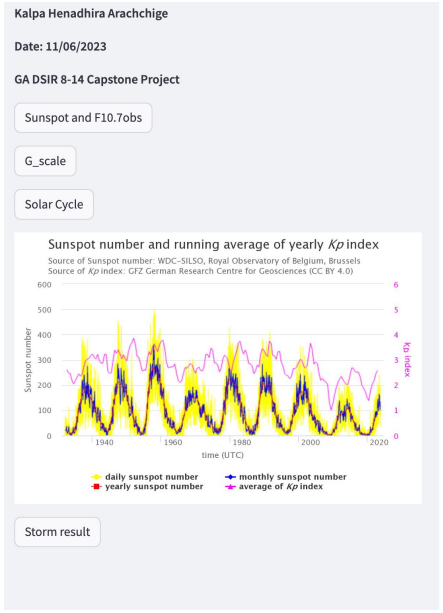


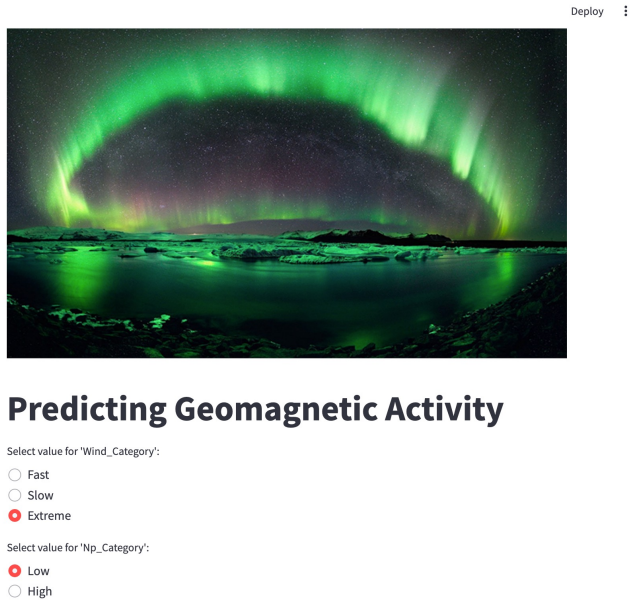
Classification Report:

	precision	recall	f1-score	support
G0	0.962	0.955	0.958	21173
G1	0.912	0.912	0.912	21173
G2	0.921	0.916	0.918	21173
G3	0.942	0.943	0.943	21173
G4	0.961	0.972	0.966	21172
G5	0.992	0.993	0.992	21173
accuracy			0.948	127037
macro avg	0.948	0.948	0.948	127037
weighted avg	0.948	0.948	0.948	127037

Conclusion & future work

- The connection between the solar indices (sunspots, solar radio flux) and solar wind properties in order to determine geomagnetic storms
- In the future, we plan to develop a model for data, including previous solar cycles 21 and 22





[Web application]