SOLAR WIND AND GEOMAGNETIC ACTIVITY PREDICTION

Using Sunspots and Solar Wind Measurements.

By Safa Mohamed Rafeeq

➢ Introduction

Geomagnetic activity, caused by interactions between the Earth's magnetosphere and solar wind, has significant impacts on satellite operations, communication systems, power grids, and navigation systems. Predicting this activity helps mitigate these effects, ensuring the stability and reliability of modern technology.

➢ The Role of Solar Wind

Solar wind is a stream of charged particles released from the upper atmosphere of the Sun, known as the corona. It varies in density, speed, and temperature, and these variations can influence geomagnetic activity. Key solar wind parameters that impact geomagnetic conditions include:

➢ Key Objectives

❑ **Integrate Solar Wind and Sunspot Data**:

1. Combine solar wind measurements from ACE and DSCOVR satellites with sunspot data.
2. Use this integrated dataset to capture the influence of solar activity on geomagnetic conditions.

❑ **Develop Predictive Models**:

1. Train models to predict the current (t0) and next (t+1) hourly Dst values based on solar wind and sunspot data.
2. Ensure models do not use historical Dst values as inputs, focusing solely on solar wind and sunspot predictors.

❑ **Calibrate Models to Solar Cycle**:

1. Incorporate the periodic variation in sunspot numbers to improve model accuracy.
2. Calibrate models to account for the 11-year solar cycle, enhancing their ability to predict geomagnetic activity during different phases of the cycle.

- **Benefits of Accurate Prediction**
- **Satellite Operations**: Predictive models help avoid damage to satellites by informing operators of potential geomagnetic storms.
- **Communication Systems**: Reliable predictions allow for adjustments in communication strategies to mitigate disruptions.
- **Power Grids**: Power companies can take preventive measures to protect the grid from geomagnetic-induced currents.
- **Navigation Systems**: Accurate forecasts help maintain the accuracy and reliability of navigation systems, crucial for aviation and maritime operations.
- By achieving these objectives, this study aims to enhance our ability to predict and respond to geomagnetic activity, thereby reducing the adverse effects on critical infrastructure and technology.

➢ Data Sources

❑ **ACE (Advanced Composition Explorer)**

- **Description**: Launched in 1997, ACE is a NASA satellite positioned at the L1 Lagrange point, about 1.5 million kilometers from Earth. This position allows ACE to continuously monitor the solar wind and interplanetary magnetic field conditions.
- **Data Provided**:
  - **Solar Wind Parameters**: Measures the speed, density, and temperature of the solar wind.
  - **Magnetic Field Data**: Provides data on the interplanetary magnetic field (IMF).
  - **Energetic Particles**: Detects and analyzes high-energy particles from the Sun and interstellar space.

❑ **DSCOVR (Deep Space Climate Observatory)**

- **Description**: Launched in 2015 by NOAA, DSCOVR is also positioned at the L1 Lagrange point. Its primary mission is to monitor space weather and provide real-time data on solar wind conditions.
- **Data Provided**:
  - **Solar Wind Parameters**: Measures the speed, density, and temperature of the solar wind, similar to ACE.
  - **Magnetic Field Data**: Monitors the interplanetary magnetic field (IMF).
  - **Earth Observations**: Provides additional data on Earth's atmosphere and climate, though its primary focus is space weather monitoring.

- **Prediction Goal**
- **Objective**: The aim is to predict the current (t0) and next (t+1) Dst values without relying on historical Dst values. This involves using other relevant data and models to forecast geomagnetic activity without using past Dst measurements directly. This approach helps in providing real-time or near-future predictions based on other available indicators.

➢Methodology

❑ Data Preparation

1. **Data Cleaning**:
    1. **Handling Missing Values**: Fill in or interpolate missing data to ensure completeness.
    2. **Outlier Detection**: Identify and address outliers to avoid skewed results.
2. **Feature Engineering**:
    1. **Temporal Features**: Extract features such as hour of the day, day of the week, and month to capture temporal patterns.
    2. **Lag Features**: Create lag features that use past observations to help in predicting future values.
3. **Normalization/Scaling**:
    1. **Normalization**: Scale the Dst values and other features to a common range to improve model performance.

- **Splitting Data**:
    1. **Training and Testing Split**: Divide the data into training and testing sets, ensuring that the test set contains future data not seen by the model during training.
- **Indexing**:
    1. **Period and Time delta Indexing**: Ensure data is correctly indexed by date and time to maintain the temporal order and granularity.
- **Hybrid Approaches**:

- **Combining Models**: Use a combination of machine learning and statistical models to leverage the strengths of both approaches for more robust predictions.

- **Key Insights & Conclusion**
- **Summary of Key Findings:**

1. Noticeable Trends in DST Over Time:

1. By plotting the DST values over time, we observe variations in DST that might indicate certain patterns or trends.
2. For instance, in the train_a period, there could be fluctuations in DST values at specific time intervals, showing either an increase or decrease in activity or measurements.

2. Significant Differences Between Periods:

1. Comparing average DST values across different periods (e.g., train_a, train_b, etc.) can highlight significant differences.
2. If train_a has an average DST value of -5 and train_b has an average DST value of -2, it suggests that train_b generally has higher DST values than train_a.

- **Final Thoughts:**
- **Overall Trends:**
  - There might be overall upward or downward trends in DST values within each period.
  - These trends can provide insights into how certain factors affect the measurements over time.
- **Period Comparison:**
  - The significant differences between periods might be due to different conditions or external factors influencing each period.
  - Understanding these differences can help in determining what factors contribute to higher or lower DST values.

Thank you !