Predicting Geomagnetic Storms with the Use of AI Algorithms in Deep Space Exploration

By Manal Siddiqui

Abstract

This research creates a predictive model for geomagnetic storm forecasting using information from the National Oceanic and Atmospheric Administration's (NOAA) Deep Space Climate Observatory (DSCOVR) satellite. The satellite DSCOVR is currently used to monitor space weather as it measures the main factors of the solar wind and the local magnetic field. These measurements are critical for understanding space weather processes and their effects on Earth.

The predictive model was built using data on solar wind speed, density, temperature, and magnetic field intensity. When it comes to sequential data, Long Short-Term Memory (LSTM) models typically perform better than different neural network patterns. This is because LSTM models have memory cells, which improves their ability to identify long-term dependencies in the data and makes them ideal for use in weather prediction models. The model tries to predict significant shifts in the magnetic field and solar wind to provide early warnings of geomagnetic storms. Forecasting these storms is critical for minimizing any interruptions because they threaten communication networks, power grids, and satellite activities.

This paper advances space weather forecasting and shows a direct use of machine learning techniques in space science. The LSTM model's predictive capabilities allow us to better prepare for geomagnetic storm events, protecting technology such as satellites and advancing our knowledge on space weather.

Past Research Done

The research I did prior to this focused on adapting communication methods for deep space missions, addressing challenges like radiation and gravitational effects. I also explored modifying communication features and using solar radiation for better prediction techniques [1], [7]. I also looked into TCP Congestion Control and Deep Reinforcement Learning to improve data transmission strategies [3] and finally my research allowed me to understand the importance of protocols like CCDS File Delivery Protocols (CFDP) for reliable file transfers in deep space [5], [6].

Recent studies exploring the integration of AI and ML within satellite communications systems [2] and [4] looked at the challenges and opportunities of onboard processing and AI techniques in forging the next generation of massive satellite networks. They talk about the application of AI to mitigate rapid elevation angle changes and tweaking channel variation predictions, stressing how important Neural Networks (NN) are for the future of space exploration forecasting.

After doing this research, my goal was to create a demonstrative model that would advance satellite communication predictions. This work led me to research the main causes of satellite communication disruptions in deep space, particularly solar winds.

The paper, ‘Space weather impacts on satellites and forecasting the Earth's electron radiation belts with SPACECAST’, talks about how geomagnetic storms and solar activity can lead to harsh and varrying radiation environments, which are especially critical for newly launched satellites that have not yet experienced conditions associated with a solar maximum. [8]. Here are some reasons listed in the paper for how exactly geomagnetic storms affect satellites:

- Solar Extreme Ultraviolet and Joule heating during geomagnetic storms can expand the Earth’s atmosphere, increasing drag on satellites in low Earth orbit (LEO), leading to potential orbital decay and uncontrolled reentry. [8]
- Satellites in higher orbits are particularly vulnerable to high-energy charged particles from cosmic rays, solar energetic particles (SEPs), and the Earth's radiation belts.
- Charged particles, especially from SEP events, can cause displacement damage in solar arrays, reducing efficiency and leading to power loss. They can also contribute to the total ionizing dose that degrades the efficiency of solar array coatings and other surface materials.
- These particles can penetrate electronic components and cause various issues, such as:
    - Single event effects like bit flips or data corruption.
    - Ionization in insulating layers results in leakage currents.
    - Increased noise and power consumption potentially lead to burnout and component failure. [8]

A full review of spacecraft anomalies in connection to space weather conditions is discussed within the publication "Space weather conditions and spacecraft anomalies in different orbits" by Iucci et al. The study covers 220 satellites and over 5700 anomalies between 1971 and 1994 and includes 49 Russian Kosmos satellites that have never been statistically examined before. [9]. The study connects space weather occurrences, like solar proton events and geomagnetic storms, to irregularities seen in satellite operations. The following summarizes how solar, wind, and geomagnetic factors affect satellite communications:

- Like the previous one, this paper mentions how strong solar proton events mainly affect satellites in high-altitude, near-polar orbits. These occurrences, which have the potential to produce higher energetic particle fluxes, are closely associated with anomalies in satellite operations, particularly in GPS-based navigation systems. Anomalies occur at a considerably higher rate during such events [9].
- Delayed Effects: Anomalies related to geomagnetic storms respond slowly, with their highest rates happening several days after the storm begins. Possible explanations for this delay include the residual effects of charged particle fluxes on satellite subsystems [9].

The following section will show an example of a Geo-Magnetic Storm Prediction model.

Code Explanation

The following section will cover the database I used for my prediction model, the pre-processing steps to prepare the data for training my model, and finally, the model itself. The database contains numerous data points, but I primarily focused on those providing the most insight into geomagnetic storms. The model is a LSTM NN, which uses several input features and predicts the designated label—detecting the level of geomagnetic storm based on other features.

Database Used

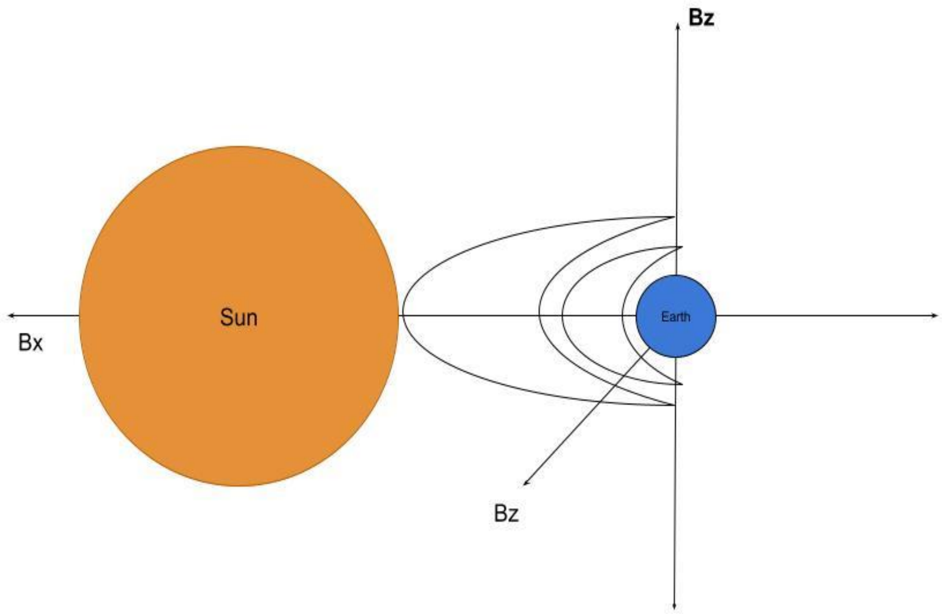
The database comes from the National Oceanic and Atmospheric Administration's (NOAA) state-of-the-art Deep Space Climate Observatory (DSCOVR) satellite. The satellite is currently located at the first Lagrange point (L1), which allows the satellite to monitor the interactions between the Earth and the Sun continuously. This is a strategic placement that takes advantage of L1's stable gravitational equilibrium to provide a clear view of solar phenomena and the space weather they cause [10].

The DSCOVR's work comes from mainly the PlasMag suite, which is an advanced equipment designed to analyze space weather details through careful observation. The fluxgate magnetometer (MAG) within the set measures the subtle variations of the local magnetic field in space. This data is important because it provides information on changes in the magnetic field that may disrupt the Earth's magnetosphere and impact power grids and satellite communications [10].

In addition to the MAG, the Faraday Cup (FC) provides information about the characteristics of the solar wind, including its temperature, density, and speed which were all used as input features [10]. Strongly charged particle streams from the Sun's atmosphere, known as the solar wind, interact with Earth's magnetic field to create the conditions for a wide range of space weather occurrences [8].

Pre-Processing

To pre-process the data, I used a NOAA database consisting of deep space satellite data. This includes key features such as the solar wind speed, density, temperature, and the Bz median value, which are for modeling geomagnetic storm predictions. Understanding variations in Bz measured in nanoteslas (nT) is necessary for predicting space weather. Monitoring Bz allows for important information about the possible effects of solar wind disruptions on satellites [11]. Bz, which is perpendicular to the ecliptic plane, is an important controller of geomagnetic activity. Given its direct contact with Earth's magnetosphere, Bz, which stands for the north-south component of the interplanetary magnetic field (IMF) [11].



[11]

The raw data is converted into a pandas DataFrame, which allows easy querying and normalization processes, narrowing down the data manipulation tasks. From the initial dataset, unnecessary columns, such as various statistical measures of the magnetic field components and solar wind properties not needed for this analysis were dropped to simplify the model inputs.

The labels were organized into four categories based on the median value of the Bz component of the IMF, which is a strong indicator of geomagnetic storm potential. These categories were designed to represent different intensities of geomagnetic storms [13] :

1. severe storms where the Bz index < -50 nT
2. strong storms where the Bz index is between 50 nT ≤ Bz < -20 nT
3. moderate storms where the Bz index is between -20 nT ≤ Bz < -10 nT
4. no storm where the Bz ≥ -10 nT

The MinMaxScaler is used to normalize the remaining feature set, which scales each feature to a given range, usually between zero and one, to make sure that all input features contribute equally to the model's training process. The scaled features and categorized labels were then ready for use in training the predictive model.

Neural Network Model

The proposed model takes account of the relationships that the solar wind's speed, density, and temperature have with the Earth's magnetic field to estimate future metrics of solar wind. It works with input data that is organized as (time_steps, features), where "features" represent the data points at each step and "time_steps" indicate the length of the sequence. This structure is designed to include important factors like elevation angle, attenuation of rain, and other noteworthy atmospheric conditions, guaranteeing a thorough approach to forecasting.

My neural network architecture includes two LSTM layers designed to capture both short-term and long-term dependencies in the data. To prevent overfitting—a common challenge with complex models—I integrate dropout layers. Within the network's final layer is a dense layer which also means all layers are fully connected, optimized for making continuous value predictions essential for the predictions. The optimizer used is 'Standard Gradient Decent' which helps in speeding up the learning process for the model. Finally, the loss function used was mean squared error which is commonly used within regression models.

Results

Validation accuracy evaluates the model's capacity to make predictions to new, unused data, whereas training accuracy evaluates the model's performance on the data it was trained on. A decrease or drop in validation accuracy indicates possible problems such as overfitting, whereas a steady increase in training accuracy signifies good learning.

Validation Accuracy is calculated by:

$$Validation\ Accuracy = \frac{Number\ of\ Correct\ \mathrm{Pr}\,e dictions\ on\ Validation\ Data}{Total\ Number\ of\ Validation\ Data\ Points}\ x\ 100$$

Training Accuracy is calculated by:

$$Training\ Accuracy = \frac{Number\ of\ Correct\ \mathrm{Pr}\,e dictions\ on\ Training\ Data}{Total\ Number\ of\ Training\ Data\ Points}\ x\ 100$$

A model is said to be overfit when it learns the training set too thoroughly, including noise and anomalies that fail to adapt to new data. Regularization methods that penalize the model for complexity, such L1 or L2 regularization, are used to correct overfitting. This also makes the model focus on key features rather than accounting for the anomalies or noise [14].

$$Regularization\ Term = \lambda \times \left(\sum_{i=1}^{n} |\ w_i\ |\ \ or \sum_{i=1}^{n} w_i^2 \right)$$

In L1 regularization, also known as Lasso regression, the regularization term penalizes the sum of the absolute values of the regression model's coefficients, favoring simplicity and feature selection by pushing fewer important coefficients towards zero. L2 regularization, or Ridge regression, penalizes the sum of squares of the coefficients, causing them to decrease more smoothly without being forced to zero. This enables L2 regularization to produce smoother models and effectively handle outliers, while L1 regularization excels at feature selection. The choice between L1 and L2 regularization can be used in the chance that overfitting is not addressed [14].

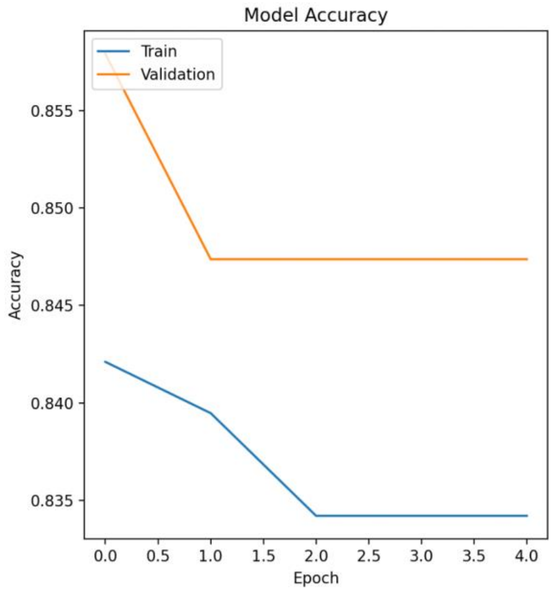


Figure 1: Model Accuracy

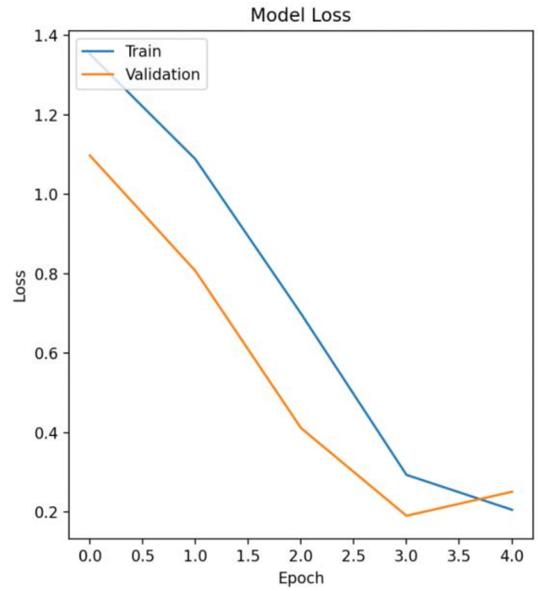


Figure 2: Model Loss

As seen in Figure 1, the accuracy for 4 Epochs ranges from 83% to 84%, with validation accuracy being much higher. This training accuracy measures the model's ability to forecast geomagnetic storms based on the training dataset, indicating successful learning when there is an increasing trend. Validation accuracy, representing the model's performance on a different dataset not seen during training, is essential for assessing its generalization abilities. A high and consistent validation accuracy indicates effective performance on untrained data. However, if training accuracy increases while validation accuracy declines, it is likely suggesting that overfitting is occurring, where the model learns noise and outliers that do not reflect true geomagnetic patterns.

Training loss, shown in Figure 2, decreases as a sign that the model's predictions are becoming more accurate. An increase in validation loss, despite a decreasing training loss, may once again be indicating overfitting. However overall, the loss is very small before the signs of overfitting begin to show, meaning that the model has perforemed quite well.

Overfitting could be addressed by adding more data, adding more features to increase input complexity, or implementing stronger L1 or L2 regularization to simplify the model and focus on main patterns of the data. A few impacts of overfitting could negatively result in the model’s ability to predict actual geomagnetic storms, which is critical if the predictions I used for important decision-making.

Conclusion & Future Work

Using LSTM neural networks and other AI capabilities, this research has presented a method for forecasting geomagnetic storms. I built a model that can use solar wind parameters and predict geomagnetic activity using DSCOVR satellite data. The results of this research have significant impact since it opens a possibility to accurate space weather forecasts, which is essential for maintaining satellites.

The training and validation results showed that my model could forecast geomagnetic storm episodes with promising accuracy. It also clarified where work still needs to be done, especially to prevent overfitting and improve the model's capacity to generalize to new data. In the future, the model will be improved by incorporating new variables and data sources, evaluating at different neural network architectures, and enhancing its capacity to produce more precise geomagnetic activity projections [12].

The goal of this research's advancement is to implement the real-time predictive model for space weather monitoring. That would include creating a fully automated end-to-end system that forecasts geomagnetic storms and offers information to communication network providers, power grid operators, and satellite operators [12].

References

Works Cited

[1] Brown, A. (2018). New Developments in Technology. In M. Johnson (Ed.), Advances in Science. https://www.intechopen.com/chapters/70180

[2] Ferreira, P. V. R., Paffenroth, R., Wyglinski, A. M., Hackett, T. M., Bilen, S. G., Reinhart, R. C., & Mortensen, D. J. (2019). Reinforcement learning for satellite communications: From LEO to deep space operations. *IEEE Communications Magazine*, *57*(5), 70-75.

[3] Ha, T., Masood, A., Na, W., & Cho, S. (2023). Intelligent Multi-Path TCP Congestion Control for video streaming in Internet of Deep Space Things communication. *ICT Express*.

[4] Ortiz, F., Monzon Baeza, V., Garces-Socarras, L. M., Vásquez-Peralvo, J. A., Gonzalez, J. L., Fontanesi, G., ... & Chatzinotas, S. (2023). Onboard processing in satellite communications using ai accelerators. *Aerospace*, *10*(2), 101.

[5] Pavale, S., Unnikrishnan, E., & Laksminarasimhan, P. (2010, December). Design, implementation and performance evaluation of CCSDS CFDP protocol. In *2010 IEEE International Conference on Computational Intelligence and Computing Research* (pp. 1-4). IEEE.

[6] Sanders, F. A., Jones, G., & Levesque, M. (2007, March). Transfer of files between the deep impact spacecrafts and the ground data system using CFDP: a case study. In *2007 IEEE Aerospace Conference* (pp. 1-5). IEEE.

[7] Smith, J. (2020). Advances in Physics Research. Journal of Physics, 2489(1), 012045. https://iopscience.iop.org/article/10.1088/1742-6596/2489/1/012045

[8] R. B. Horne *et al.*, "Space weather impacts on satellites and forecasting the Earth's electron radiation belts with SPACECAST," *Space Weather*, vol. 11, no. 4, pp. 169–186, Apr. 2013, doi: https://doi.org/10.1002/swe.20023.

[9] N. Iucci *et al.*, "Space weather conditions and spacecraft anomalies in different orbits," *Space Weather*, vol. 3, no. 1, p. n/a-n/a, Jan. 2005, doi: https://doi.org/10.1029/2003sw000056.

[10] SPACE WEATHER PREDICTION CENTER, "Real Time Solar Wind | NOAA / NWS Space Weather Prediction Center," *Noaa.gov*, 2019. https://www.swpc.noaa.gov/products/real-time-solar-wind

[11] Parsec vzw, "The Interplanetary Magnetic Field (IMF)," www.spaceweatherlive.com. Available: https://www.spaceweatherlive.com/en/help/the-interplanetary-magnetic-field-imf.html

[12] manalsiddiqui, "manalsiddiqui/honours-thesis-geomagnetic-prediction-model," GitHub, Apr. 16, 2024. Available: https://github.com/manalsiddiqui/honours-thesis-geomagnetic-prediction-model. [Accessed: Apr. 17, 2024]

[13] "The Kp-index," www.spaceweatherlive.com. Available: https://www.spaceweatherlive.com/en/help/the-kp-index.html

[14] F. vl, "Interpreting Training/Validation Accuracy and Loss," *Medium*, Nov. 08, 2023. Available: https://medium.com/@frederik.vl/interpreting-training-validation-accuracy-and-loss-cf16f0d5329f