Probabilistic Predictions of Geomagnetic Field Perturbations and Using Explainability Methods to Uncover Drivers of the Localization Effect

Mike Coughlan [1,2] Amy Keesee [1,2] Victor Pinto [3] Raman Mukundan [1,2] Jose Paulo Marchezi [2]
Jeremiah Johnson [4] Mayowa Adewuyi [1,2] Joel Tibbetts [1,2] Hyunju Connor [5] Don Hampton [6]

[1]University of New Hampshire - Department of Physics & Astronomy [2]University of New Hampshire - Institute for Earth, Oceans & Space
[3]Departamento de Fisica, Universidad de Santiago de Chile [4]University of New Hampshire - Department of Electrical and Computer Engineering
[5]NASA Goddard Space Flight Center [6]University of Alaska Fairbanks - Geophysical Institute

MAGICIAN
Machine Learning Algorithms for Geomagnetically Induced Currents in Alaska and New Hampshire

Introduction

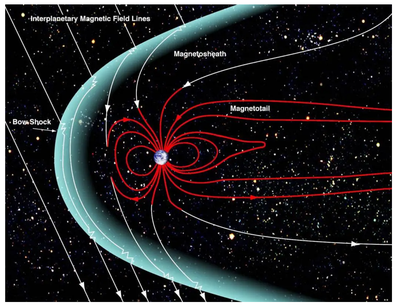

Figure 1. The Earth's Magnetosphere.

Rapid changes in the ground magnetic field **(dB/dt)** can lead to **Geomagnetically Induced Currents (GICs)**. dB/dt often exhibits large, highly **localized** spikes, showing substantial differences in measurements within small spatial areas which can be attributed to drivers originating in the ionosphere and magnetosphere [3]. **Region-to-Specific Difference (RSD)** [2] is a parameter that can be used to quantify this localization effect. Here we explore the effect of including **ion temperature maps** from the Two Wide-Angle Imaging Neutral-Atom Spectrometers (TWINS) into predictions of large dB/dt and RSD in a particular region. We then use **SHapley Additive exPlination (SHAP)** values to examine the driving features of these phenomena.

Method

- Two **Convolutional Neural Networks (CNN)** were trained, one forecasting extreme threshold crossings of dB/dt (dB/dt model) and one forecasting times of extreme localization (RSD model). Dimensionality reduction of the maps was done using MaxPooling.
- The model predicts a probability that the $(dB/dt)_{max}$ or RSD_{max} in the region will exceed the 99^{th} percentile (y^0).
- It outputs a mean (μ) and a standard deviation (σ) of a Gaussian distribution using the **Continuous Rank Probability Score (CRPS)**, enabling us to understand the model's uncertainty in it's predictions.

$$\mathrm{CRPS}(\mu,\sigma,y^0)=\sigma\left[\frac{y^0-\mu}{\sigma}\mathrm{erf}\left(\frac{y^0-\mu}{\sqrt{2}\sigma}\right)+\sqrt{\frac{2}{\pi}}\exp\left(\frac{-(y^0-\mu)^2}{2\sigma^2}\right)-\sqrt{\frac{1}{\pi}}\right]$$

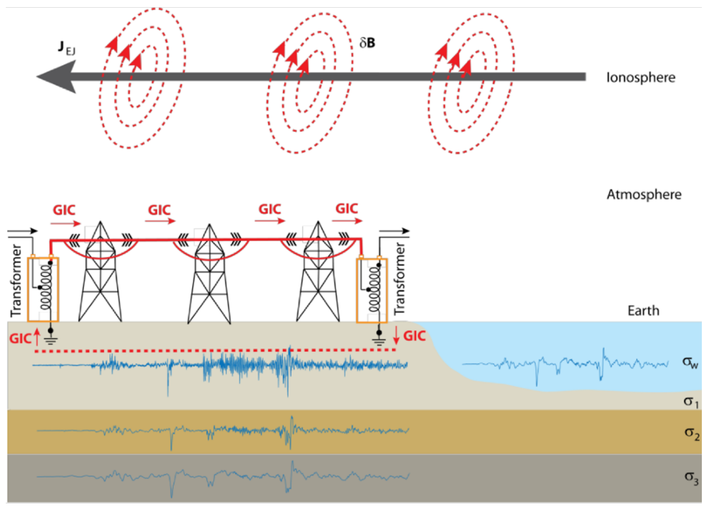

Figure 2. Demonstration of how ionospheric currents can cause GICs when interacting with long conducting infrastructure. Figure also shows the role played by the different conductivity's of ocean water and land mass.

Model Architecture

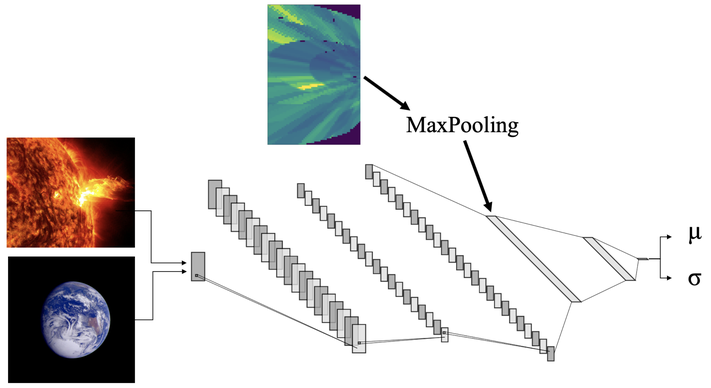

Figure 3. Layout of Convolutional Neural Network model design. Model utilizes two Convolutional layers processing the solar wind and magnetometer time series data. The ion temperature data is then concatenated after going through a maxpooling layer, before the data is sent through two fully connected layers. Based off of model used in [1]

Model Predictions

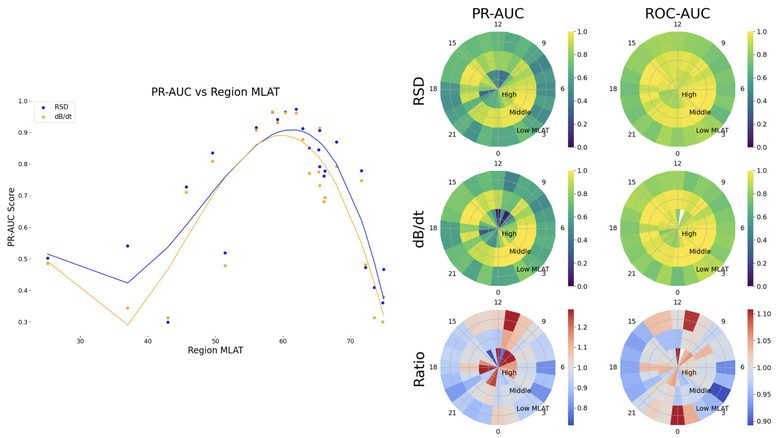

Figure 4. Model results. Left hand figure shows the area under the precision-recall curve (PR-AUC) for both models as a function of MLAT for each region. The figure on the right shows both the PR-AUC (left column) as well as the area under the receiver-operator characteristic curve (ROC-AUC) (right column) as a function of both Magnetic Latitude (MLAT) (radial) and Magnetic Local Time (MLT) (theta).The top row are the results for the RSD model, the middle the dB/dt model, and the bottom row the ratio of the two with positive values indicating the RSD achieved a higher score.

Conclusions

- Both versions of the model show skill in their forecasting ability.
- Models show best skill in the 55-65 MLAT range with degraded performance at low and high magnetic latitudes.
- Ion temperature in the magnetotail only has significant influence at higher temperatures.
- A feature's overall influence on the model varies as a function of Magnetic Latitude and Magnetic Local Time.
- Further investigation is necessary to determine if the drivers of these two phenomena differ significantly.

RSD Model DeepSHAP Contributions

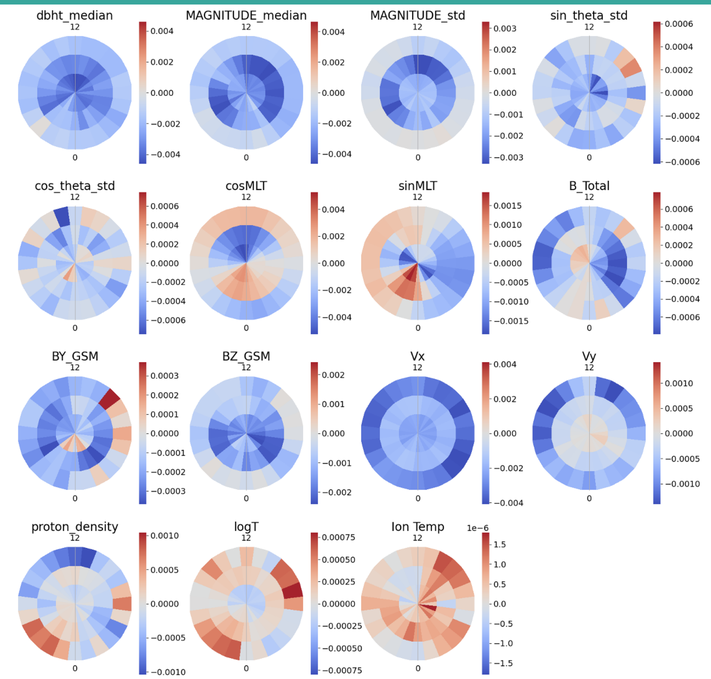

Figure 5. Averaged feature importance for each input parameter across all testing data as calculated by the DeepSHAP method for the RSD model. Results are plotted as a function of binned MLAT in the radial direction, with the center being the highest MLAT and the outer ring being the Lowest MLAT, and 1-hour binned MLT. Positive values (red) indicate an increase in model output and negative values (blue) indicate a decrease.

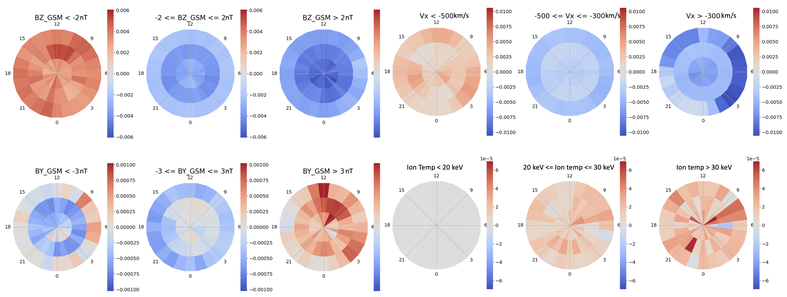

Figure 6. SHAP importance broken down by input parameter value for the Z (top left) and Y (bottom left) component of the Interplanetary Magnetic Field (IMF). the X component of solar wind velocity (top right) and the magnetotail ion temperature (bottom right).

References/Acknowledgements

[1] M. Coughlan, A. Keesee, V. Pinto, R. Mukundan, J. P. Marchezi, J. Johnson, H. Connor, and D. Hampton. Probabilistic Forecasting of Ground Magnetic Perturbation Spikes at Mid-Latitude Stations. *Space Weather*, 21(6), June 2023.

[2] A. P. Dimmock, L. Rosenqvist, D. T. Welling, A. Viljanen, I. Honkonen, R. J. Boynton, and E. Yordanova. On the Regional Variability of d *B* /d *t* and Its Significance to GIC. *Space Weather*, 18(8), Aug. 2020.

[3] C. M. Ngwira, D. Sibeck, M. V. D. Silveira, M. Georgiou, J. M. Weygand, Y. Nishimura, and D. Hampton. A Study of Intense Local d *B* /d *t* Variations During Two Geomagnetic Storms. *Space Weather*, 16(6):676–693, June 2018.

We thank all members of the MAGICIAN team at UNH and UAF that participated in the discussions leading to this work. We also thank the OMNIWeb and SuperMAG teams for providing the data.