Towards an Interpretable Machine Learning Model of Localized Geomagnetic Disturbances in Terms of Solar Wind Drivers

SM43B-2554

Raman Mukundan[1,2], Amy Keesee[1,2], José Marchezi[3], Mike Coughlan[4], Victor Pinto[5]

AGU Fall Meeting 2025

[1]Department of Physics and Astronomy, University of New Hampshire
[2]Institute for the Study of Earth, Oceans, and Space, University of New Hampshire
[3]Space Weather Division, National Institute for Space Research, Brazil
[4]NASA GSFC, Greenbelt, Maryland
[5]Departamento de Física, Universidad de Santiago de Chile

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

Introduction

- The solar wind's interaction with the Earth's magnetic field can cause Geomagnetically Induced Currents (GICs), which are hazardous to power and communications infrastructure.
- Changes in the horizontal component of the ground magnetic field ($d\mathbf{B}_H$) or its time derivative are used as a proxy for GIC measurements.
- Our goal is to obtain a $d\mathbf{B}_H$ forecast at every point in a large region using the Spherical Elementary Current Systems interpolation technique.
- We used a neural network to model the physics that connects the solar wind to localized $d\mathbf{B}_H$.
- We used Integrated Gradients to explain forecasts of $d\mathbf{B}_H$ and its regional roughness R_{ms} in terms of their solar wind drivers.

Conclusions and Key Insights

- Preliminary results include:
  - v_x tends to not become important (except for pulses) until several hours into storm main phase
  - Reconnection E-field ($||\mathbf{v} \times \mathbf{B}_z||$) is influential at long lookback times (>20 min. ago).
  - B_z strongly promotes localized $d\mathbf{B}_H$ (large R_{ms}). across all lookback times.
- So far only solar wind inputs are used. The next iteration will include ionospheric data from DMSP, SuperDARN, and AMPERE.
- A differentiable contour-finding algorithm is desired so that contours' properties can be included in the Integrated Gradients analysis.

Acknowledgements and References

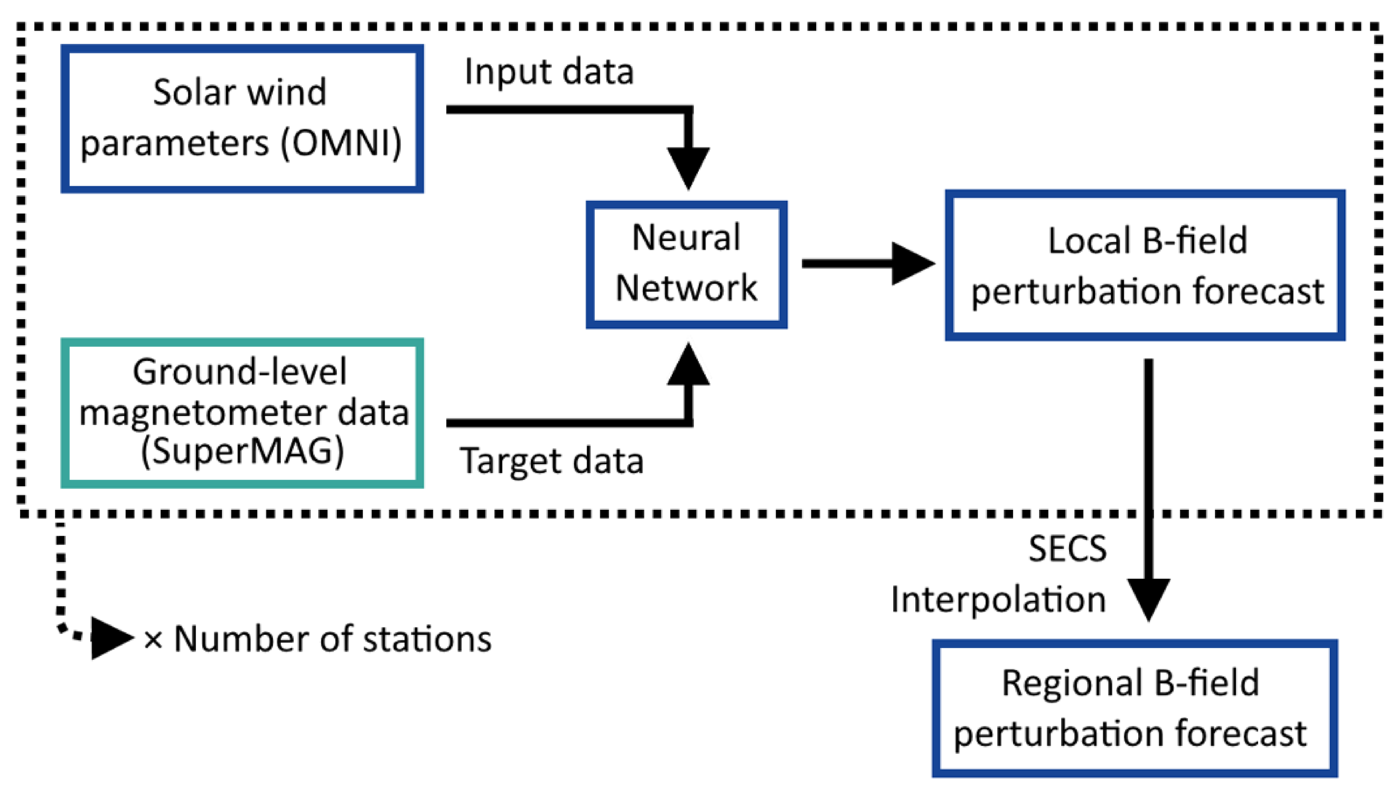
This work is supported by NSF EPSCoR T-2 Award 1920965 & NSF Award 2331527

- OMNIWeb data from omniweb.gsfc.nasa.gov
- SuperMAG data obtained from supermag.jhuapl.edu

Website

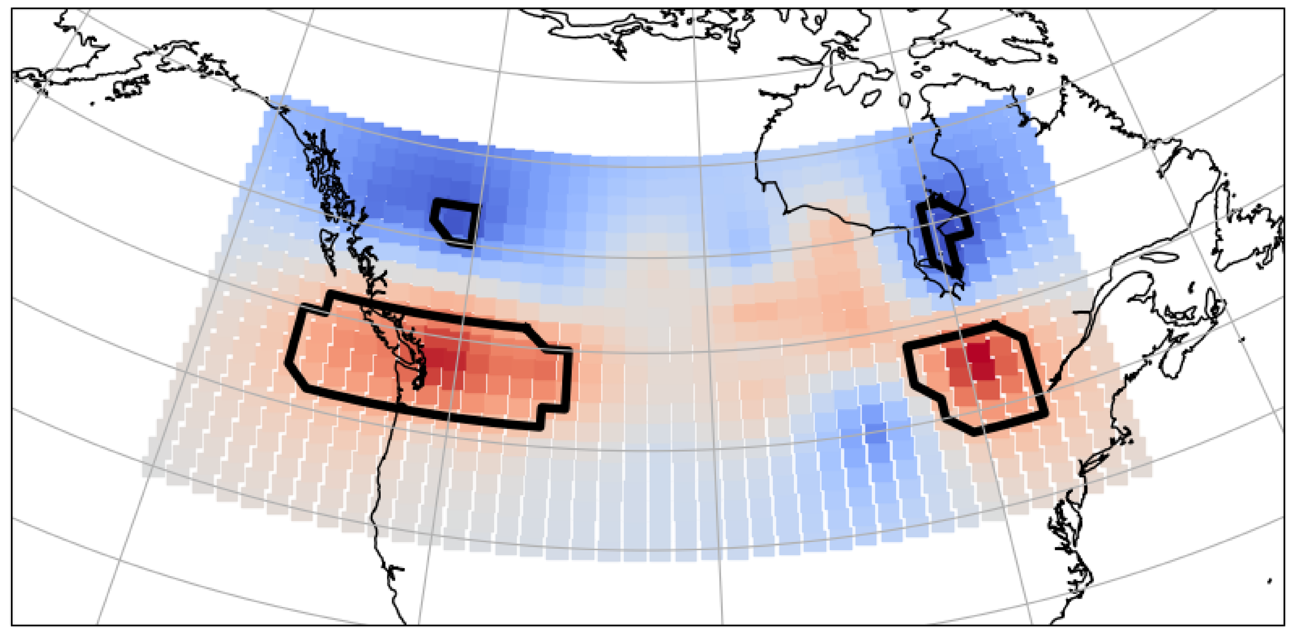
Read our related paper!

Magnetic Multi-Model Prediction



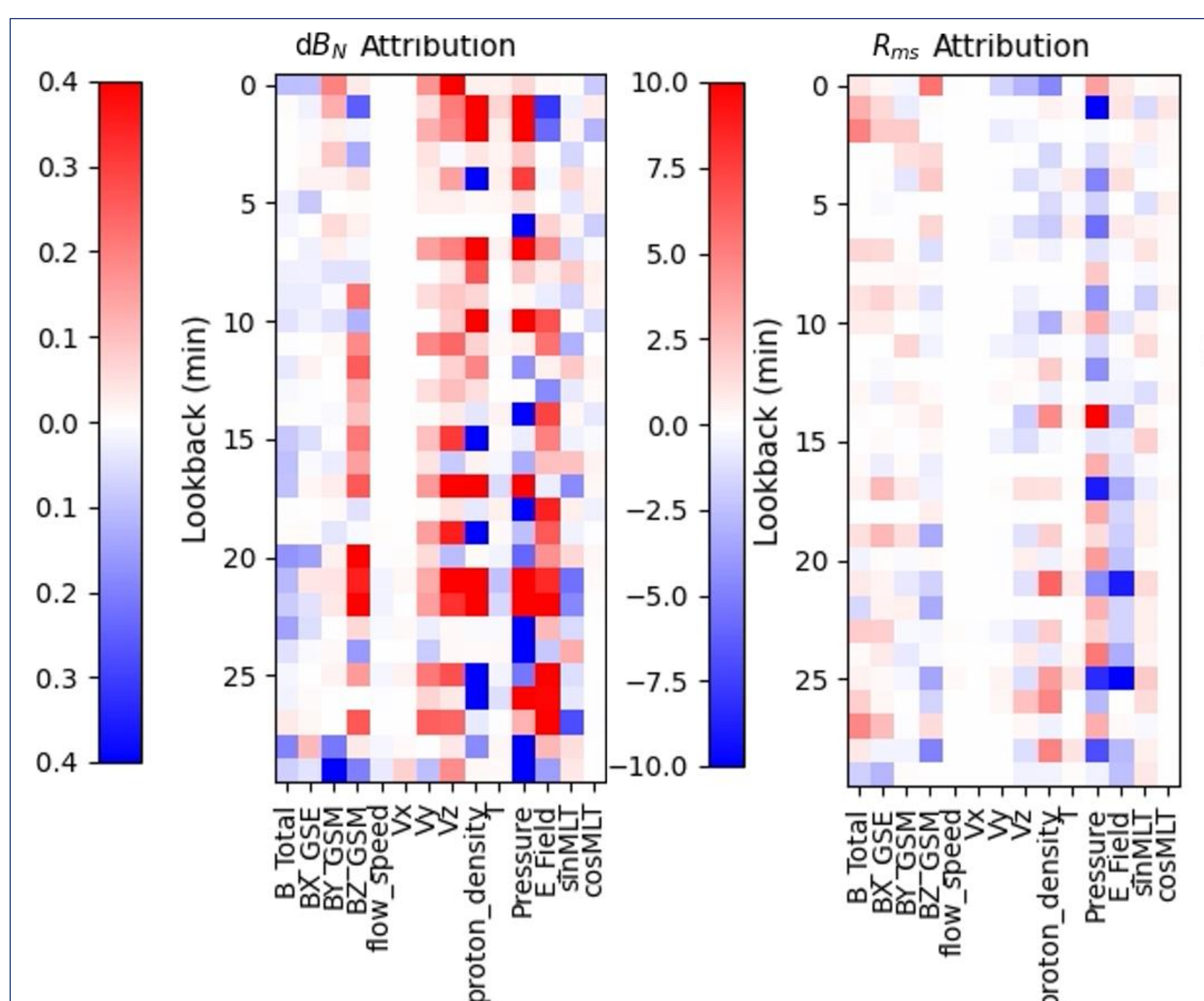
- Each of the 24 magnetometer stations has an associated convolutional neural network (CNN) trained to predict only that station's measurements using 30-minute time history (similar to Pinto+ 2022; Coughlan+, 2023). Spatial interpolation is performed afterward.
- **Theoretical Advantage:** Each model has a relatively simple task, so each model needs to learn less information than if a single larger CNN were used.

Spatial Interpolation as a Non-Learnable Layer

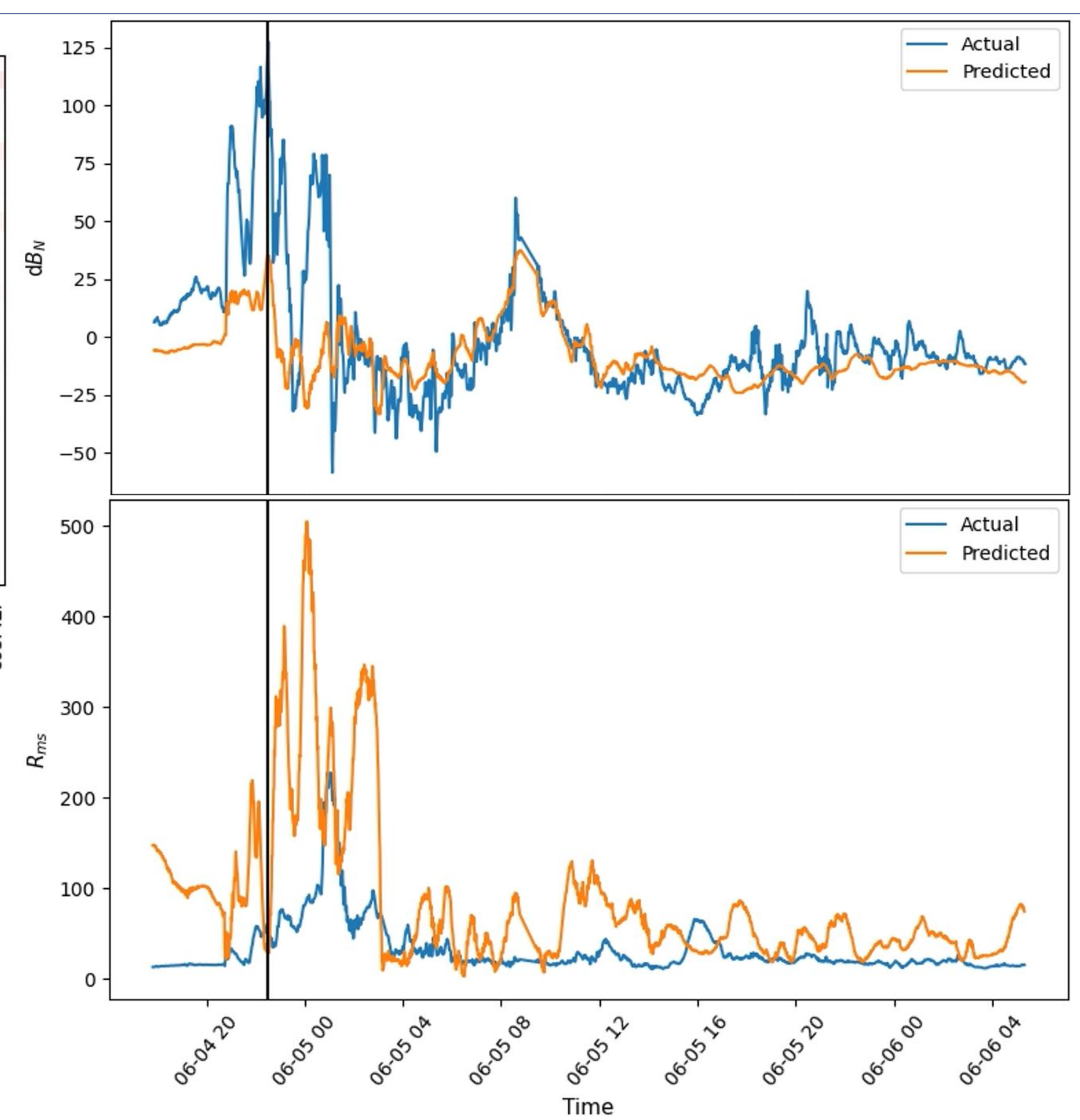


- The magnetometer forecast vector (CNN outputs) is used to fit a grid of Spherical Elementary Current Systems, representing ionospheric currents, to derive 1-minute cadence $d\mathbf{B}_H$ values across the region during storm time in Solar Cycle 24 (2009-2019).
- Interpolations are validated by leaving each station out 1 by 1 and at comparing at that station's site.
- The entire process is differentiable (though not learnable), which enables attribution methods such as Integrated Gradients.
- An automated algorithm identifies localized geomagnetic disturbances (**black contours**) in the prediction.

Which Inputs Influence the Model's Prediction at any Given Moment?



Attribution map movies



Integrated Gradients in 3 Easy Steps

1. For each input tensor, create a series of artificial inputs x_i' which lie along a path between that input x and the average solar wind conditions x_0'
2. Calculate $\partial F(x_i')/\partial x_i'$ for every i and every feature
3. Take the mean of those gradients

Plotted are predictions for $d\mathbf{B}_N$ and R_{ms}, the root mean square deviation of each $d\mathbf{B}_N$ map from its mean, μ.

$$R_{ms} = \left(\frac{1}{IJ} \sum_i^I \sum_j^J (x_{ij} - \mu)^2 \right)^{1/2}$$

R_{ms} is a measure of the roughness of a $d\mathbf{B}_N$ map like the one above.

Red pixels mean those features tend to increase the value of $d\mathbf{B}_N$ or R_{ms}; blue pixels mean those features tend to decrease (make more negative) that value.