57: Comparison of Time Series Techniques to Model Connections Between Solar Wind Input and Geomagnetically Induced Currents

Amy Keesee[1a], Victor Pinto[1], Michael Coughlan[1], Connor Lennox[1], Md Shaad Mahmud[1], Hyunju Connor[2]
[1]*University of New Hampshire, Durham, NH 03824*
[2]*University of Alaska Fairbanks, Fairbanks, AK 99775*
[a] Amy.Keesee@unh.edu

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

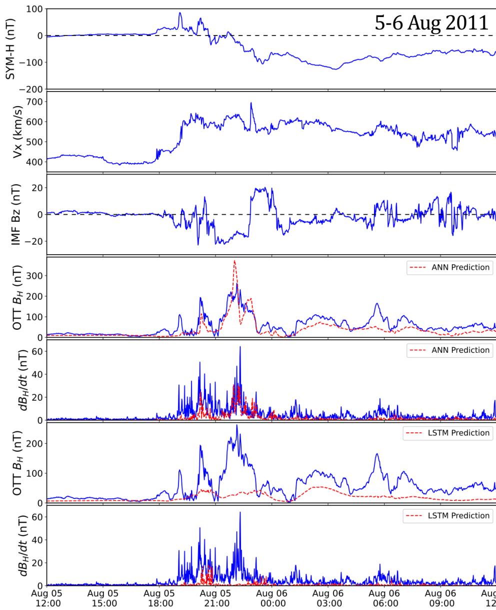
Abstract

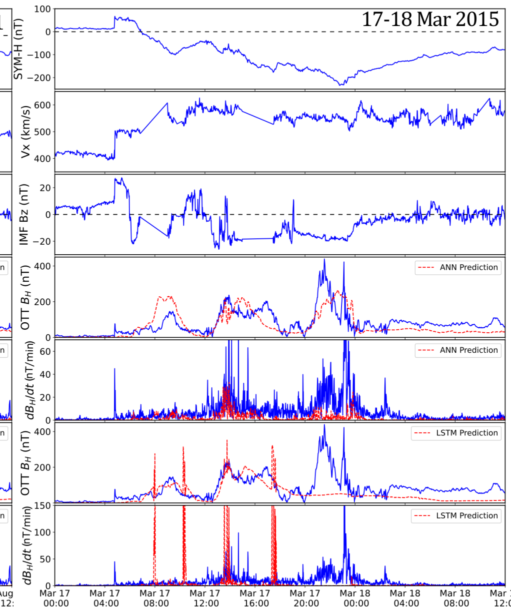
Geomagnetically induced currents (GIC) can drive power outages and damage power grid components while also affecting pipelines and train systems. Developing the ability to predict local GICs is important to protecting infrastructure and limiting the impact of geomagnetic storms on public safety and the economy. While GIC data is not readily available, variations in the magnetic field, dB/dt, measured by ground magnetometers can be used as a proxy for GICs. We are developing a set of neural networks to predict the east and north components of the magnetic field, B_E and B_N, from which the horizontal component, B_H, and its variation in time, dB_H/dt, are calculated. We apply two techniques for time series analysis to study the connection of solar wind and interplanetary magnetic field properties obtained from the OMNI dataset to the ground magnetic field perturbations. The analysis techniques include a feed-forward artificial neural network (ANN) and a long-short term memory (LSTM) neural network. Here we present a comparison of both models' performance when predicting the B_H component of the Ottawa (OTT) ground magnetometer for the year 2011 and 2015 and then when attempting to reconstruct the time series of B_H for two geomagnetic storms that occurred on 5 August 2011 and 17 March 2015.

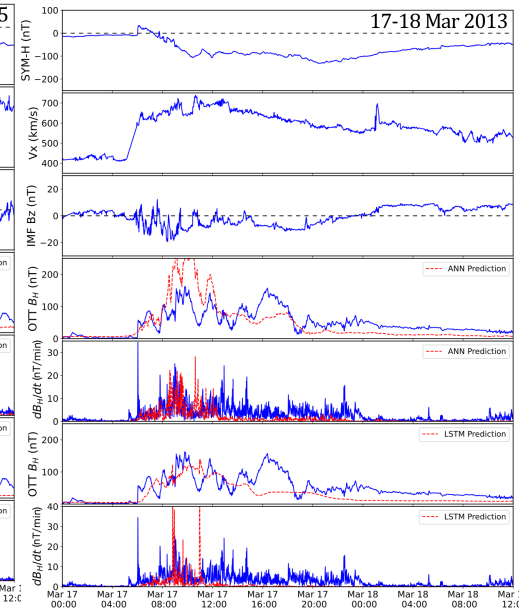
Methodology

- OMNI data and SuperMag data from the Ottawa (OTT) station from the years 1995-2010 were used. A full linear interpolation was done on the missing data points.
- Artificial Neural Network (ANN) and Long Short-Term Memory (LSTM) models were trained in Tensorflow-Keras.
- The input feature vector includes solar wind speed (V, Vx, Vy, Vz), IMF (B, Bx, By, Bz), proton density, dynamic pressure, temperature and solar wind electric field using a 1-min cadence for the first 12 preceding minutes (i.e., up to t − 12) plus 10-min averages over the previous 2 hours. The ground magnetometer sin(MLT) and cos(MLT) values have been included to ensure a cyclical dependence over the Earth's rotation and solar zenith angle as a proxy of both latitude and yearly seasonality.
- The output value is not included as an input feature in the LSTM for direct comparison to the ANN. We also want to determine the ability for a forecasting model, requiring a time delay for the model input.

Storm Results







Metrics

- Each storm interval is divided into 20-min windows, within which the maximum measured dB_H/dt is determined to see whether it crosses threshold values of 18, 42, 66, and 90 nT/min.
- The modeled dB_H/dt is calculated with two methods described by Pulkkinen et al, 2013 and Tóth et al., 2014, respectively:
  - $$\left(\frac{dB}{dt}\right)_H = \sqrt{\frac{dB_N}{dt}^2 + \frac{dB_E}{dt}^2}$$
  - $$\left(\frac{dB}{dt}\right)_H = \left(\frac{B_H}{248\,\text{nT}}\right)^{1.04} \frac{\text{nT}}{\text{s}}$$
- These values are used to determine the probability of detection (POD), probability of false detection (PFD), percentage correct (PC), and Heidke Skill Score (HSS).

		Pulkkinen				Tóth			
		18	42	66	90	18	42	66	90
2011 Storm	ANN POD	0.33	0.00	—	—	0.50	0.50	—	—
	LSTM POD	0.08	0.00	—	—	0.00	0.00	—	—
	ANN PFD	0.02	0.00	0.00	0.00	0.02	0.03	0.03	0.01
	LSTM PFD	0.08	0.00	—	—	0.00	0.00	—	—
	ANN PC	0.88	0.97	1.00	1.00	0.90	0.96	0.97	0.99
	LSTM PC	0.85	0.97	1.00	1.00	0.84	0.97	1.00	1.00
	ANN HSS	0.41	0.00	—	—	0.58	0.37	—	—
	LSTM HSS	0.13	0.00	—	—	0.00	0.00	—	—
2015 Storm	ANN POD	0.09	0.00	0.00	0.00	0.69	0.77	0.00	0.00
	LSTM POD	0.13	0.11	0.20	0.00	0.38	0.33	0.20	0.00
	ANN PFD	0.00	0.00	0.00	0.00	0.13	0.13	0.00	0.00
	LSTM PFD	0.05	0.07	0.07	0.07	0.19	0.08	0.03	0.00
	ANN PC	0.73	0.92	0.95	0.98	0.82	0.86	0.95	0.98
	LSTM PC	0.71	0.86	0.90	0.92	0.68	0.87	0.94	0.98
	ANN HSS	0.13	0.00	0.00	0.00	0.56	0.42	0.00	0.00
	LSTM HSS	0.09	0.04	0.10	0.00	0.19	0.23	0.19	0.00

Four threshold values in nT/min are used.

Conclusions

- There is some ability for each of the models to predict the timing of magnetic field perturbations, though this ability is not consistently better for either model between the storms and neither is able to predict the magnitude of the enhancements or predict enhancements later in the storm.
- Validation metrics indicate that the LSTM is barely more skilled than random or constant predictions, and that using an empirical fitting improves HSS as it does for first principles-based models.

Further information

MAGICIAN Team Posters at this workshop: Connor (2), Coughlan (53)

This work is Funded by NSF EPSCoR Award OIA-1920965

Keesee AM, Pinto V, Coughlan M, Lennox C, Mahmud MS and Connor HK (2020) *Front. Astron. Space Sci.* 7:550874. doi: 10.3389/fspas.2020.550874