

Introduction

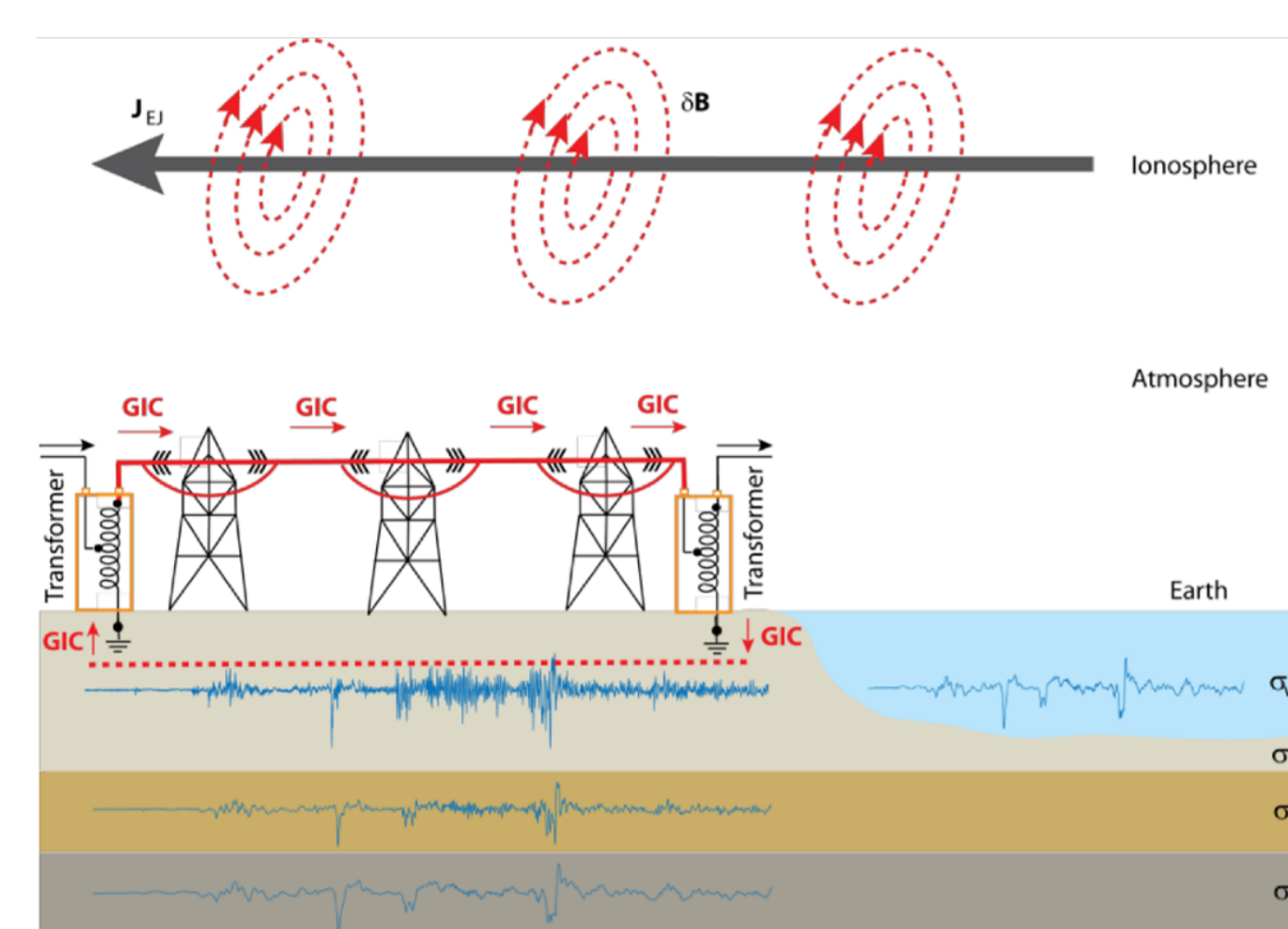


Figure 1. Demonstration of how ionospheric currents can cause GICs when interacting with long conducting infrastructure. Figure also shows the role played by the different conductivity's of ocean water and land mass.

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 use of combining **SHapley Additive exPlication (SHAP)** values with Superposed Epoch Analysis (SEA) 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).
- The model predicts a probability that the $(dB/dt)_{max}$ or RSD_{max} in the region will exceed the 99th percentile (y^0).

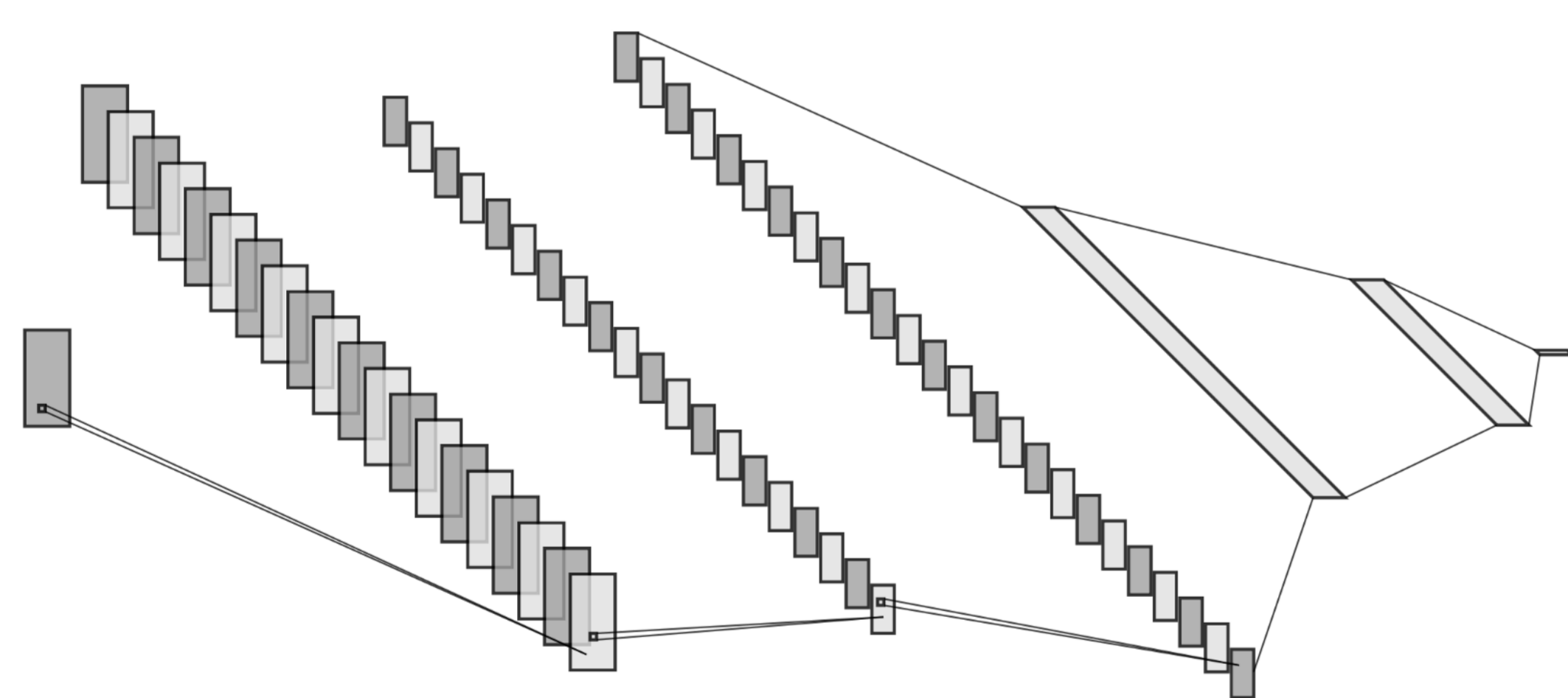
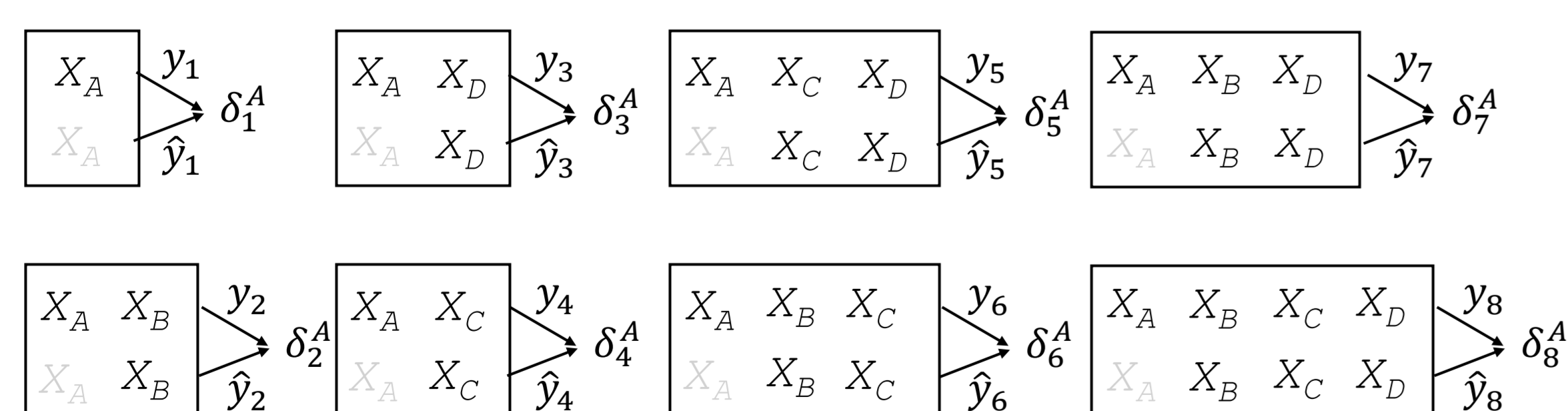


Figure 2. Layout of Convolutional Neural Network model design. Model utilizes two Convolutional layers processing the solar wind and magnetometer time series data. The results are then sent through two fully connected layers. Based off of model used in [1]

Shapley Values



y_i = model output with X_A present

\hat{y}_i = model output without X_A

$$\delta_i^A = y_i - \hat{y}_i$$

$$\text{Shapley Value} = \phi_A = \frac{1}{n} \sum_i \delta_i^A \quad n = 2^{\# \text{ features}}$$

Model Predictions

2013-03-17 Storm

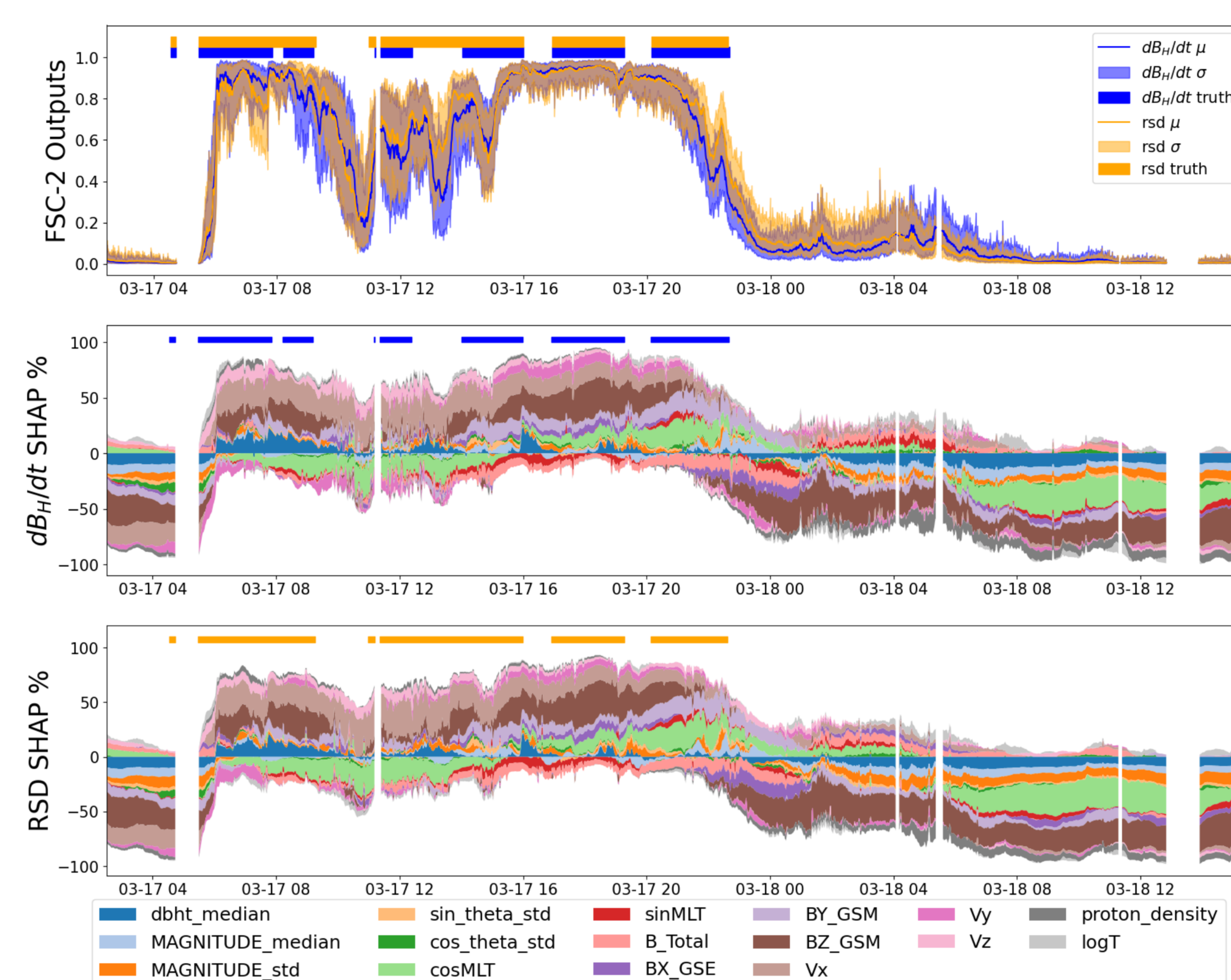


Figure 4. Top panel shows the model outputs for the FSC-2 region during the March 17 2013 Storm, where the blue line indicates the mean probabilities of the dB/dt model and the blue shaded area the uncertainty. Orange is the same for the RSD model. The colored bars at the top of the plot show the times where the ground truth crosses the threshold. The other two panels show the SHAP values for each model as a time series, with values above zero raising the model output probabilities and below zero lowering the model outputs.

Metric Scores

Metric scores for both model types show good performance in most sectors of MLT and MLAT. Model explainability tools are only useful for skillful models. Here we display the area under the calibrated precision-recall curve (pr-auc), the area under the receiver operator curve (roc-auc), and the root mean square error between the mean and the actual values (rmse). The results are presented as a function of Magnetic Local Time (MLT) and Magnetic Latitude (MLAT) for all of the regions identified in this work.

Conclusions

- Combining SHAP values with Superposed Epoch Analysis allows us to gather a statistical picture of the drivers and highlight areas of particular importance.
- Decreases in B_z before SYM-H minimum show high positive influence on model output. V_x and V_y shows a bifurcation around SYM-H minimum, with negative SHAP values dominating before and positive after as V_x and V_y increases.
- Other parameters show positive SHAP values during times of elevated activity.

SHapley Additive exPlication (SHAP)

SHAP values are based on Shapley Values but solve the issue of computation expense by taking a sample of all the possible sub-sets, and replacing missing values with background samples. SHAP is the only model explainability method that satisfies the requirements of:

- Local Accuracy:** Requires the simplified explanation model to at least match the output of the original model for the simplified input x' (which corresponds to the original input x).
- Missingness:** SHAP values for features that are missing or irrelevant are zero.
- Consistency:** If a model changes such that a simplified input's contribution increases or stays the same regardless of the other inputs, that input's attribution should not decrease.

RSD Model Superposed Epoch Analysis

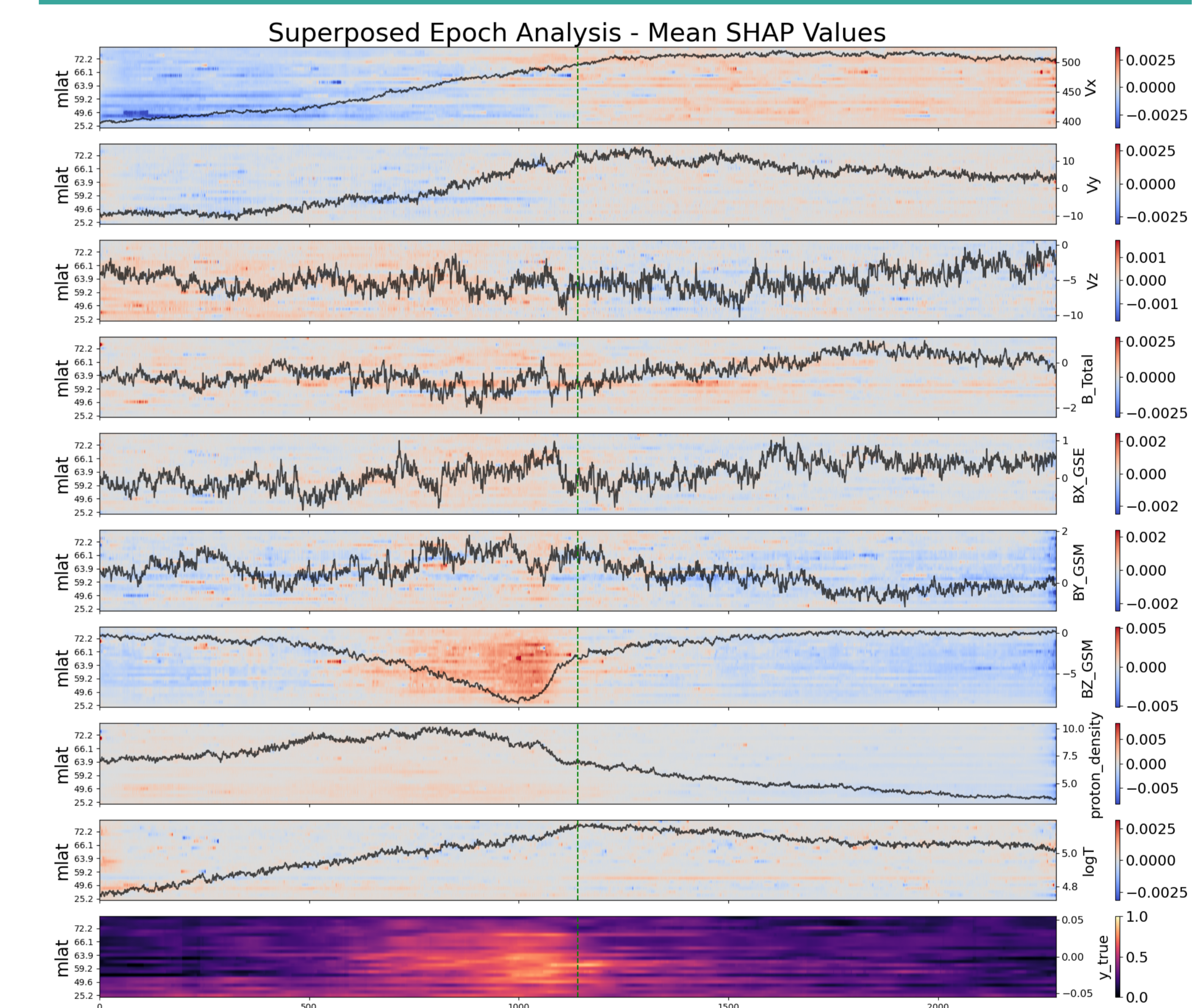


Figure 5. Each panel shows the mean SHAP values (red-blue) and the median parameter values (black) for the solar wind parameters over all the test storms. The green line indicates the point of SYM-H minimum which was used as the zero-epoch. The SHAP values are displayed in order of MLAT with the lowest latitude regions at the bottom of each panel and the highest at the top. The final panel shows the percentage of true values within the test storms for each region. Several of the parameters show clear cut differences around the zero-epoch, notably V_x (first panel) and B_z (7th panel).

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 dB/dt 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 dB/dt 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.