Characterizing Earth's Plasma Sheet through Multi-Spacecraft Observations and Machine Learning

Savvas Raptis[1], Connor O'Brien[2], Kareem Sorathia[1], Viacheslav Merkin[1]
Louis Richard[3], Simon Wing[1]

[1]The Johns Hopkins University Applied Physics Laboratory, Laurel, MD, USA
[2]Center for Space Physics, Boston University, Boston, MA, United States
[3]Swedish Institute of Space Physics, Uppsala, Sweden

Abstract

Earth's plasma sheet is a highly dynamic environment characterized by multiscale phenomena, fast plasma flows and reconnection, making it exceptionally complex to model. While physics-based simulations and empirical models based on in-situ observations have been employed, advanced statistical and machine learning (ML) modeling has not been widely used in this environment.

In this work, we leverage extensive in-situ data from the Geotail and Magnetospheric Multiscale (MMS) missions to address this gap. We use the vast Geotail dataset, comprising millions of data points, to develop ML-driven 2D maps of plasma sheet properties, specifically focusing on plasma density (n) and ion temperature (Ti). Additionally, we use the state-of-the-art MMS measurements of electron and ion moments to model the ion-to-electron temperature ratio (Ti/Te). This ratio is not only physically insightful but also crucial for initializing particle distributions in physical models.

For this regression task, our model inputs include the time history of solar wind data (propagated OMNIWeb values or direct L1 Wind measurements) and geomagnetic indices such as DST and AE/AL. The model's output is a particle moment that describes the large-scale properties of the plasma sheet under these driving conditions. We evaluated a range of modeling techniques, from linear regression to gradient boosting and Long Short-Term Memory (LSTM) models, and compared their performance against existing empirical models derived from Geotail and THEMIS observations.

Our results indicate that model complexity, architecture, and input variables have a limited effect on performance, suggesting an under-determined system. Even so, our ML approaches systematically outperformed previous empirical efforts across all metrics, demonstrating increased applicability. ML models additionally introduce physically accurate spatial asymmetries such as an elevated dawnside density and an increased duskside ion temperature, which were not captured by previous empirical models that contained predefined symmetries.

Finally, testing the model against THEMIS observations during a geomagnetic storm revealed that sparse measurements of rare, strong events can limit the accuracy of such modeling efforts. This suggests that the most promising path forward may be a fusion between physics-based and ML data-driven models.