Machine-/Deep-Learning Applications to Space Science

M. Maimaiti (Muhammad Rafiq)

Virginia Tech

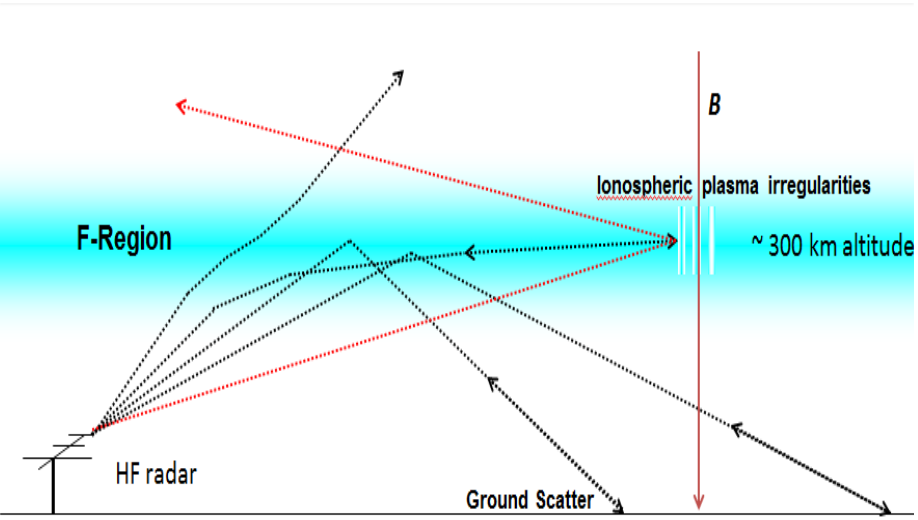
Presentation Outline

- **Project I**: Characterizing SuperDARN Backscatter Echoes using Machine Learning Algorithm (*manuscript in preparation*)

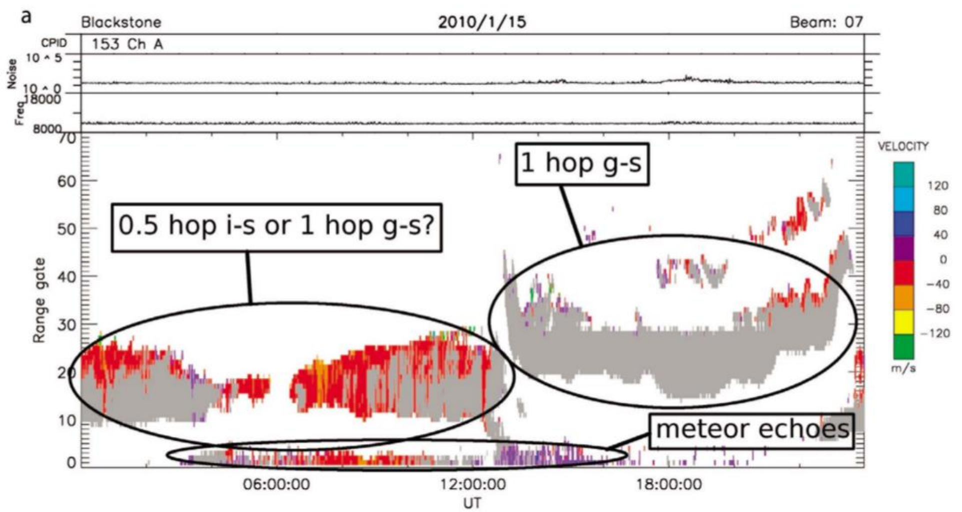
- **Project II**: A Deep Learning Based Approach to Forecast GPS TEC over the North American Sector (*manuscript in preparation*)

- **Project III**: A deep Learning Based Approach to Forecast the Onset of Magnetic Substorms (*manuscript submitted to Space Weather*)

Project I: Characterizing SuperDARN Backscatter Echoes using Machine Learning Algorithm – Background



HF rays are refracted in the ionosphere as they encounter gradients in electron density



RTI Plot showing SuperDARN Blackstone Radar Data

- Separating different types of scatters (ionospheric, ground, meteor and mixed) is one of the major challenges [Riberio et al. 2011; Burrell et al. 2018] and still an open area of research in SuperDARN.
- **Problem statement**: Use unsupervised machine learning (Gaussian Mixture Model – GMM) to improve SuperDARN ground/ionospheric scatter classification.

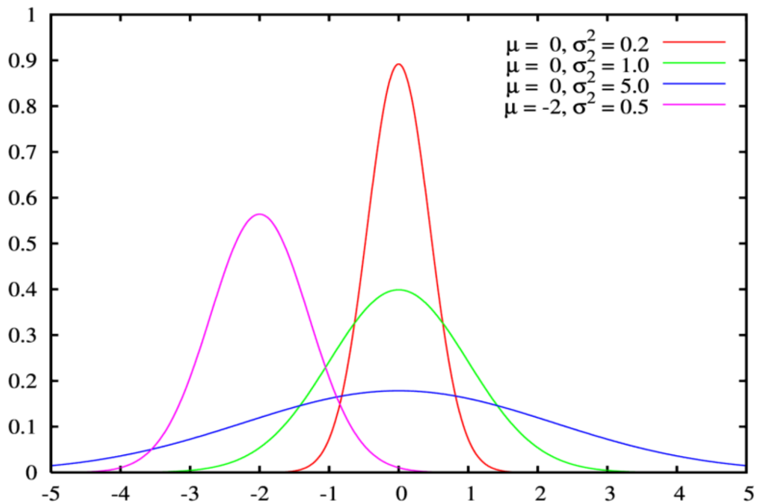
Project I: Characterizing SuperDARN Backscatter Echoes using Machine Learning Algorithm – Dataset & Algorithms

Dataset: SuperDARN high latitude Saskatoon and mid latitude Christmas Valley radars.

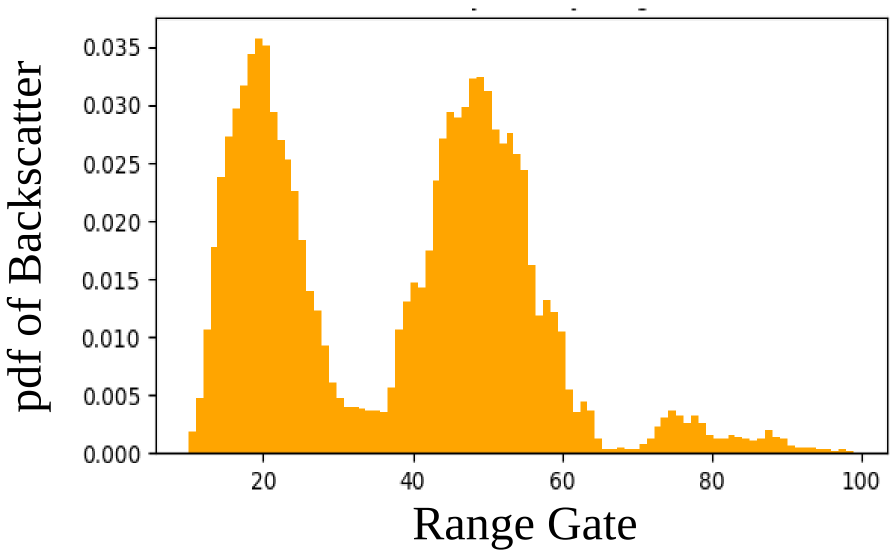
Features: Doppler Velocity, Spectral Width, Range Gate, Beam Number, Time of day.

Gaussian Mixture Model: GMM is a probabilistic model that assumes all the data points are generated from a mixture of a finite number of Gaussian distributions with unknown parameters.

- Useful for non-spatial feature clustering (Advantage).
- Can create clusters with arbitrarily high variance, which often don't make physical sense (Disadvantage).
- Have to tune number of clusters (hyper parameter) for each day (Limitation).



Example of GMM



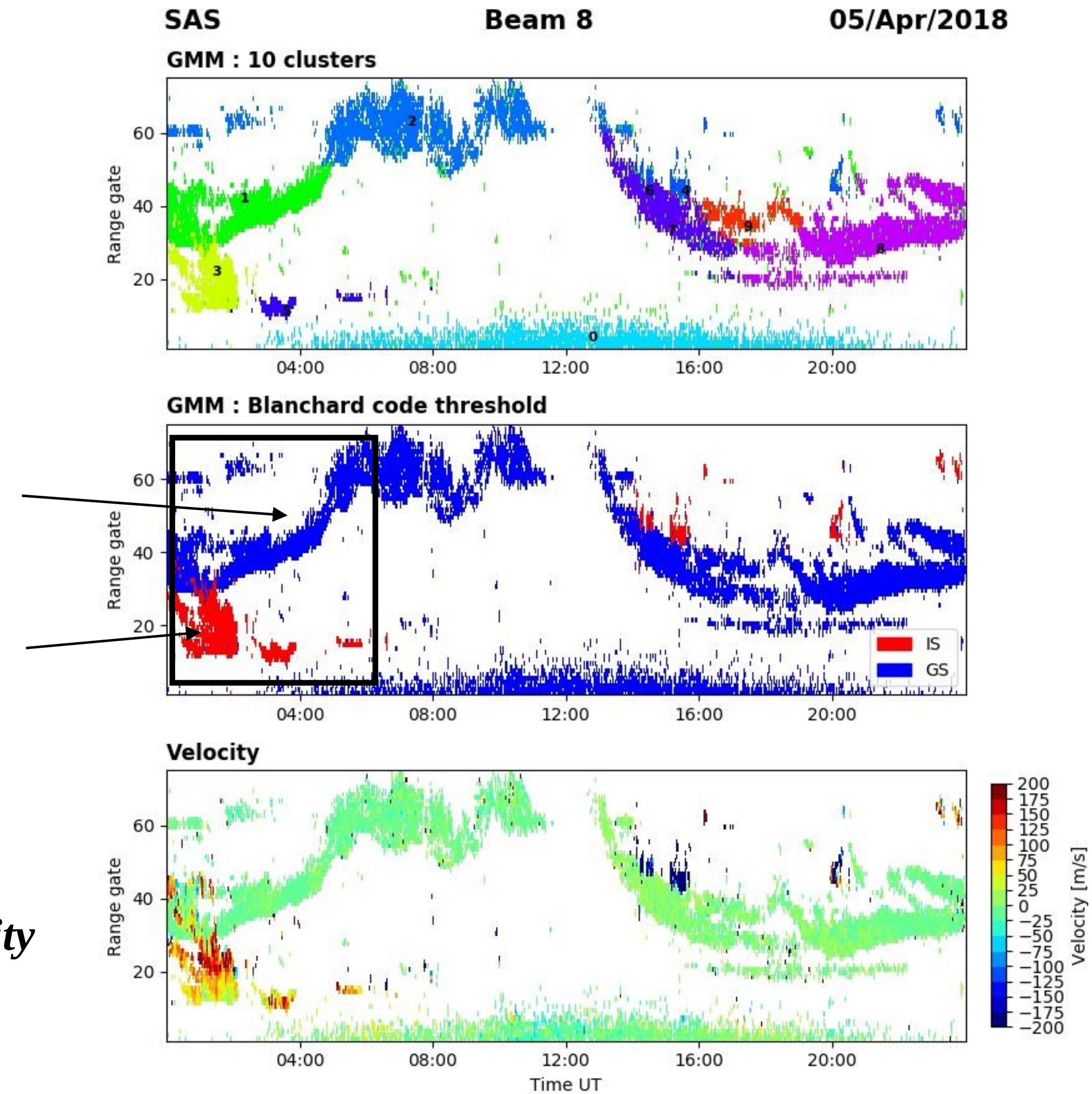
Project I: Characterizing SuperDARN Backscatter Echoes using Machine Learning Algorithm – Sample Results

GMM:10 Clusters

Ground Scatter

Ionospheric Scatter

Line-of-Sight Velocity

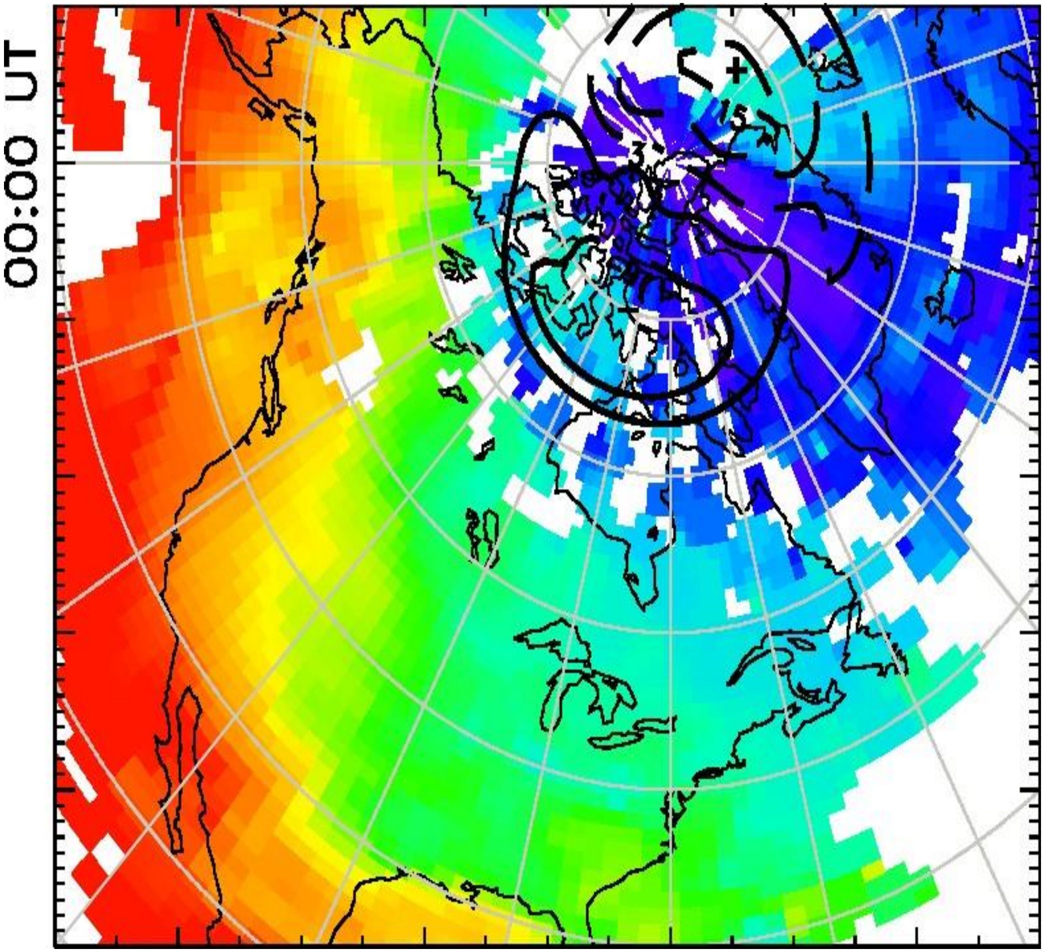


Shi, X., Robb, E., Chakraborty, S., Ruohoniemi, J. M., Baker, J. B. H. , Burrell, A. G., Maimaiti, M. (TBD). Characterizing SuperDARN Backscatter Echoes using Machine Learning Algorithm. *Radio Science* (in preparation)

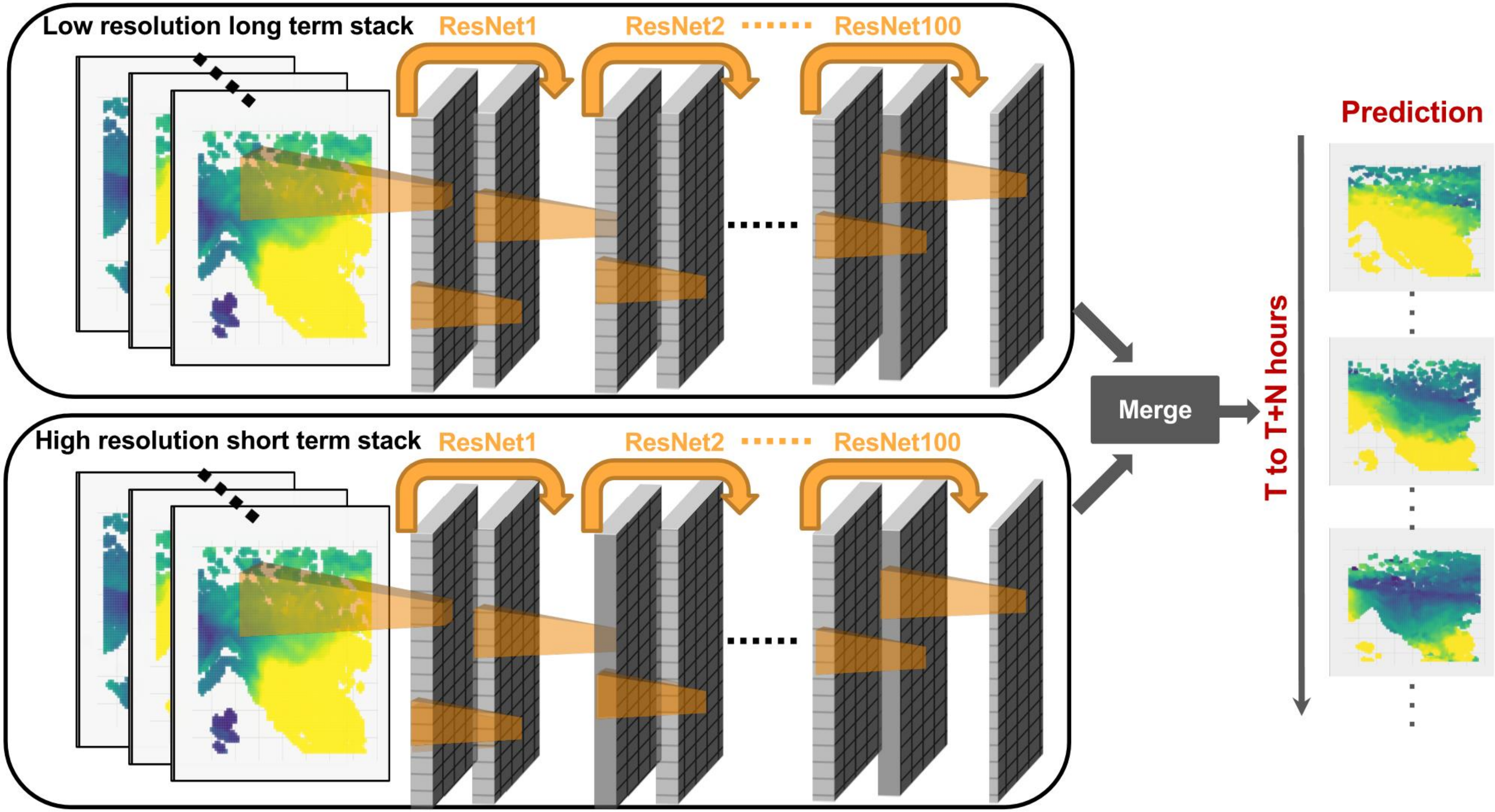
Project II: A Deep Learning Based Approach to Forecast GPS TEC over the North American Sector – Background

- Total Electron Content (TEC): the integral of electron density in a 1 m^2 column along the signal transmission path.
- 1x2 (MLAT, MLON) bins of GPS TEC values.
- Median filtered GPS TEC values over the North American Sector are used.
- **Problem statement**: Given a history of GPS TEC values (maps), predict TEC values over the entire North American continent for the next 1, 2 and 4 hours.

GPS TEC map



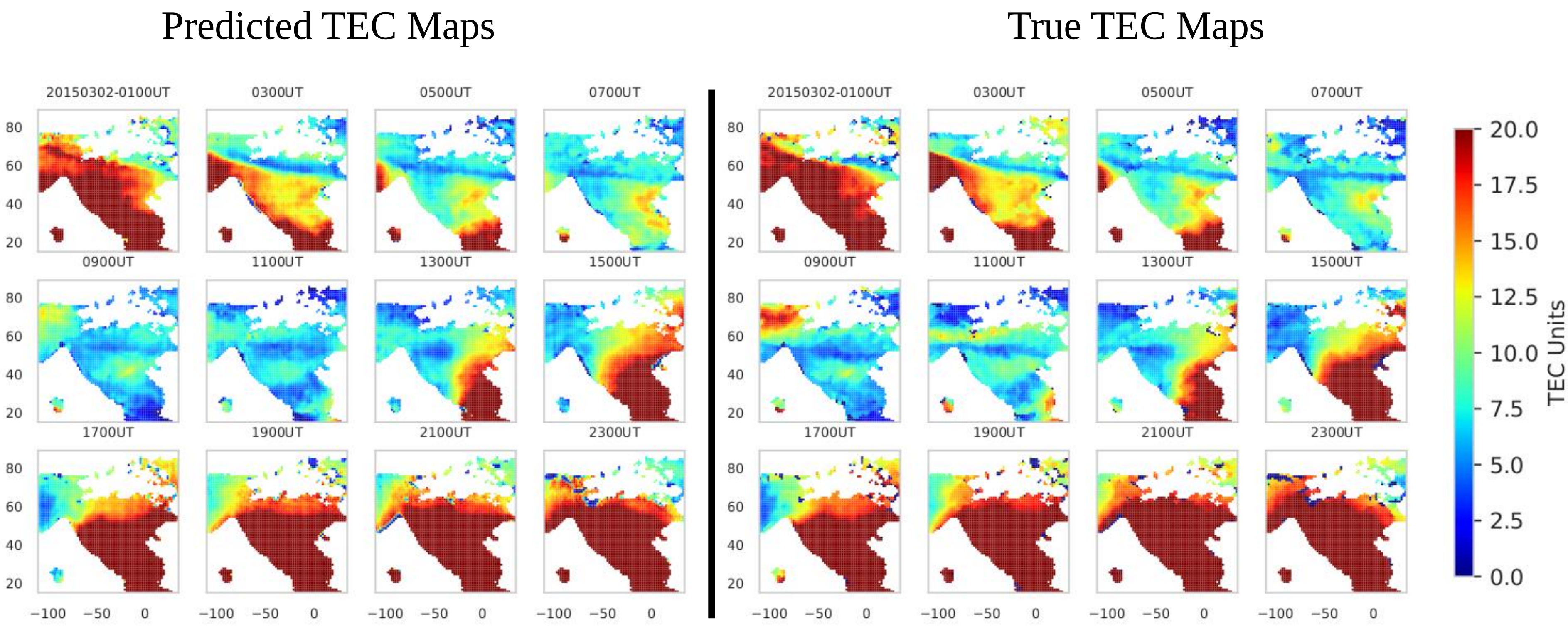
Project II: A Deep Learning Based Approach to Forecast GPS TEC over the North American Sector – Model Architecture



Architecture of the deep-learning model used for GPS TEC prediction.

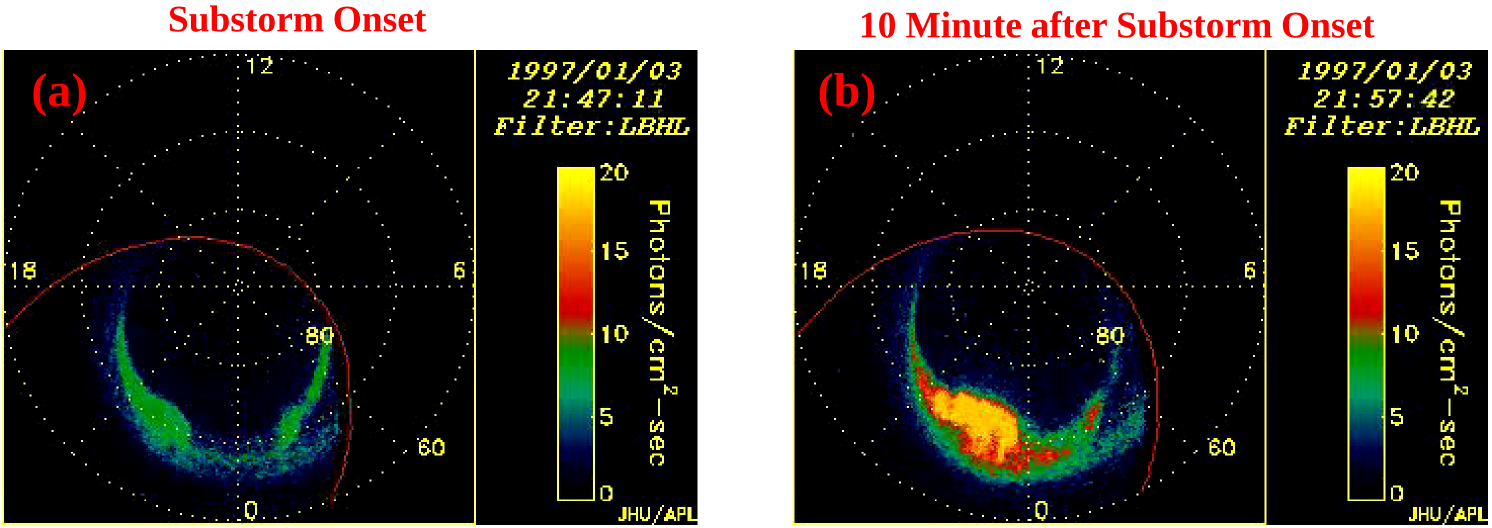
Project II: A Deep Learning Based Approach to Forecast GPS TEC over the North American Sector – Sample Predictions

Model prediction on Mar-02-2015

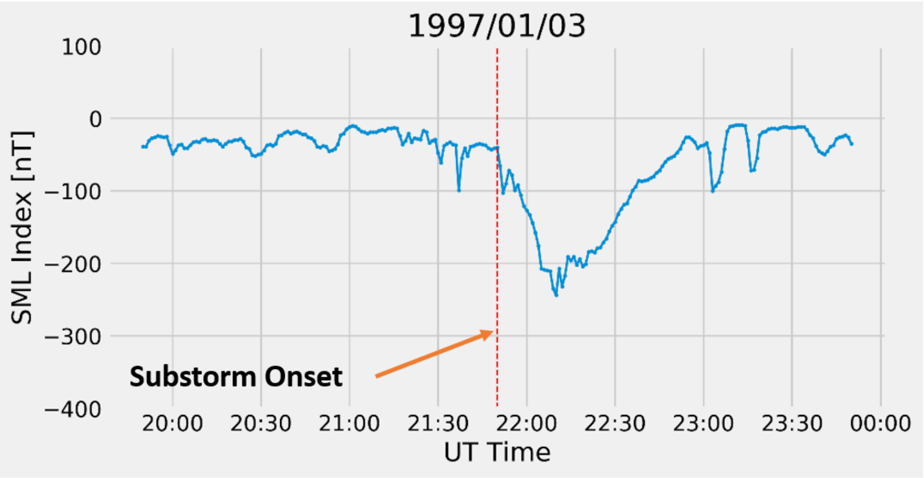


B. Kunduri, M. Maimaiti, J. B. H. Baker, J. M. Ruohoniemi (2019). A deep learning based approach to forecast the onset of magnetic substorms. *Space Weather*. (<u>In preparation</u>)

Project III: A deep Learning Based Approach to Forecast the Onset of Magnetic Substorms – Background



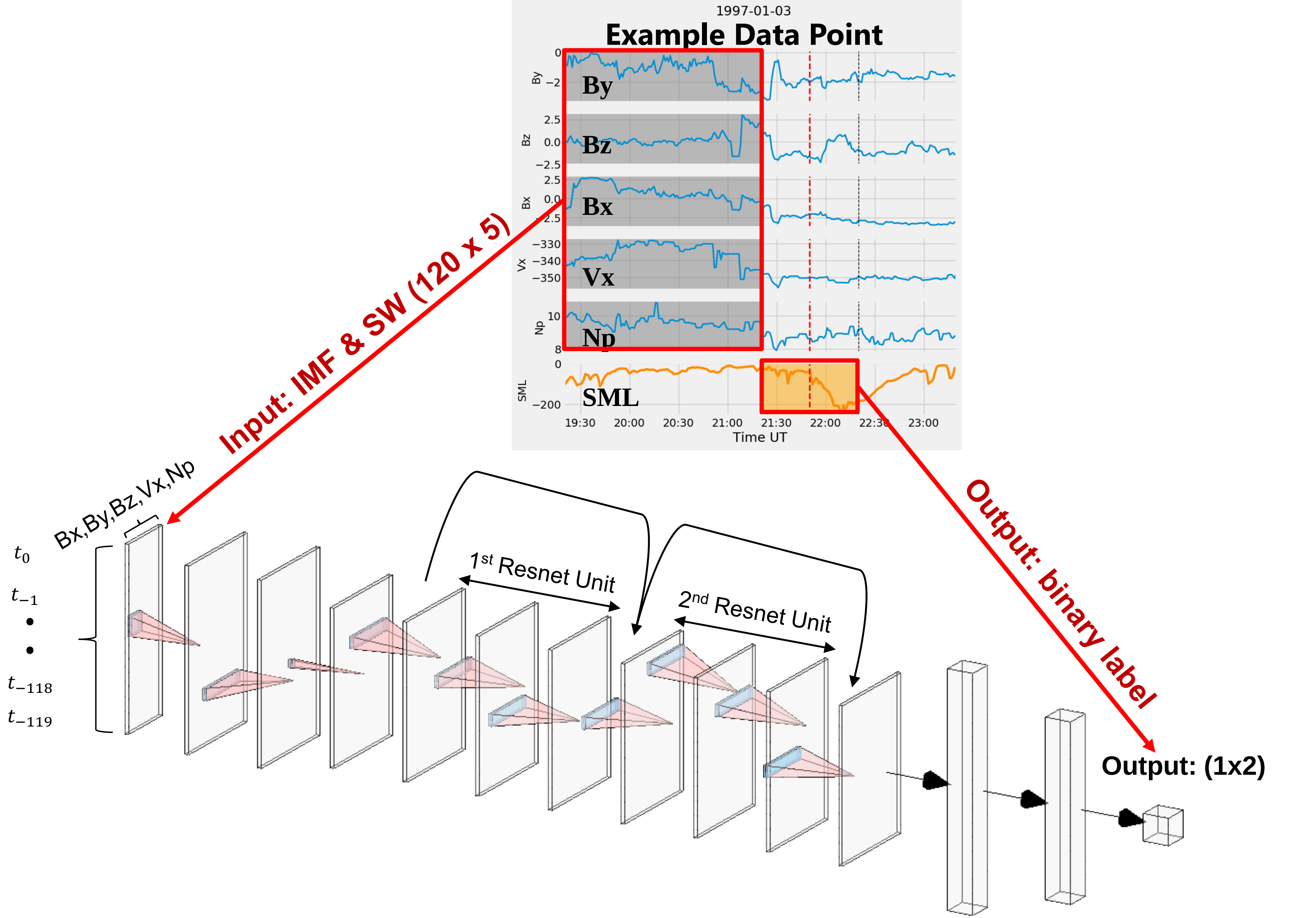
An example of an auroral substorm occurred at **21:47 UT** on January 3rd, 1997, captured by cameras onboard of Polar UVI satellites.



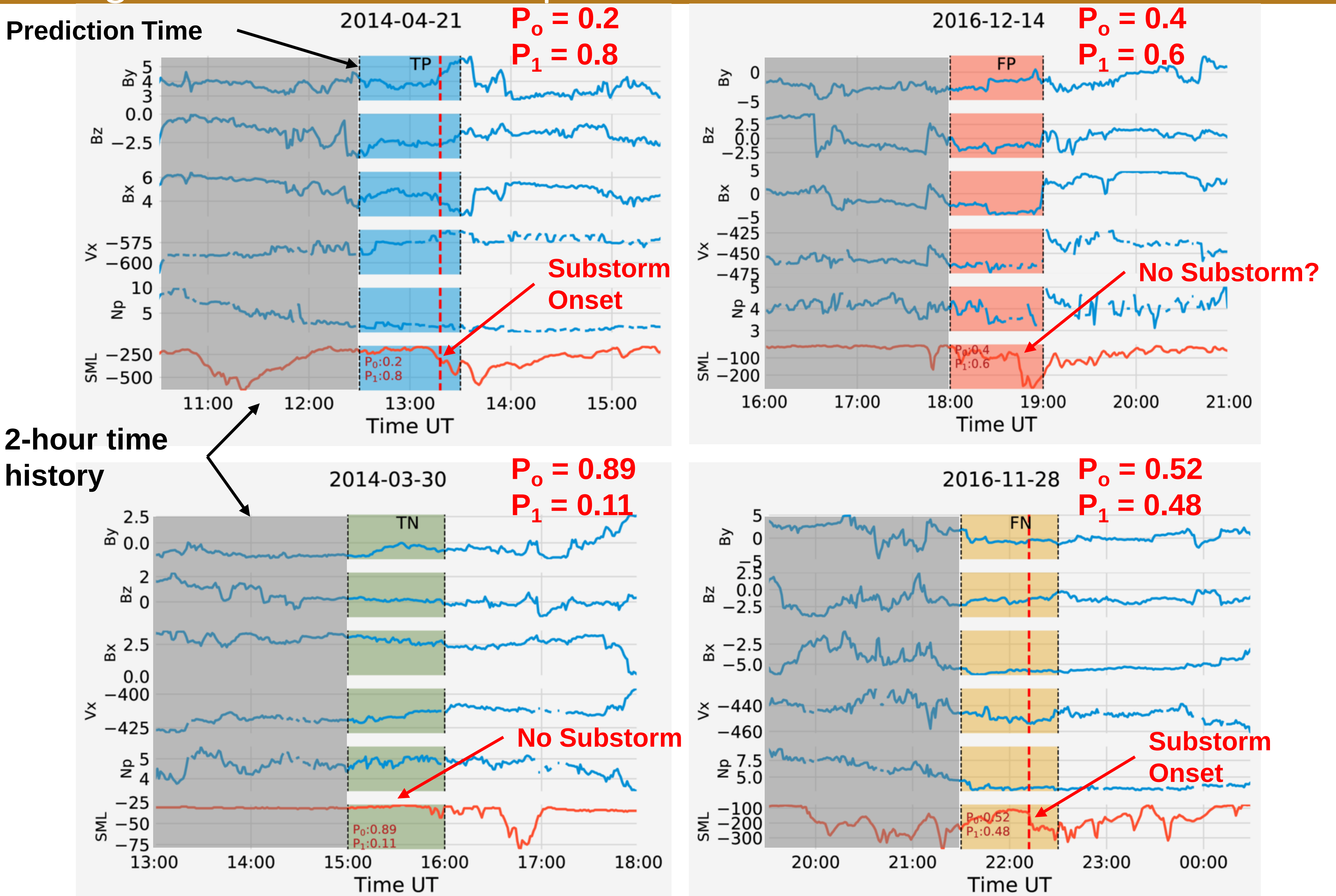
An example of a magnetic substorm occurred at **21:50 UT** on January 3rd, 1997, as seen from SML index.

<u>Problem Statement</u>: Use 120-minute time history of solar wind bulk speed (Vx), proton number density (Np), and interplanetary magnetic field (IMF) components (Bx, By, Bz) to predict the occurrence probability of a magnetic substorm onset.

Project III: A deep Learning Based Approach to Forecast the Onset of Magnetic Substorms – Model Architecture



Project III: A deep Learning Based Approach to Forecast the Onset of Magnetic Substorms – Sample Predictions



Applications of Data Science and Engineering to Magnetospheric Physics and Aeronomy

SPA-Magnetospheric Physics (SM005)

Space science is a data-rich field, comprising data from many space- and ground-based instruments developed over the last several decades. This dataset is made up of numerous important parameters measured in the **near Earth space environment** and provide crucial information required to answer important questions in the field. However, working with such massive datasets comes with its own set of challenges, both in terms of **data engineering and analysis**. Recently there has been a remarkable advancement in the field of **data science** which can be attributed to better infrastructure and availability of many **open source libraries**. This session solicits contributions describing applications of **data science and engineering, data mining, and machine learning to space science**. We particularly welcome contributions focusing on applications of tools and techniques that provide **real time analysis**, assist scientific operation and discovery, or contribute to the advancement of science in general.

AGU Fall Meeting 2019

AGU Fall Meeting 2019
American Geophysical Union Fall Meeting
Date: 12/9/2019 - 12/13/2019
Venue: Moscone Center, San Francisco CA, United States

Abstracts due: July 31, 2019

Contact:
Bharat Kunduri (bharatr@vt.edu)
Muhammad Rafiq (rafiq@vt.edu)
Asti Bhatt (asti.bhatt@sri.com)
Ashton Reimer (ashton.reimer@sri.com)

Thank You!