Novel approaches to geospace particle transfer in the digital age: Progress through data science

Ryan McGranaghan\textsuperscript{1}, Enrico Camporeale\textsuperscript{2}, Kristina Lynch\textsuperscript{3}, Jesper Gjerloev\textsuperscript{4}, Téo Bloch\textsuperscript{5}, Spencer Hatch\textsuperscript{6}, Binzheng Zhang\textsuperscript{7}, Pete Riley\textsuperscript{8}, Mathew Owens\textsuperscript{5}, Yuri Shprits\textsuperscript{9}, Irina Zhelavskaya\textsuperscript{10}, and Susan Skone\textsuperscript{11}

\textsuperscript{1}ASTRA LLC
\textsuperscript{2}NOAA
\textsuperscript{3}Dartmouth College
\textsuperscript{4}JHU APL
\textsuperscript{5}University of Reading
\textsuperscript{6}University of Bergen
\textsuperscript{7}Hong Kong University
\textsuperscript{8}Predictive Science Inc.
\textsuperscript{9}University of California Los Angeles
\textsuperscript{10}GFZ Potsdam
\textsuperscript{11}University of Calgary

November 24, 2022

Abstract

The magnetosphere, ionosphere and thermosphere (MIT) act as a coherently integrated system (geospace), driven in part by solar influences and characterized by variability and complexity. Among the most important and yet uncertain aspects of the geospace system is energy and momentum coupling between regions, which is, in part, accomplished by the transfer of charged particles from the magnetosphere to the ionosphere in a process known as particle precipitation, and in the opposite direction by ion outflow. Both processes are inherently multiscale and manifest the variabilities and complexities of the geospace system. Despite the importance of the transfer of particles, existing models are increasingly ill-equipped to provide the specification necessary for the growing demand for geospace now- and forecasts. Due to recent trends in the availability of data, we now face an exciting opportunity to progress particle transfer in geospace through the intersection of traditional approaches and state-of-the-art data-driven sciences. We reveal novel particle transfer models utilizing machine learning (ML), present results from the models, and provide an evaluation of their capabilities including comparisons with observations and the current 'state-of-the-art' models (e.g., OVATION Prime for particle precipitation and the Gamera-Ionosphere Polar Wind Model for ion outflow). We detail the data wrangling required to utilize the available geospace observations to make progress on the long-standing challenge of particle transfer and place specific emphasis on the discovery possible when ML models are appropriate and robustly interrogated in the context of physical understanding. Our presentation helps illustrate the trends in the application of data science in space science.
Energy and momentum transfer via particles between the magnetosphere and the ionosphere (particle precipitation down and ion outflow up) represents one of the great unsolved topics across Heliophysics and Space Weather, and an ideal use case to advance the systems science perspective. We utilized data science (i.e., the full data science pipeline) to bring traditional approaches and state-of-the-art data science techniques from e.g., Silicon Valley to bring data scientists and scientists closer together. The result was not only cutting edge advances, but also the emergence of new ethos of novel collaboration for radically interdisciplinary teams, the utility of a data science-driven approach in the Earth and Space Sciences, and a flourishing Community of Practice (CoP). This poster reveals the particle precipitation component of the ISSI results.

**Background and the need for data science**

We carried out a research team study at the International Space Sciences Institute (ISSI) to evaluate and a flourishing Community of Practice (CoP). This poster reveals the potential of novel collaboration for radically interdisciplinary teams, the utility of a data science-driven approach in the Earth and Space Sciences, and a flourishing Community of Practice (CoP). This poster reveals the particle precipitation component of the ISSI results.

**Unique approach to collaboration**

Due to recent trends in the availability of data, we now face an exciting opportunity to progress space science understanding through the intersection of traditional approaches and state-of-the-art data science-driven sciences (McGranaghan et al., 2017). We utilized data science (i.e., the full data science pipeline) and novel collaboration techniques from e.g., Silicon Valley to bring data scientists and scientists closer together.

**The particle precipitation challenge**

Different sets of input parameters used in the particle precipitation models rely on a difference in philosophy of approach and lead to a difference in capabilities. Existing models are limited in their ability to reproduce observed features that are associated with large spatial gradients and that occur rapidly.

We present a new machine learning model (hereafter M2019) that utilizes the expressive power of deep neural networks to incorporate both solar wind and magnetosphere-ionosphere (MI) state descriptors and to be capable of specifying substorm-scale (space and time) phenomena.

**Analysis Ready Data**

Defense Meteorological Satellite Program (DMSP) in-situ particle precipitation observations

Collate solar wind and geomagnetic activity data as input ‘features’

Reveal that data contain observations of substorm phenomena

Include time history of input features

Compile into ‘Analysis Ready Data’ [McGranaghan et al., 2018]

**Interrogating the model**

Integrate physical understanding with the data science-driven approach

**What is the impact of data volume?**

Add each year’s observations to the dataset: more data improves model performance on novel data across standard assessment metrics

Higher data volumes reduce error over all metrics.

Finer granularity needed at large volumes to determine benefit of additional DMSP data

**What is the impact of model hyperparameters?**

Increasing the expressive power of a model is to expand the breadth of ideas that can be represented and communicated

**Explore with NASA Frontier Development Laboratory:**

- Potential new challenge concept in 2020 (proposed)
- What is the impact of model hyperparameters?
- Explore with NASA Frontier Development Laboratory:
- What's next?

**Acknowledgements**

The International Space Sciences Institute and the fantastic Particle Transport Team (http://www.isissi.ch/team/multiagencypartnership.htm)

NASA LWS for partial support

DAIR-SEE program for partial support

SuperMAG and Rob Barnes for providing the POLAR VIS data

Full list of references and additional materials online (https://github.com/mcgranaghan/Ag2019)

**A Flourishing Community of Practice**

Community of Practice (CoP)

- Long-term: Multi-disciplinary, multi-institutional
- Short-term: Project-driven

Can we transfer the methodology to other challenges?

- Ionospheric ion outflow
- What's next?

**Help shape a flourishing space physics community**

- Town Hall: Antidisciplinary: Science and engineering in the digital age
- Join a newly interdisciplinary group to shape the Future for the Earth and Space Sciences

**Scale and extend**

- Grow the database: Fast Auroral SnapshotT (FAST) observations
- Incorporate ground-based magnetometers: ‘Greatest potential for predicting auralor’ [Newall et al., 2018]
- Incorporate auroral imagery: Space- and Ground-based
- Compare with state-of-the-art magnetosphere model: GAMERA precipitation