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

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

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

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

2019 International Space Sciences Institute Team

Ryan McGranaghan, Enrico Camporeale, Kristina Lynch, Spencer Hatch, Susan Skone, Bin Zhang, Irina Zhelavskaya, Mathew Owens, Teo Bloch, Jesper Gjerloev, Pete Riley, Yuri Shprits

@AeroSciengineer
rmcgranaghan
ryan.mcgranaghan@gmail.com

This poster in 30 seconds…

1. Demonstrate a new 'state-of-the-art' particle precipitation model capable of mesoscale (substorm-scale) specification and prediction
2. Reveal progress enabled by 'data science' approach (i.e., considering the full data lifecycle)
3. Show that data science-driven mindset permits new models of collaboration

Background and the need for data 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 unresolved topics across Heliophysics and Space Weather, and an ideal use case to advance the systems science perspective.

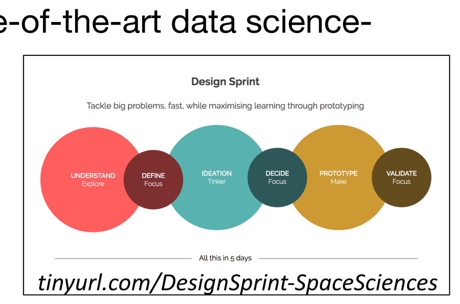
We carried out a research team study at the International Space Sciences Institute (ISSI) to evaluate the existing data landscape for particle precipitation and ion outflow. Our focus was to bring to bear novel data science tools and technologies to produce profound new insight and capability for particle transport specification and prediction.

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.

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 lifecycle approach) and novel collaboration techniques from e.g., Silicon Valley to bring data scientists and scientists closer together



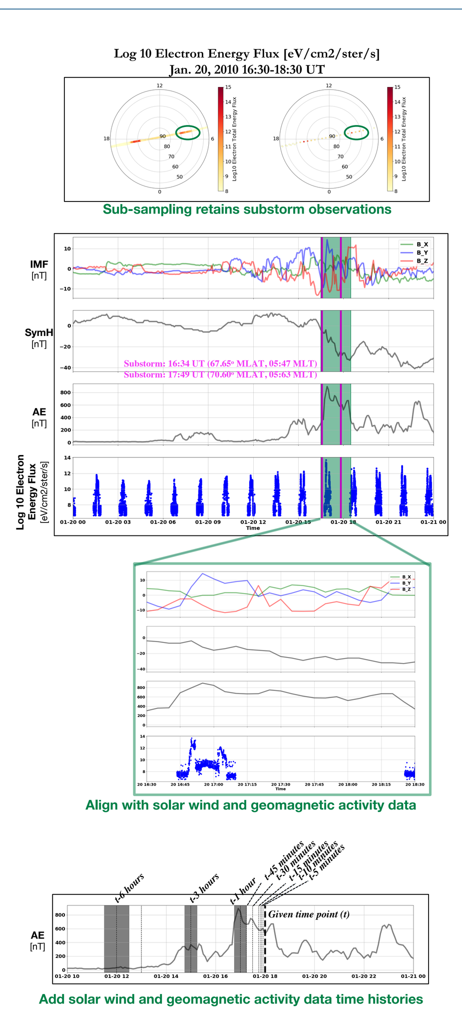
The particle precipitation challenge

Different sets of input parameters used in the particle precipitation models rely of 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]

Progress on the grand challenges of space physics in the digital age

Magnetosphere-Ionosphere particle transport at substorm-scale

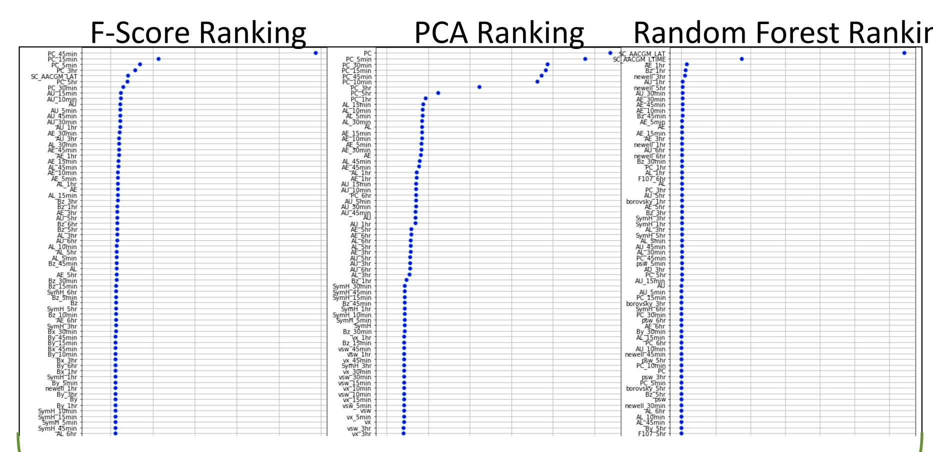
A data-driven approach to the massive data and model design space

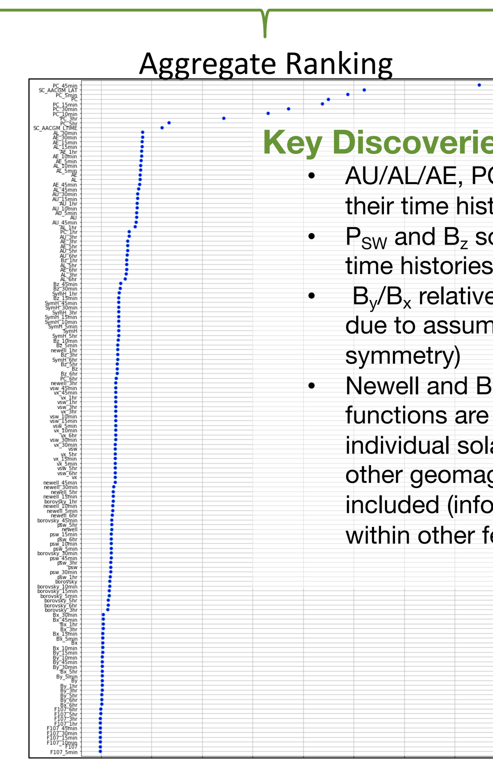
Explore critical questions with data science experiments

What input 'features' are most important?*

'The major improvements in precipitation models have always arisen from improved geophysical insights into the choice of organizing parameters' [*Newell et al.*, 2015]

*Explore the full results online





Key Discoveries

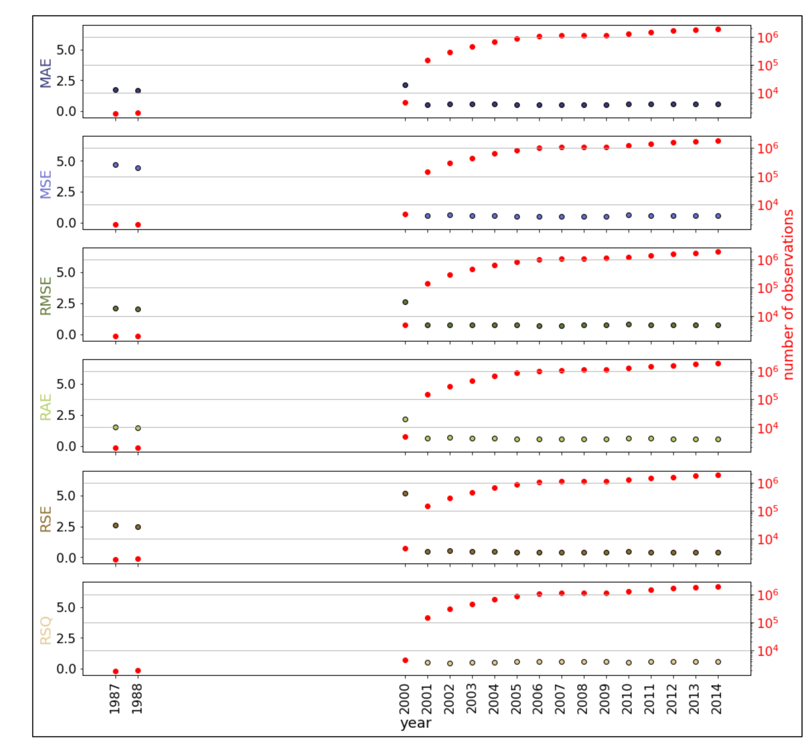
- AU/AL/AE, PC, and Sym-H indices (and their time histories) are important
- P_{SW} and B_z solar wind variables and time histories are important
- B_y/B_x relatively uninformative (likely due to assumption of hemispheric symmetry)
- Newell and Borovsky coupling functions are unimportant when individual solar wind variables and other geomagnetic activity indices are included (information is contained within other features)

What is the impact of data volume?

Add each year's observations cumulatively and assess ML model performance on novel data across standard assessment metrics

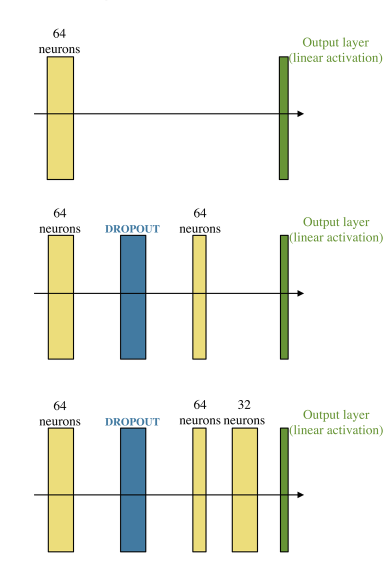
Higher data volumes reduce error over all metrics

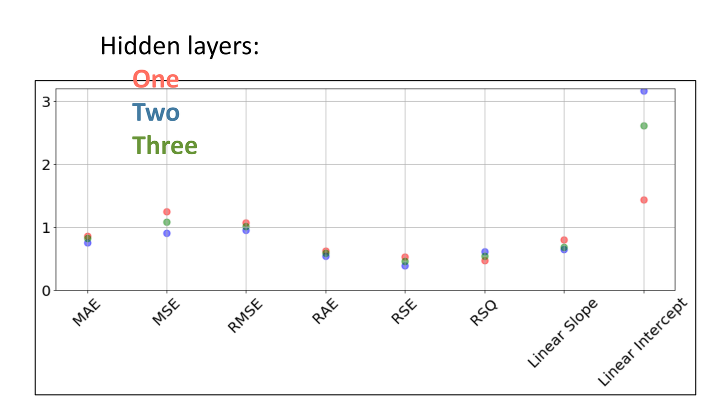
Finer granularity needed at large volumes to determine benefit of additional DMSP data



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





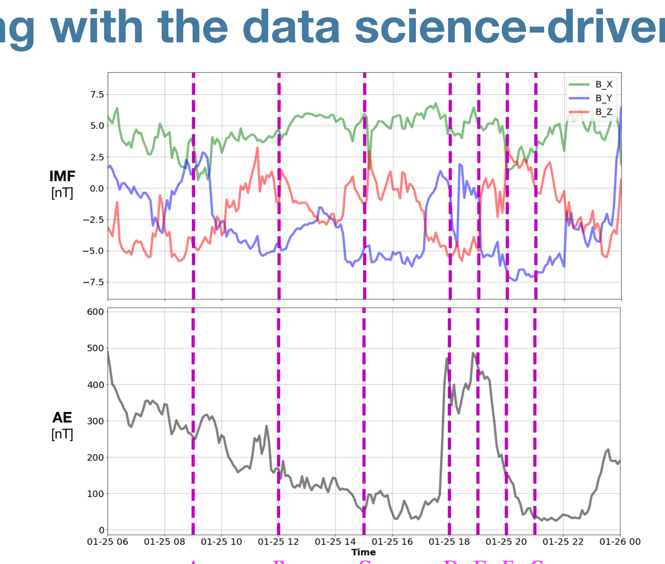
*Explore more architectures and hyperparameter variations online

Bias-Variance tradeoff apparent

Three hidden layers architecture best balances performance across metrics

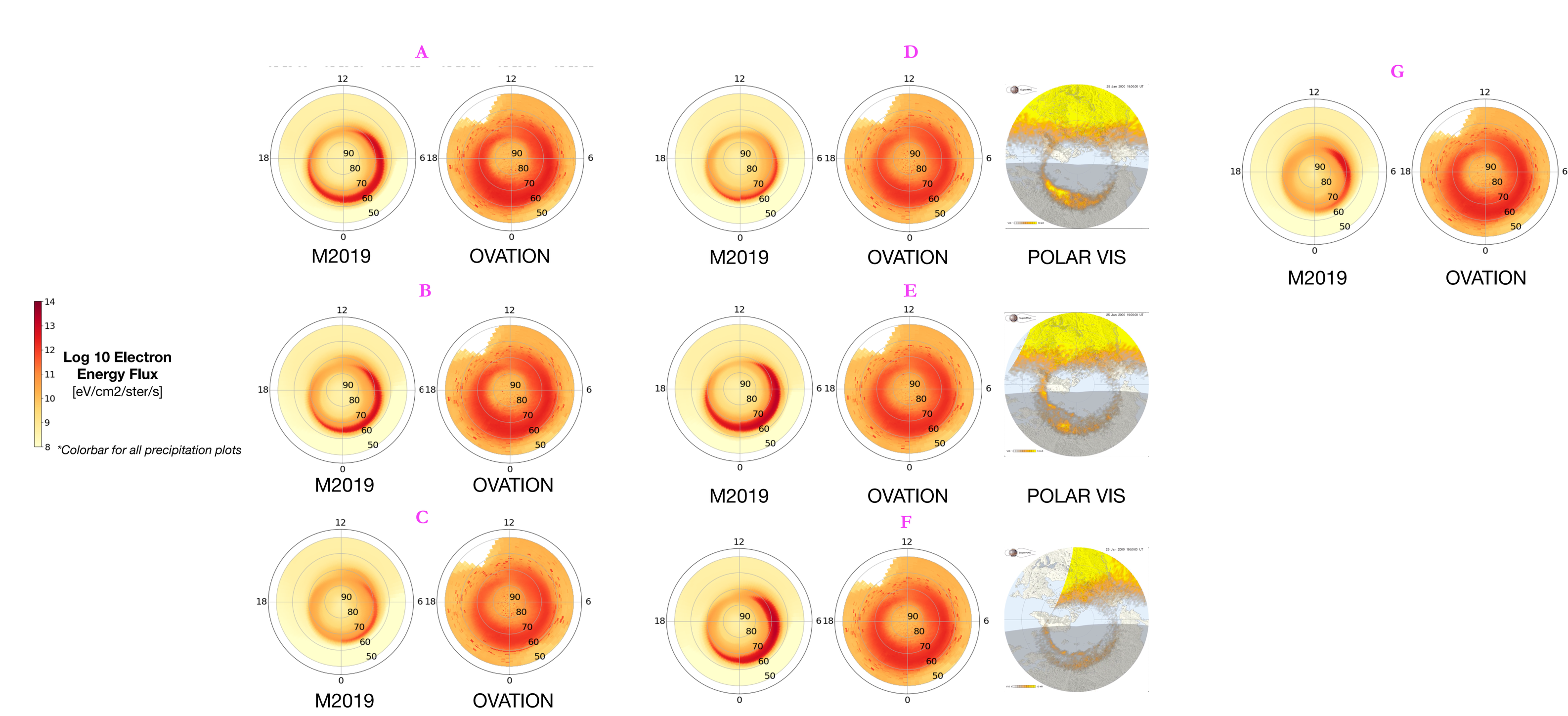
Interrogating the model

Integrate physical understanding with the data science-driven approach*



*Explore our robust interrogation across multiple 'levels' online, including:

- *Through standard assessment metrics on novel data [Liemohn et al., 2018]*
- *Capability to specify auroral boundaries and hemispheric power*
- *Capability to reproduce known physical phenomena*
- *With 'next-generation' metrics (e.g., image-to-image comparison metrics)*



What's next?

Scale and extend

Grow the database:
Fast Auroral SnapshoT (FAST) observations

Incorporate ground-based magnetometers:
'Greatest potential for predicting auroral power' [*Newell et al.*, 2018]

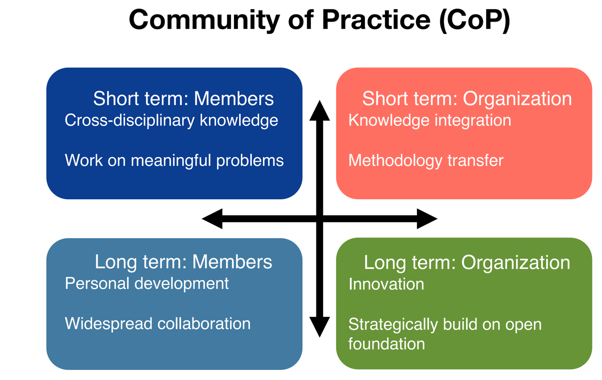
Incorporate auroral imagery:
Space- and Ground-based

Compare with state-of-the-art magnetosphere model:
GAMERA precipitation

Explore with NASA Frontier Development Laboratory:
Potential new challenge concept in 2020 (*proposed*)

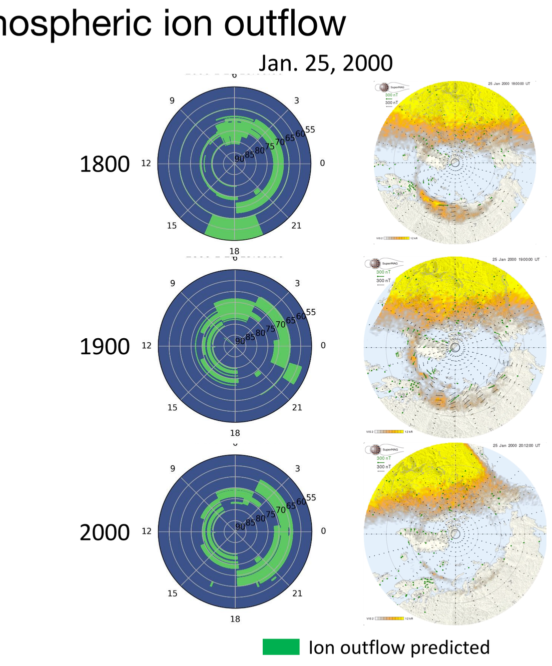
Additionally, our ISSI team's ethos of 'open by default' led to new data sets and software that will allow the community to amplify these efforts and join the community of practice…

A Flourishing Community of Practice



Can we transfer the methodology to other challenges?

Ionospheric ion outflow



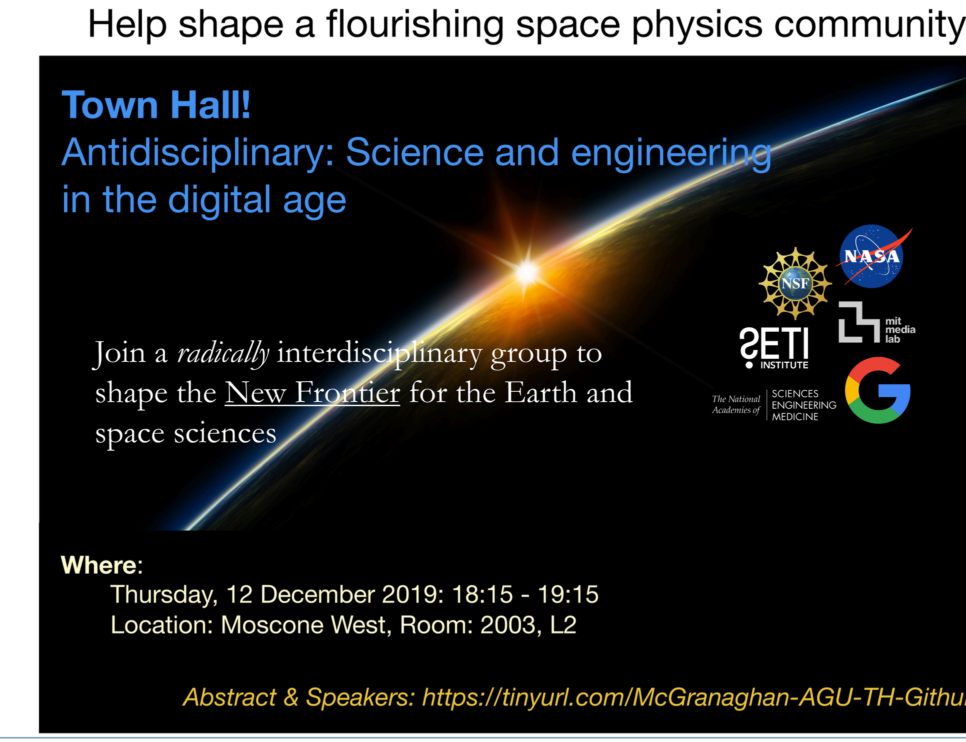
Classification algorithm using similar inputs as particle precipitation model

Hot spots of predicted energetic ion outflow are in the cusp and pre-midnight, as they should be

Initial results show correspondence with substorm behavior

*Courtesy of Téo Bloch and Spencer Hatch

Help shape a flourishing space physics community



Acknowledgements

The International Space Sciences Institute and the ***fantastic*** Particle Transport Team (http://www.issibern.ch/teams/multigeopartransfer/)

NASA LWS for partial support

DARPA 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/rmcgranaghan/AGU2019)