Featurizing an In-situ and Imagery Conjunction Database for Auroral Current Closure Studies: Implications for Magnetosphere-Ionosphere Coupling Studies

Alexander Mule, Kristina Lynch, Maia Kawamura, Grace Connolly, and the multiple providers of the data we use

Abstract:

Featurizing an In-situ and Imagery Conjunction Database for Auroral Current Closure Studies: Implications for Magnetosphere-Ionosphere Coupling Studies

The relationship between flow, Field Aligned Current (FAC), and conductance in an auroral arc is governed by the current continuity equation. In an idealized, sheetlike discrete arc, a single satellite measuring electric and magnetic field along a trajectory can reconstruct much of the detail of current closure in the arc. Even relaxing sheetlikeness assumptions, the combination of in-situ and imagery data for simultaneous use in analyzing current closure provides a powerful tool for studying ionospheric system science.

We have developed several tools to help extract as much information as possible from conjugate spacecraft and auroral imagery data - these include a wavelet transform based featurization of the Swarm ion flow and B field cuts, a routine that uses Swarm B/E ratios to determine conductance in sufficiently sheetlike arcs, and a tool to automatically distinguish, track, and featurize arcs in imagery.

We have used these tools, together with published databases of Swarm spacecraft data and THEMIS-ASI imagery queried with the Aurora-X tool, to construct a large, featurized database of conjunctions which will allow us to study in-situ and visible arc statistics and develop ML-based predictive tools that effectively use the available, heterogeneous data.

These tools and analyses allow us to address several open questions in the field. Featurized conjugate in-situ and imagery data can be used to examine the statistics of visible discrete arcs and their associated in-situ current sheet and flow structures. More abstractly, the relationship between flow and FAC (with boundary conditions partially informed by imagery) can give information on when the magnetosphere acts more as a current or voltage source, at the time and length scales of discrete arcs. In the future, the problem of predicting in-situ ionospheric conditions from imagery data ("reading the aurora") can be attacked.

Outline:

We are collecting and featurizing data from conjunctions between THEMIS-ASI/DASC imagers and Swarm satellites, to use for statistical and ML studies.

Left Column: Background and Context
- Background
- Conjunction Database

Middle Column: Tools
- Data featurization, wavelet transform
- ASI Arc featurization and tracking
- The relationship between B, E, and conductance

Right Column: Results
- Predicting conductance profile across arc from Swarm B/E data
- Signatures of Alfven reflection
- Planned studies on these data

Scientific Background

Auroral current continuity, though it can be greatly simplified in situations with high symmetry, is a complex system that is controlled by the small scale behavior of Magnetosphere-Ionosphere (M-I) coupling, which is an area of active research. Investigations into the details of how current continuity is satisfied in a given auroral event can therefore provide information on the nature of M-I coupling and the dynamics of the auroral ionosphere.

We are using in-situ data from the Swarm satellites, white light imagery data from the THEMIS-ASI imagers, and color imagery from the Poker Flat DASC imagers to investigate this problem.

Conjunction Database:

As part of collecting these data, we are forming a publicly available database of conjunctions for use in other studies. The database will include, for each valid crossing:

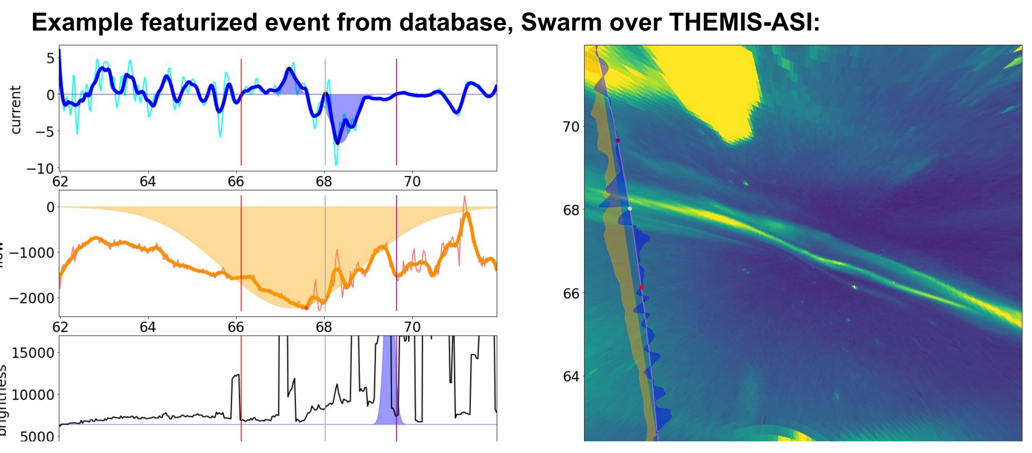
- The **raw and featurized in-situ data cuts**
- **Brightness cut**, for white light, or **RGB cuts** for color imagery
- **Full image frame** and skymap
- Attilla Danko (ClearDarkSky) predicted **cloud cover index**
- Moon marker, red 'X' on ASI frame
- In progress: RGB along-cut keogram, B/E derived product

Currently 2013-2020, ≈13000 conjunctions with ≈750 manually flagged as discrete arcs. 72 conjunctions with full color Poker DASC imagery.

Scan this to see PDFs of conjunctions in the database

We hope to expand this to include other quantities, like number of arcs or arc proper motion, but the full image featurization is still in development.

Example featured event from database, Swarm over THEMIS-ASI:

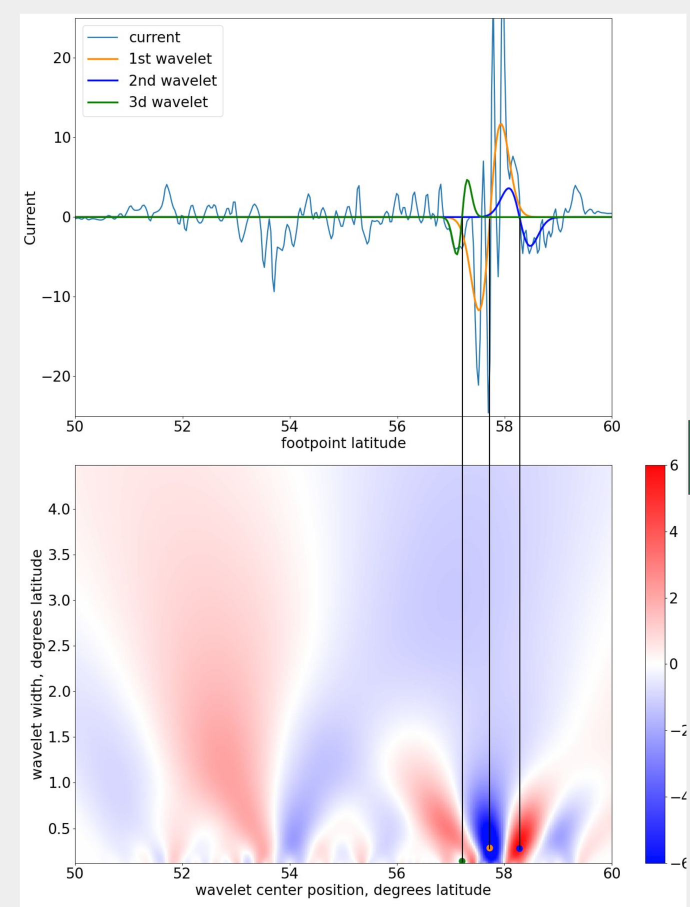


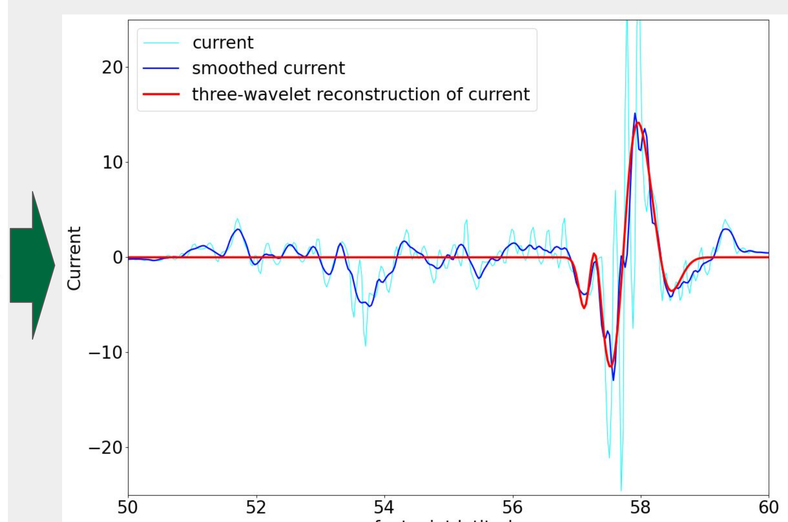
Featurization Tool 1: In-situ Cuts

- Each FAC or flow data cut requires ~200 numbers to specify.
- There are at max ~3000 valid Swarm-THEMIS conjunctions per year, far fewer have visible discrete arcs. There are also only 72 three color DASC conjunctions
- If we want to train a predictive model, high dimensional input + few examples leads to very poor training, so dimensionality must be reduced. Pulling out relevant features also enables more effective statistical analysis

Heuristic Gaussian Fits:
- One gaussian curve is fit to each of precipitating current, return current, and flow.
- Peaks in each quantity are automatically chosen by a heuristic algorithm defined by analogy to topographic prominence

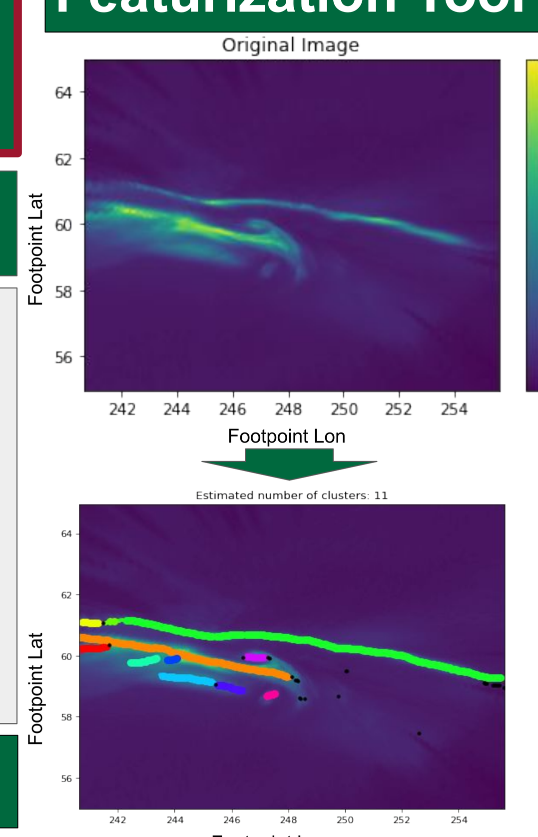
Continuous Wavelet Transform (CWT)
- The CWT allows the examination of structures localized in both frequency and time by projecting a data cut onto an overcomplete basis of wavelets.
- We use derivative-of-gaussian wavelets to featurize structures that are expected to be bipolar (FAC) and gaussian wavelets to featurize cuts that may have unipolar structures (plasma flow)
- **By successively projecting and removing the most significant wavelet components, we can economically represent the cut with as few wavelets as possible.**





This 200-dimensional FAC cut is reduced to a 9-dimensional (3x amplitude,center,width) wavelet representation!

Featurization Tool 2: ASI Arc Recognition and Tracking



Featurization of the ASI Image:

- Scale-sensitive "ridgelikeness" metric based on Frangi vesselness filter [5] used to highlight potential discrete arcs
- Tops of "second derivative ridges" are identified, as in the Shadden lab tutorial on FTLEs
- A local linear fit is found at every point, and used to create an arc aligned coordinate system.
- With a weight applied to the cross-arc direction, DBSCAN [4] is used to cluster the arc points and separate them into arcs, for featurization and tracking.

Scan this to see movies of arc recognition and tracking

Science Tool: Quasistatic 1d Current Continuity in a Sheetlike Arc, Solving For Conductance

Start with the height integrated (thin-ionosphere) current continuity equation from Marghitu [1]:

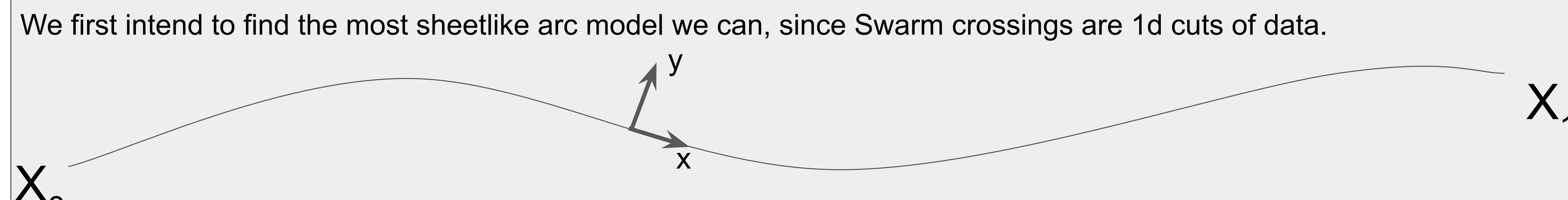
$$\nabla \cdot \mathbf{J}_{\perp} = j_{//}$$

Take **B** to point along the earth's magnetic field and call that direction z:

$$\nabla \cdot (\Sigma_P \mathbf{E} + \Sigma_H \mathbf{e}_z \times \mathbf{E}) = j_{//}$$

$$\Sigma_P \nabla \cdot \mathbf{E} + (\nabla \Sigma_P) \cdot \mathbf{E} + \Sigma_H \nabla \cdot (\mathbf{e}_z \times \mathbf{E}) + (\nabla \Sigma_P) \cdot (\mathbf{e}_z \times \mathbf{E}) = ((\nabla \times \mathbf{B}_{\perp})/\mu_0) \cdot \mathbf{e}_z$$

We first intend to find the most sheetlike arc model we can, since Swarm crossings are 1d cuts of data.



If we can find a sheetlike section of the arc parameterized by x with minimal along-x variation, the equations simplify. Letting $\mathbf{e}_x \times \mathbf{e}_y = \mathbf{e}_z$, we approximate that all quantities vary only along y, as in Marghitu [1]. An **E** pointing along x may also polarize the arc and induce secondary currents via the Cowling effect, so we neglect it for simplicity. Then we get:

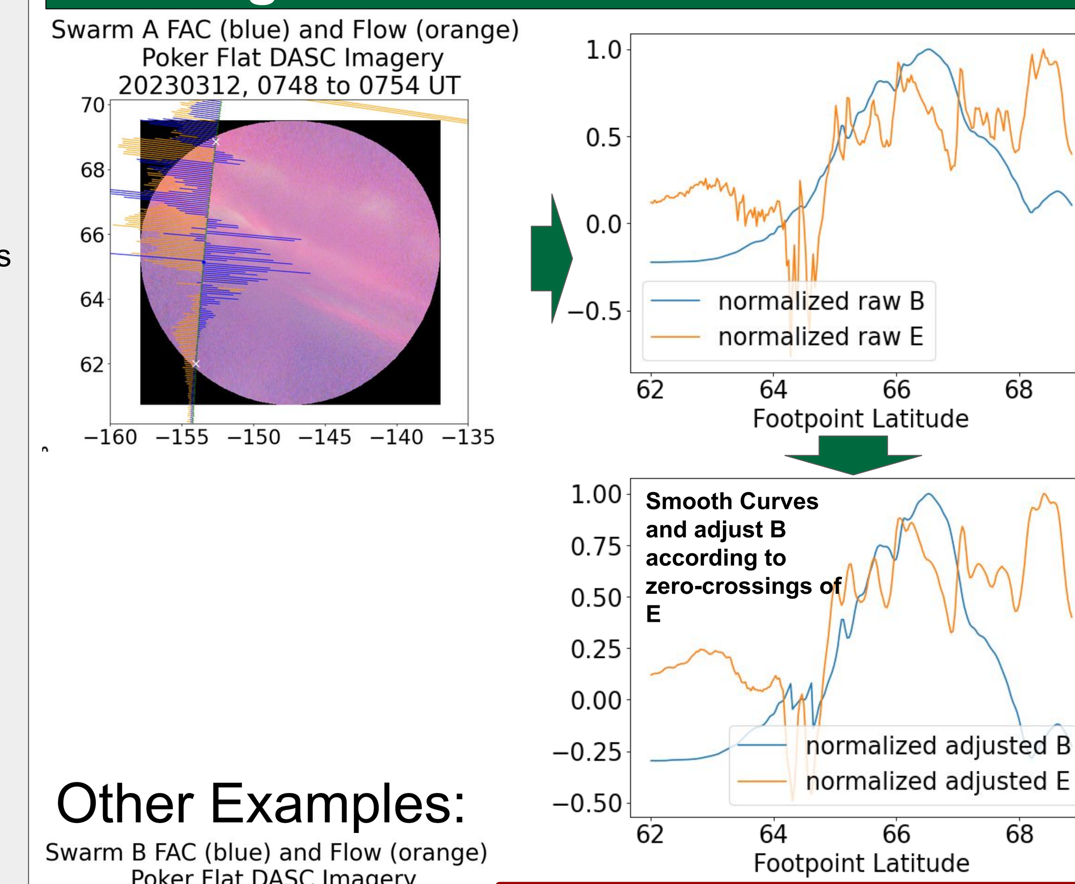
$$\Sigma_P(\partial_y E_y) + E_y(\partial_y \Sigma_P) + 0 + 0 = \partial_y B_x/\mu_0$$

$$\partial_y(E_y \Sigma_P) = \partial_y(B_x/\mu_0)$$

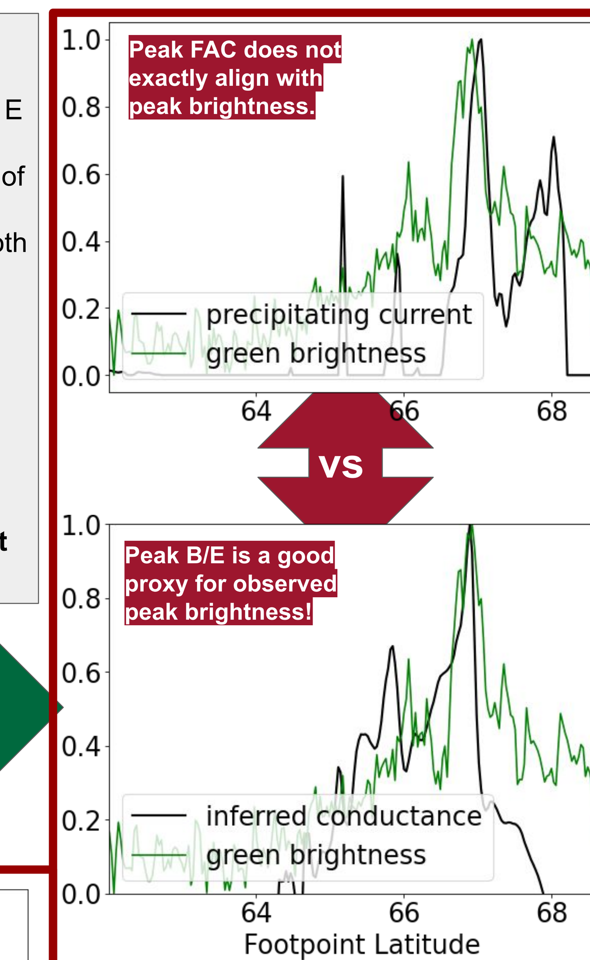
If we find a point y_0 where the electric field crosses zero, the solution takes a particularly simple form:

$$\Sigma_P(y) = (1/\mu_0)([B_x(y) - B_x(y_0)] / E_y(y))$$

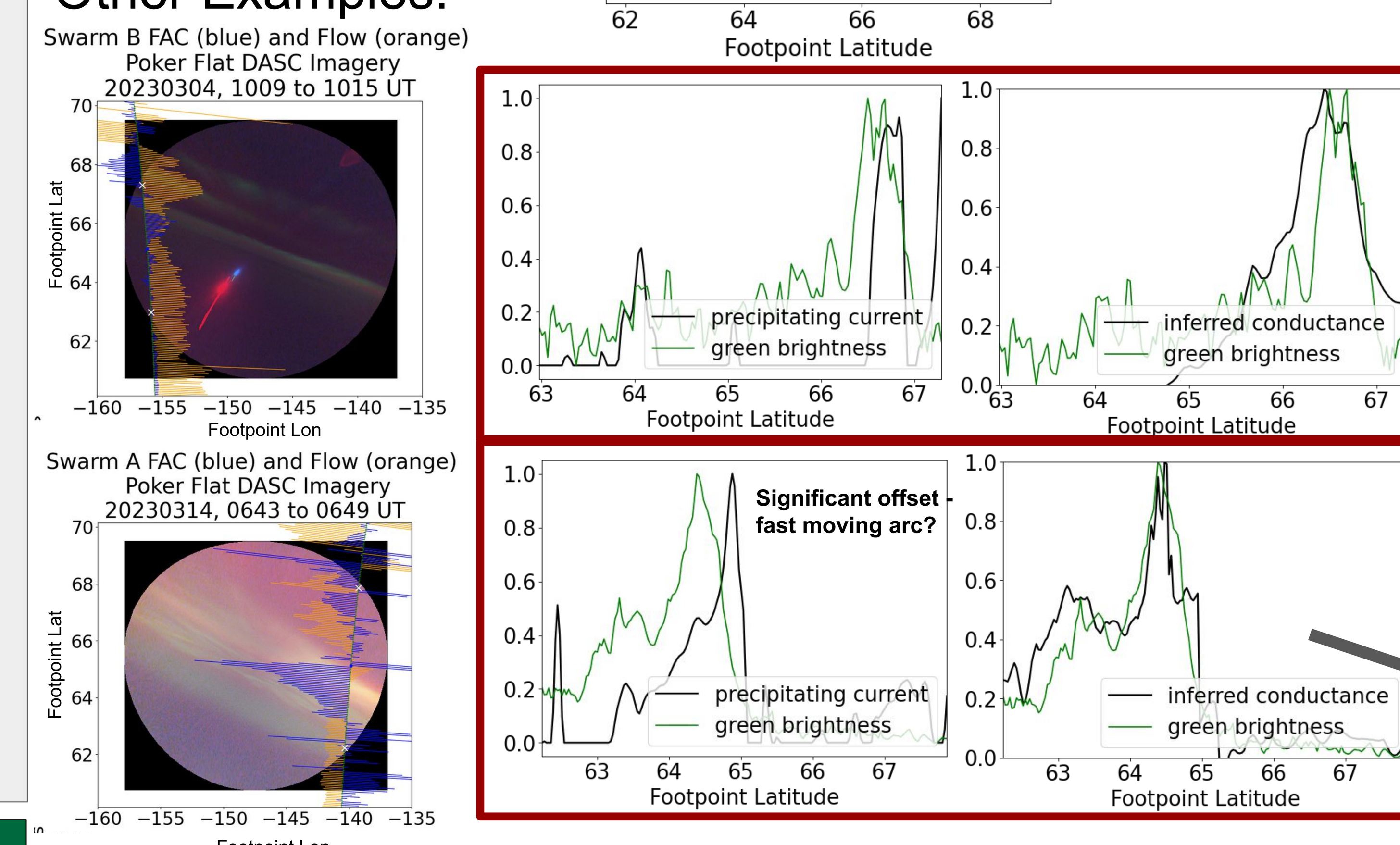
Inferring Small-Scale Conductance Patterns from Swarm Data Cuts



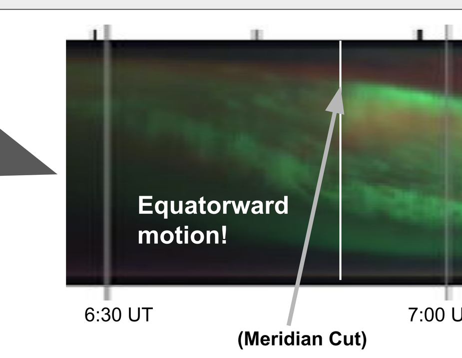
Solving Current Continuity
- Any zero crossing of E determines the boundary conditions of current continuity,
- In practice, due to both theoretical and experimental deviations, multiple zero crossings suggest different boundary conditions for the same cut
- **We use the closest zero crossing to set offsets to B locally!**



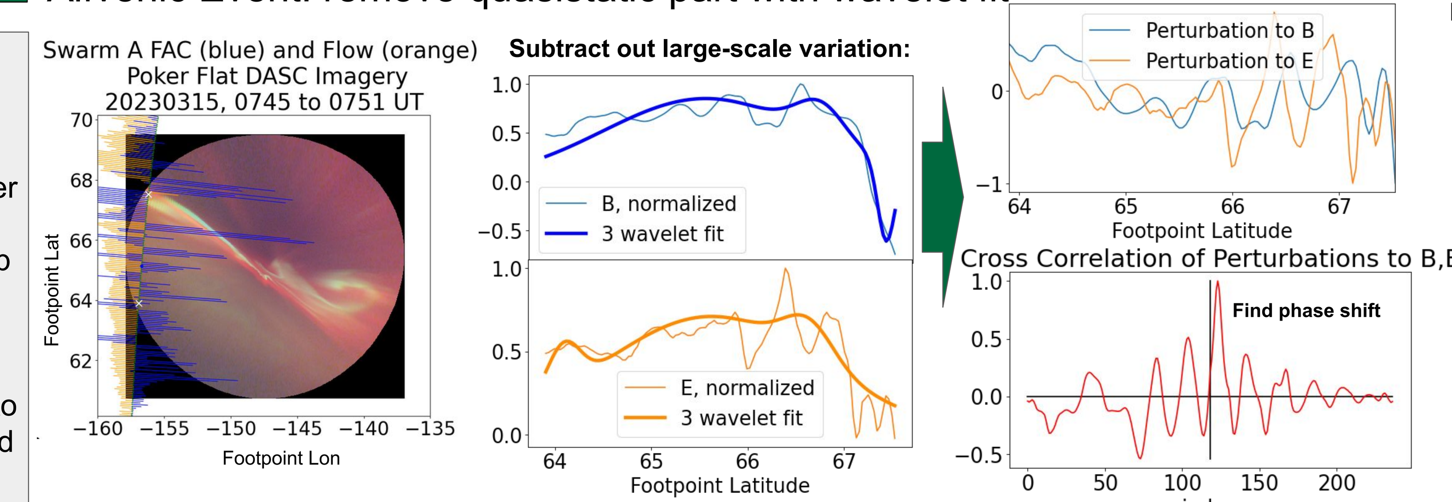
Other Examples:

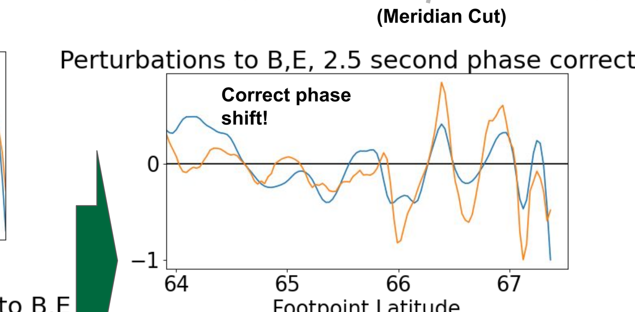


- **Peak visible brightness** appears to **more closely follow peak inferred conductance than peak precipitating current**
- **Brightness/conductance is often equatorward of current**, suggesting peak current may lag behind a moving band of precipitation



Alfvenic Event: remove quasistatic part with wavelet fit





- Phase shift is signature of Alfven reflection
- B/E ratios, Alfven conductance can now be studied!

Future work:

- The (n=3)-wavelet representation is a useful, highly nonlinear way of reducing the dimensionality of data cuts. **Through unsupervised ML (clustering), we can attempt to find a quantitative classification scheme of arc morphology, extending the qualitative work of Marklund [2].**
- With supervised ML, **we can test the predictive power of the wavelet coefficients towards other physical quantities**. In either case, the usefulness of the successive projection wavelet coefficients can be compared against other means of dimensionality reduction, both linear (PCA) and nonlinear (autoencoders, etc).
- We intend to **incorporate time dependence of the aurora into the database** with the use of ASI Image featurization
- Due to the wealth of available Swarm data when not limited by conjunctions, we will **consider starting an in-situ alone study** and new database of Swarm measurements
- Collaborations with future users of this database (Swarm over Poker + Swarm over THEMIS-ASI)
- Working with Jules Van Irsel on using in-situ data to provide physically realistic drivers to the GEMINI ionospheric model

Acknowledgements:

- Maia Kawamura and Kristina Lynch, for starting this project before I arrived at Dartmouth
- Lynette Gelinas, for inspiration and reading material regarding wavelet transforms
- Jules Van Irsel, for constructive discussion and for running the simulation side of this project
- Undergraduate students Maia Kawamura, Grace Connolly, Meggie Bond, Katherine Pommerening, and Anaisa Rowe for their assistance in the project
- Attilla Danko's generous sharing of cloud cover prediction data for GBO sites
- Darren Chaddock, assistance in using Skyfield to locate the moon in ASI images
- ESA Swarm data: https://earth.esa.int/eogateway/missions/swarm/data
- University of Calgary - AuroraX Conjunction finder: https://aurorax.space
- University of Calgary - Themis ASI data: https://www.ucalgary.ca/aurora/projects/themis
- University of Alaska Fairbanks - Poker Flat DASC data: http://optics.gi.alaska.edu/optics/archive

References:

- 1. Marghitu, Octav. "Auroral arc electrodynamics: Review and outlook." *Relationship between auroral phenomenology and magnetospheric processes: Earth and other planets, edited by: Keiling, A., Donovan, E., Bagenal, F., and Karlsson, T., Geophysical Monograph* 197 (2012): 143-158.
- 2. Marklund,Göran. "Auroral arc classification scheme based on the observed arc-associated electric field pattern." *Planetary and Space Science* 32.2 (1984): 193-211
- 3. Ester, M., Kriegel, H.-P., Sander, J., Xu, X., & others. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In kdd (Vol. 96, pp. 226–231).
- 4. Shadden, Shawn C., Francois Lekien, and Jerrold E. Marsden. "Definition and properties of Lagrangian coherent structures from finite-time Lyapunov exponents in two-dimensional aperiodic flows." *Physica D: Nonlinear Phenomena* 212.3-4 (2005): 271-304.
- 5. Frangi, Alejandro F., et al. "Multiscale vessel enhancement filtering." *International conference on medical image computing and computer-assisted intervention*. Springer, Berlin, Heidelberg, 1998.