Assessing Our Predictive Understanding of Geospace Physics at GEM

Katherine Garcia-Sage
NASA GSFC

Thanks to:
GEM Modeling Methods and Validation co-chairs:
Lutz Rastaetter, Mike Liemohn, Rob Redmon
Community Coordinated Modeling Center
iCCMC-LWS Tracking Progress team led by Alexa Halford
iCCMC-LWS Auroral Assessment team led by Bob Robinson
and the whole GEM community

What is GEM?

“Geospace Environment Modeling (GEM) is a broad-based, community-initiated research program on the ***physics*** of the Earth's magnetosphere and the ***coupling*** of the magnetosphere to the atmosphere and to the solar wind. The purpose of the GEM program is to support ***basic research*** into the dynamical and structural properties of geospace, leading to the ***construction*** of a global Geospace General Circulation ***Model*** (GGCM) with ***predictive capability***. This GGCM model will be ***modularized*** and will complement parallel developments of magnetohydrodynamic models. The strategy for achieving GEM goals is to undertake a series of campaigns and ***focus groups***, in both ***theory and observation***al modes, each focusing on particular aspects of the geospace environment.”
-from GEM wiki, emphasis mine

What is GEM?

physics

coupling

basic research

construction

Model

predictive capability

modularized

focus groups

theory and observation

GEM involves ***communication between different regimes*** for ***predictive understanding***.

What is GEM?

coupling
construction
modularized
focus groups
theory and observation

How? ***Validation!***

physics
basic research

Model
predictive capability

GEM involves ***communication between different regimes*** for ***predictive understanding***.

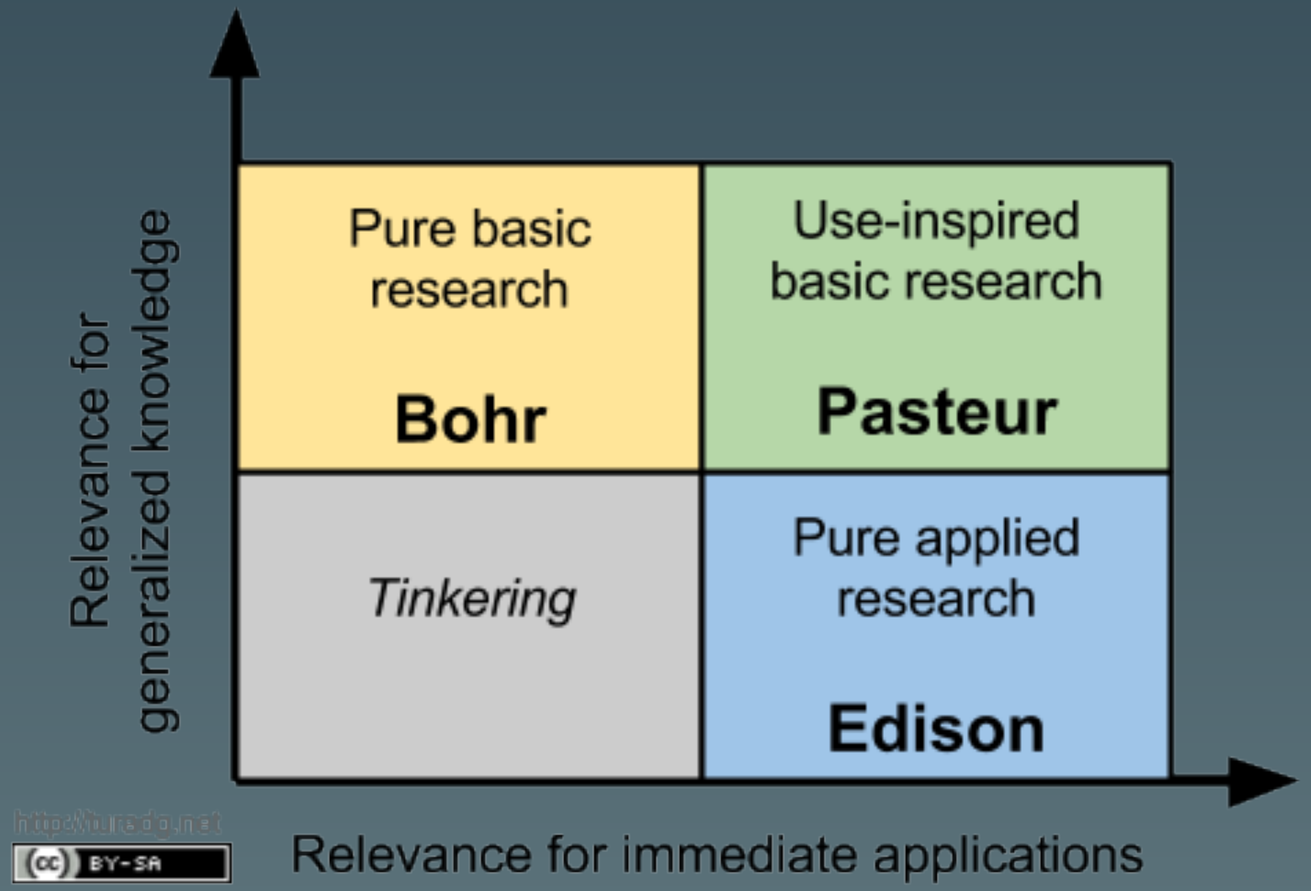
Is this about space weather?

Yes, and more

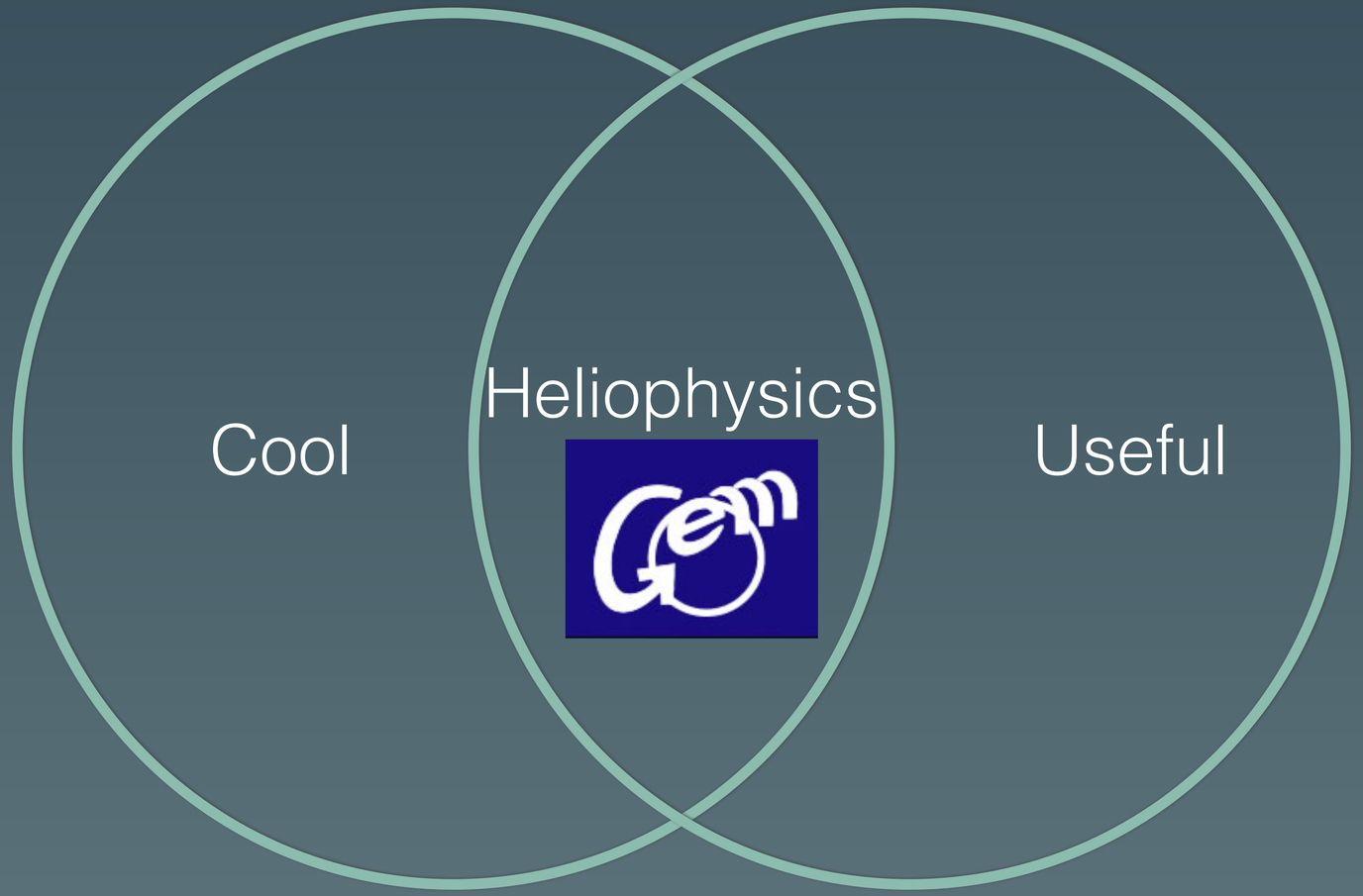
How do I know whether this model, theory, or data is *valid* for this application?

Validation requires context, application, and reproducibility.

Pasteur's Quadrant



An alternate view



Outline

Photo by Martin Adams on Unsplash

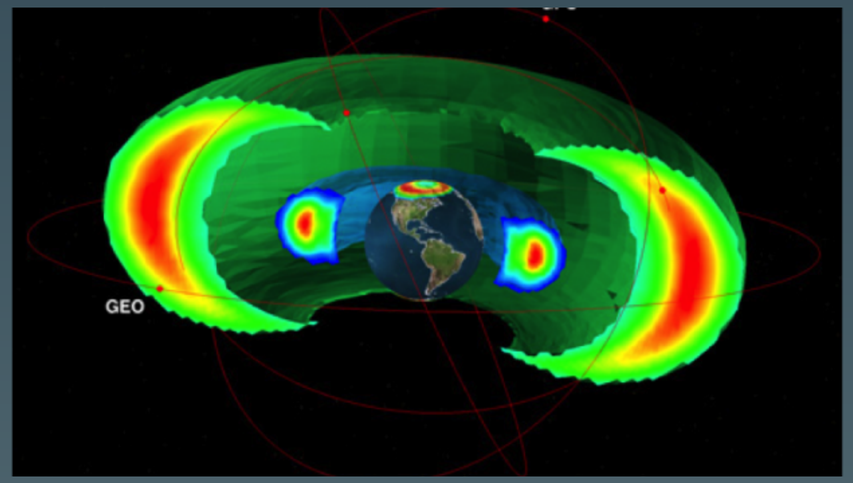
- Applications in our field
- Asking the validation question
- Validation processes and pathways
- Validation tools
- Validation needs, challenges, and opportunities
- Modeling Methods and Validation Focus Group

Applications for Geospace Modeling (a.k.a. Who Cares About Validation?)

- Space Weather operational users
- ITM scientists
- Geospace data and instrument scientists
- Other Geospace modelers
- Planetary Magnetospheres
- Exoplanetary Magnetospheres

Applications for Geospace Modeling (a.k.a. Who Cares About Validation?)

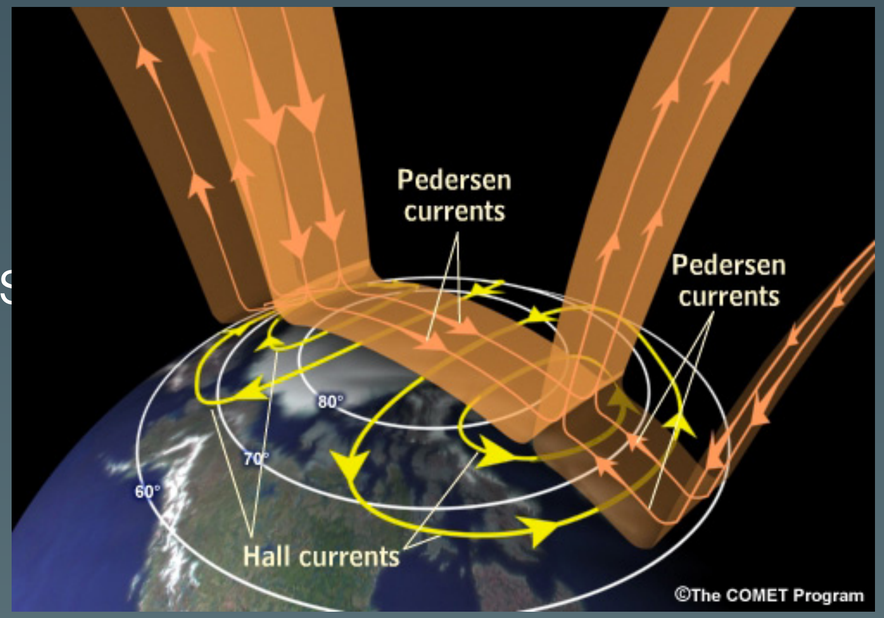
- Space Weather operational users
- ITM scientists
- Geospace data and instrument scientists
- Other Geospace modelers
- Planetary Magnetospheres
- Exoplanetary Magnetospheres



from NOAA website

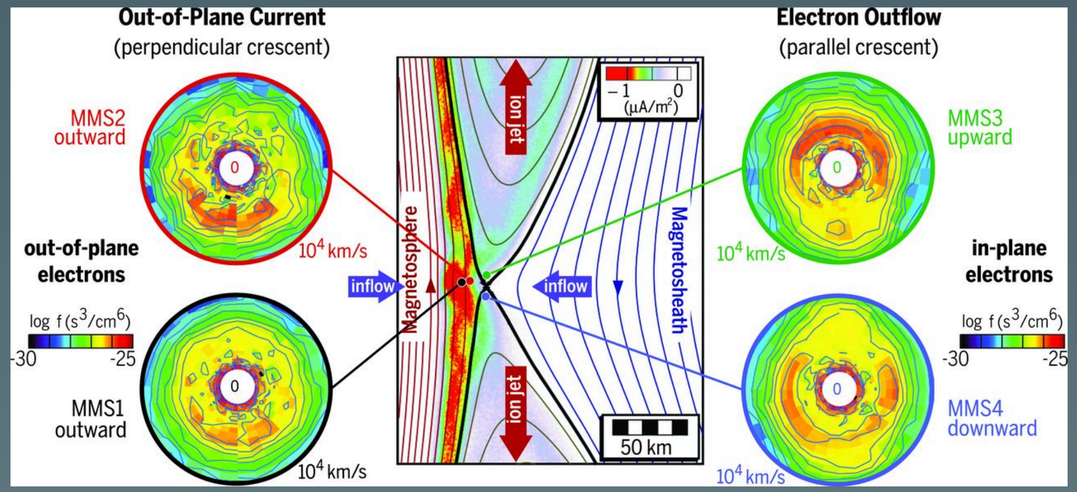
Applications for Geospace Modeling (a.k.a. Who Cares About Validation?)

- Space Weather operational users
- ITM scientists
- Geospace data and instrument scientis
- Other Geospace modelers
- Planetary Magnetospheres
- Exoplanetary Magnetospheres



Applications for Geospace Modeling (a.k.a. Who Cares About Validation?)

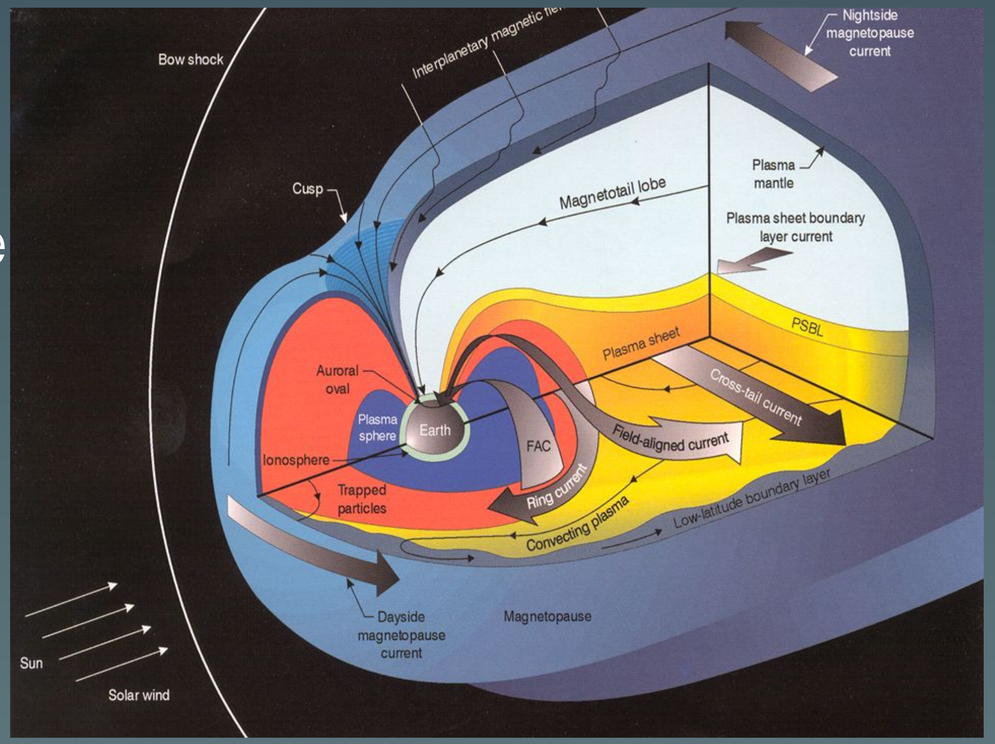
- Space Weather operational users
- ITM scientists
- Geospace data and instrument scientists
- Other Geospace modelers
- Planetary Magnetospheres
- Exoplanetary Magnetospheres



Burch et al., 2016

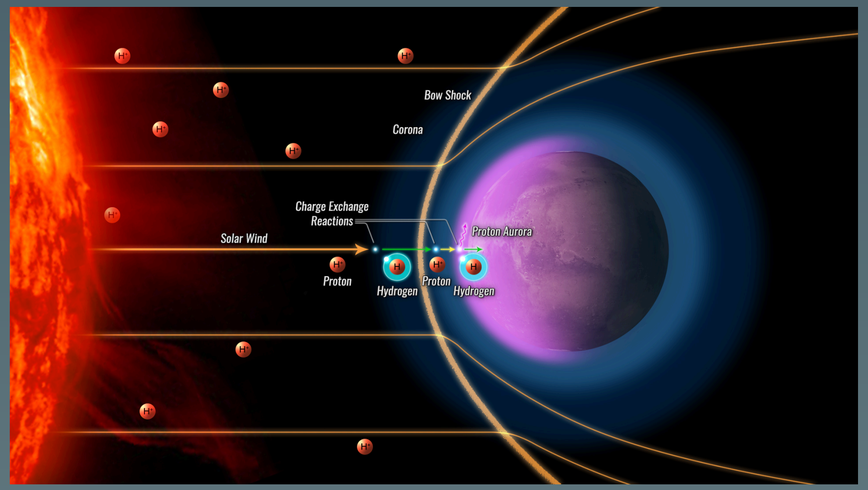
Applications for Geospace Modeling (a.k.a. Who Cares About Validation?)

- Space Weather operational users
- ITM scientists
- Geospace data and instrument scie
- Other Geospace modelers
- Planetary Magnetospheres
- Exoplanetary Magnetospheres



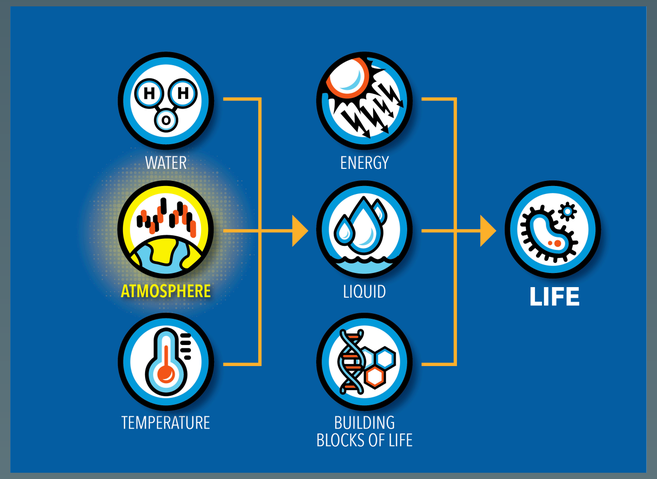
Applications for Geospace Modeling (a.k.a. Who Cares About Validation?)

- Space Weather operational users
- ITM scientists
- Geospace data and instrument scientists
- Other Geospace modelers
- Planetary Magnetospheres
- Exoplanetary Magnetospheres



Applications for Geospace Modeling (a.k.a. Who Cares About Validation?)

- Space Weather operational users
- ITM scientists
- Geospace data and instrument scientists
- Other Geospace modelers
- Planetary Magnetospheres
- Exoplanetary Magnetospheres



Validation Drives Progress

- Improve our understanding of universal physics by pushing our model capabilities past initial assumptions

- Use this improved understanding to improve models

Asking the validation question

How do I know whether this model, theory, or data is *valid* for this application?

GEM: How do we apply models, theories, and data to each other for a predictive understanding of geospace?

Asking the validation question

How do I know whether this model, theory, or data is ***valiɔ*** for this application?

CCMC: How can heliophysics models reach their broadest *appropriate* application?

Asking the validation question

How do I know whether this model, theory, or data is *valid* for this application?

NOAA: How can heliophysics models be applied to predict space weather human and technological impacts?

What is your validation question?

How do I know whether this model, theory, or data is *valid* for this application?

What model can be applied to provide context for my data?

Can this data be applied to refine my model uncertainties?

How can my theory be applied to the needs of modelers?

Outline

Photo by Martin Adams on Unsplash

- Applications in our field
- Asking the validation question
- Validation processes and pathways
- Validation tools
- Validation needs, challenges, and opportunities
- Modeling Methods and Validation Focus Group

Processes and pathways: terminology

- Validation vs. Verification (from Roache, 1998):
  - *Verification* checks that you are solving the equations right.
  - *Validation* checks that you are solving the right equations.
- *Metrics* - a quantitative measure of performance [Halford et al., in press]
- "*Metrics* and *validation* are both important and are complementary: metrics provide a quantitative measure of the improving predictive strength of a model, while scientific validation systematically identifies model shortcomings and strengths and informs ongoing model improvement." [Spence et al., 2004]

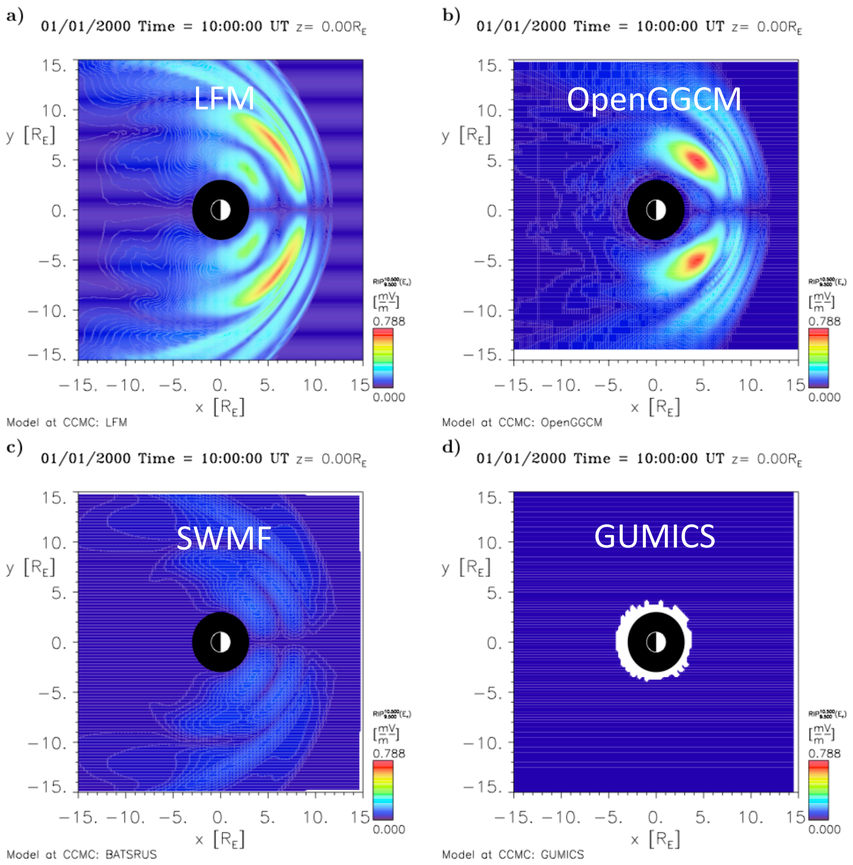
What’s missing?

- What assumptions are in the model and/or data that we are using?
- Validation and metrics are about accuracy with reproducibility.
- Reproducible results require **clear** assumptions.

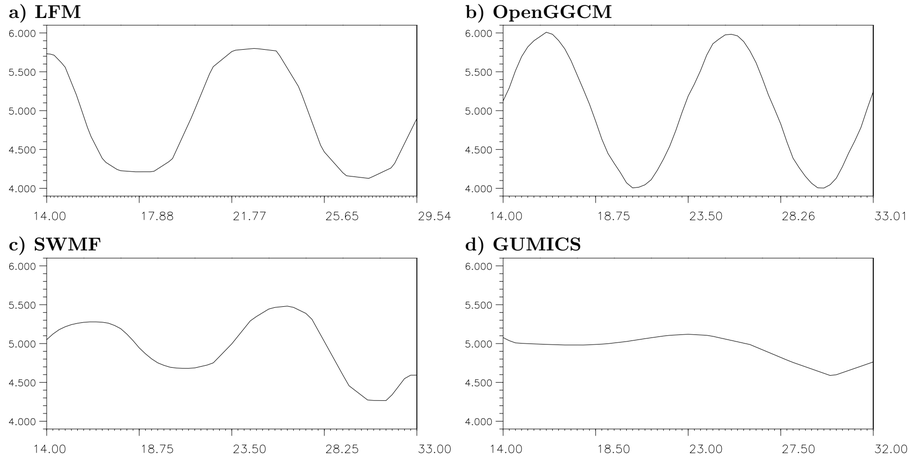
ULF wave modeling challenge

ULF wave modeling at CCMC

Power at 10 mHz at Z=0



10 mHz solar wind signal between X=33 and X=14
Driver at 33 R_E: 4/cm^3 to 6/cm^3



Signal at X=14

Model	$max(N)$	$min(N)$	Atten. Ratio	$\max\left(RIP(E_r)_{9.5}^{10.5}\right)$ constant conductances	auroral conductances
LFM	5.749	4.221	0.764	0.501	0.519
OpenGGCM	6.010	3.990	1.010	0.788	0.959
SWMF	5.269	4.726	0.272	0.0945	0.1923
GUMICS	5.098	4.983	0.0577	0.00410	0.00402

Grid resolution upstream of bow shock matters!

Conductance model matters!

Lutz Rastaetter, et al. (2019 in prep.)

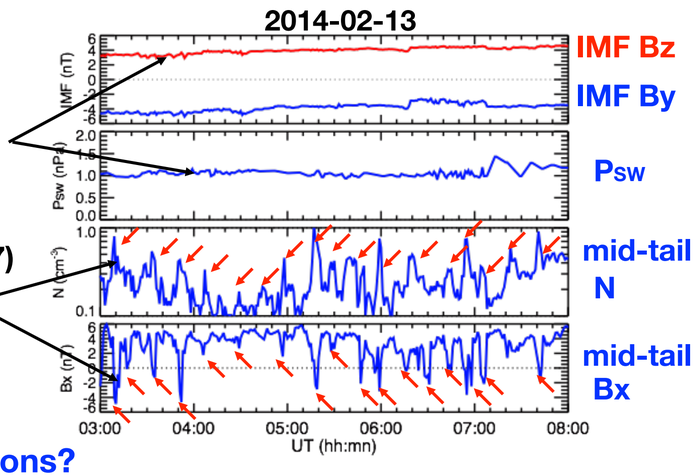
Mid-tail modeling challenge

Mid-Tail (X ~ –60 R_E) Modeling Challenge

(Tail Environment and Dynamics at Lunar Distances)

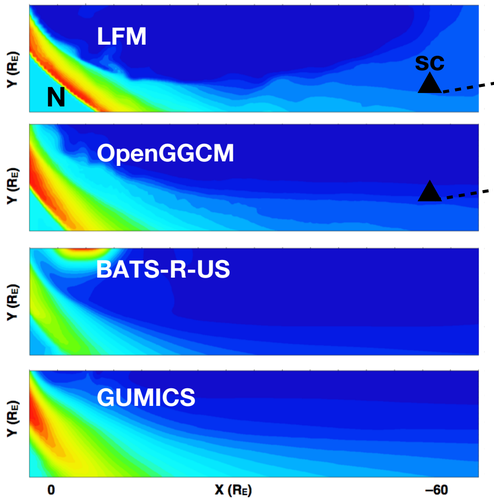
~5 hr nearly constant northward IMF and solar wind

Two ARTEMIS in mid-tail lobe (X ~ –58 and 61, Y ~17) saw large mesoscale (10-30 min, a few R_E) fluctuations in plasma and B (N increase and B) reduced)

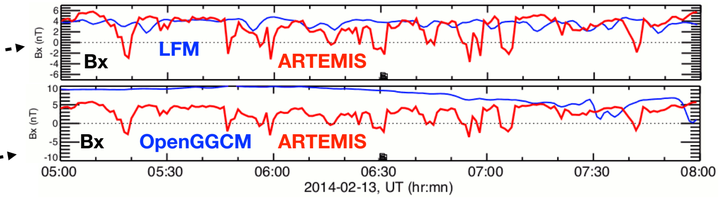


What process causes these mesoscale perturbations?

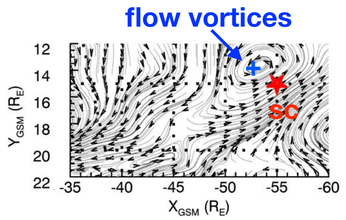
Event simulations

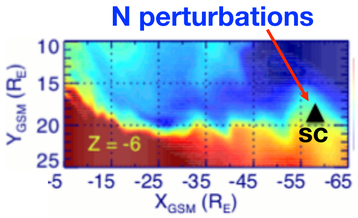


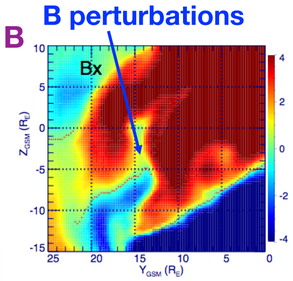
mid-tail mesoscale perturbations only seen in LFM



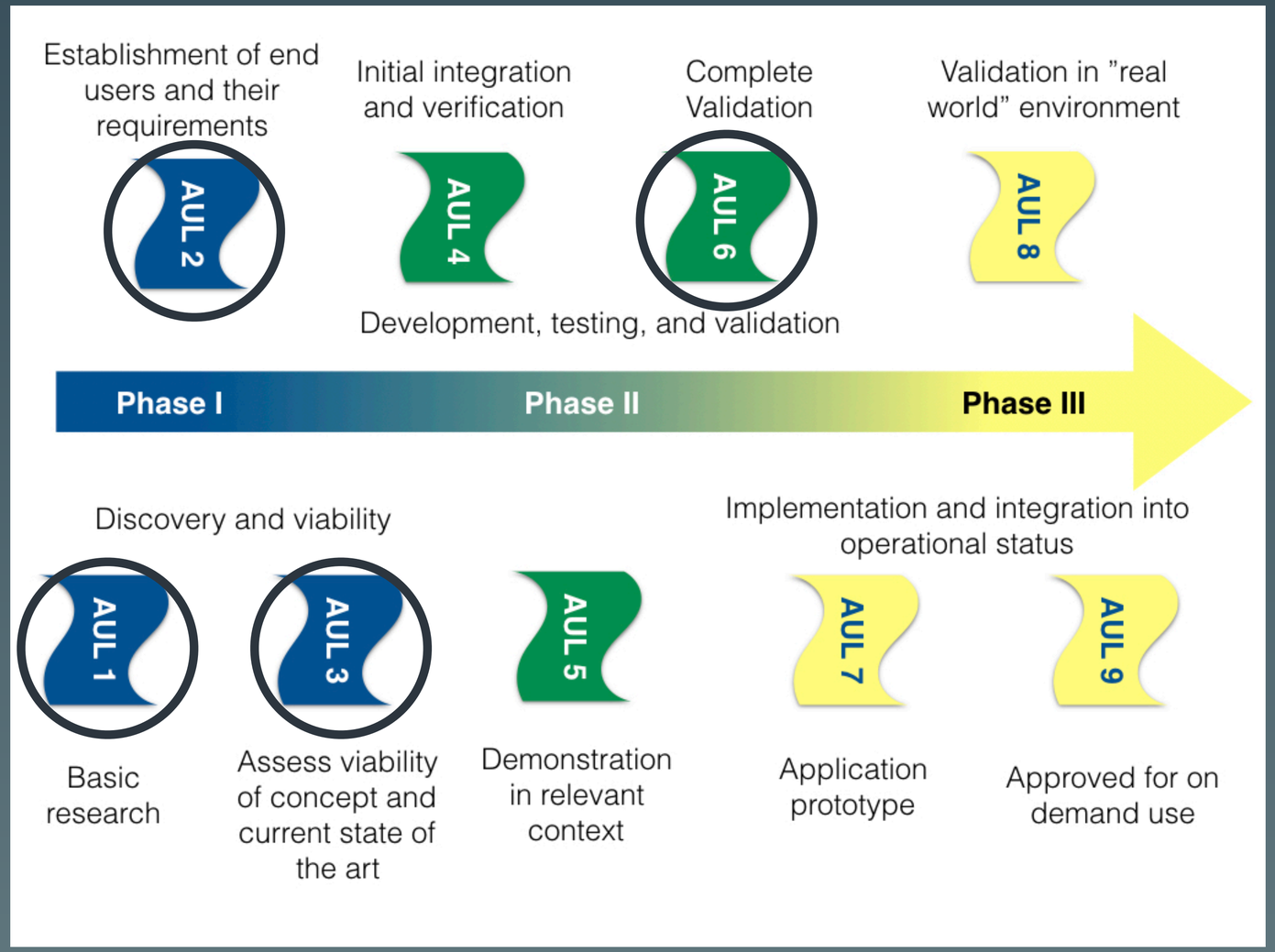
perturbations due to:
tailward propagation of K-H flow vortices –>
high N plasma penetrate deep to lobe –> reduce B







Processes and pathways: Application Usability Levels

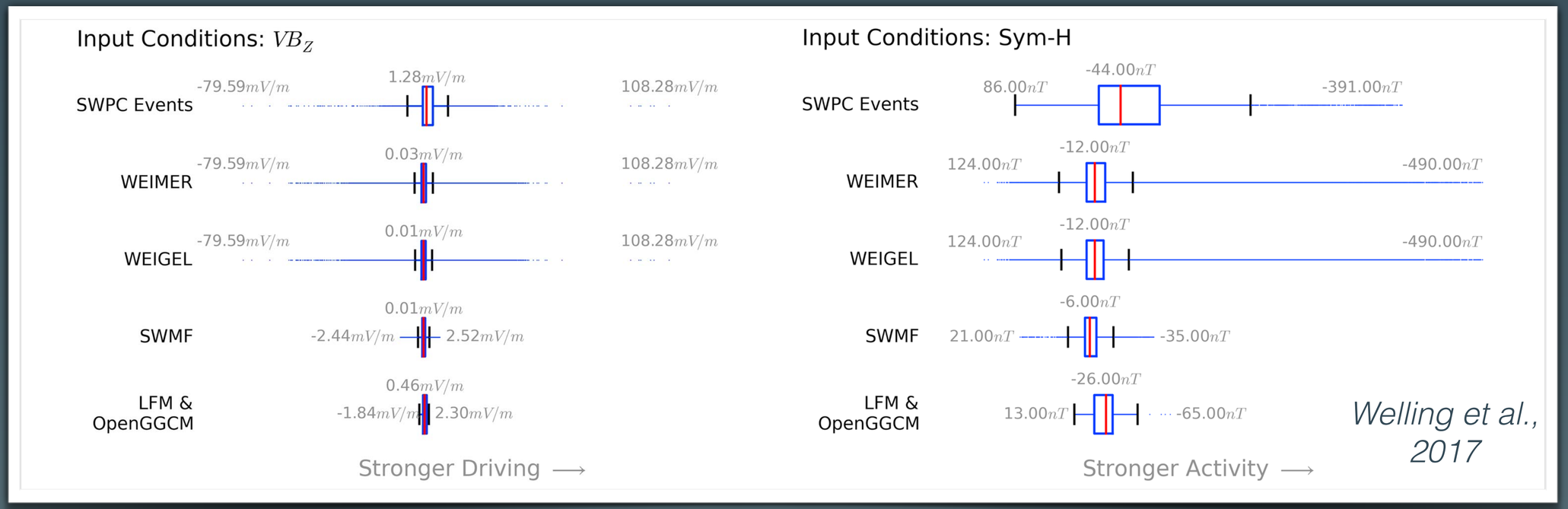


Halford et al., JSWSC, in press

Processes and pathways: data uses

- Training/development data
- Testing data
- Validation data
- “Operational” data

Processes and pathways: data uses



- Welling et al. reanalyzed the results of the GEM Ground Magnetic Perturbations Challenge
- Validation events included a range of driving conditions
- Validation events include a much broader range of drivers and geomagnetic conditions compared to the events used to construct the conductance models of global MHD codes

Processes and pathways: choosing metrics

Space Weather

Space Weather publishes original research articles and commentaries devoted to understanding and forecasting space weather and other interactions of solar processes with the Earth environment, and their impacts on telecommunications, electric power, satellite navigation, and other systems.

HOME | BROWSE | HIGHLIGHTS | COMMENTARIES | COLLECTIONS | ABOUT

Space Weather Capabilities Assessment

Special Issues | First published: **1 April 2018** | Last updated: **27 August 2018**

This special section highlights the progress of working teams of the International Forum on Space Weather Capabilities Assessment. The Forum activities were initiated during the International CCMC - LWS Working Meeting: “Assessing Space Weather Understanding and Applications”, April 3 - 7, 2017, Cape Canaveral, Florida. The Forum addressed the need to quantify and to track progress over time in the field of space weather and to establish internationally recognized metrics for objective model evaluations. Such metrics must be meaningful to end-users of space weather information, model developers, and decision makers. Expected outcomes of the Forum activities included developing means to aid in tracking the progress of the LWS program towards its goals. Topics covered by the special collection include: defining metrics for essential space weather quantities; benchmarking the current state of space environment models, applications and forecasting techniques; addressing challenges in data-model comparisons; tracking progress in incorporation of scientific ideas into space weather applications. While the focus of this special collection is contributions from the Forum working teams, related manuscripts from the community were also invited.

Processes and pathways: choosing metrics

AGU PUBLICATIONS

Space Weather

RESEARCH ARTICLE

10.1002/2017SW001669

Measures of Model Performance Based On the Log Accuracy Ratio

S. K. Morley[1], T. V. Brito[1,2], and D. T. Welling[3]

[1]Space Science and Applications, Los Alamos National Laboratory, Los Alamos, NM, USA
University of Helsinki, Helsinki, Finland, [3]Climate and Space Sciences and Engineering D
Michigan, Ann Arbor, MI, USA

Key Points:

- The median symmetric accuracy and symmetric signed percentage bias are introduced to address some drawbacks of current metrics
- The spread of a multiplicative linear model can be robustly estimated using the log accuracy ratio

Table 1
A Summary of Key Metrics

Metric	Definition	Symmetry	Scale/Order dependent	Comments
			Error metrics	
ε	$y - x$	Y	Scale	Forecast error
Q	y/x	N	Order	Ratio; complement of forecast relative error
			Accuracy metrics	
MSE	$\left(\frac{1}{n}\sum_{i=1}^{n}\varepsilon_i^2\right)$	Y	Scale	Different units/scale; quadratic penalty
RMSE	$\sqrt{\left(\frac{1}{n}\sum_{i=1}^{n}\varepsilon_i^2\right)}$	Y	Scale	Same units as x, y; quadratic penalty
MAE	$\frac{1}{n}\sum_{i=1}^{n}\lvert\varepsilon_i\rvert$	Y	Scale	Same units as x, y; linear penalty
MdAE	$M\lvert\varepsilon_i\rvert$	Y	Scale	Same units as x, y; linear penalty; robust and resistant
MAPE	$\frac{100}{n}\sum_{i=1}^{n}\left\lvert\frac{\varepsilon_i}{x_i}\right\rvert$	N	Order	Percentage; penalizes overprediction more heavily
sMAPE	$100\,\frac{1}{n}\sum_{i=1}^{n}\left\lvert\frac{y_i - x_i}{(x_i + y_i)/2}\right\rvert$	Y	Order	Percentage; unintuitive normalization; handles $x = 0$
ζ	$100\left(e^{(M(\lvert\log_e(Q_i)\rvert))} - 1\right)$	Y	Order	Percentage; robust and resistant
			Bias metrics	
ME	$\frac{1}{n}\sum_{i=1}^{n}\varepsilon_i$	Y	Scale	Same units as x, y
MPE	$\frac{100}{n}\sum_{i=1}^{n}\frac{\varepsilon_i}{x_i}$	N	Order	Percentage; penalizes overprediction more heavily
MdLQ	$M\log_e(Q_i)$	Y	Order	Different scale
SSPB	$100\,\mathrm{sgn}(\mathrm{MdLQ})\,(e^{(\lvert\mathrm{MdLQ}\rvert)} - 1)$	Y	Order	Percentage; robust and resistant

Note. The columns give, in order, the abbreviation or symbol of the metric (as used in the text), the definition, whether the penalty is symmetric, whether the metric is scale or order dependent, and selected attributes.

?

Processes and pathways: choosing metrics

Metric	Definition	Symmetry	Scale/Order dependent	Comments
ε	$y - x$	Y	Scale	Forecast error
Q	y/x	N	Order	Ratio; complement of forecast relative error
MAPE	$\frac{100}{n}\sum_{i=1}^{n}\left\lvert\frac{\varepsilon_i}{x_i}\right\rvert$	N	Order	Percentage; penalizes overprediction more heavily
ζ	$100\left(e^{(M(\lvert\log_e(Q_i)\rvert))} - 1\right)$	Y	Order	Percentage; robust and resistant

Processes and pathways: choosing metrics

Metric	Definition	Symmetry	Scale/Order dependent	Comments
ε	$y - x$	Y	Scale	Forecast error
Q	y/x	N	Order	Ratio; complement of forecast relative error
MAPE	$\frac{100}{n}\sum_{i=1}^{n}\left\lvert\frac{\varepsilon_i}{x_i}\right\rvert$	N	Order	Percentage; penalizes overprediction more heavily
ζ	$100\left(e^{(M(\lvert\log_e(Q_i)\rvert))} - 1\right)$	Y	Order	Percentage; robust and resistant

Mean Absolute Percentage Error

Processes and pathways: choosing metrics

Metric	Definition	Symmetry	Scale/Order dependent	Comments
ε	$y - x$	Y	Scale	Forecast error
Q	y/x	N	Order	Ratio; complement of forecast relative error
MAPE	$\frac{100}{n}\sum_{i=1}^{n}\left\lvert\frac{\varepsilon_i}{x_i}\right\rvert$	N	Order	Percentage; penalizes overprediction more heavily
ζ	$100\left(e^{(M(\lvert\log_e(Q_i)\rvert))} - 1\right)$	Y	Order	Percentage; robust and resistant

Median Symmetric Accuracy

Processes and pathways: choosing metrics

Metric	Definition	Symmetry	Scale/Order dependent	Comments
ε	$y - x$	Y	Scale	Forecast error
Q	y/x	N	Order	Ratio; complement of forecast relative error
MPE	$\frac{100}{n}\sum_{i=1}^{n}\frac{\varepsilon_i}{x_i}$	N	Order	Percentage; penalizes overprediction more heavily
SSPB	$100\ \mathrm{sgn}(\mathrm{MdLQ})\ (e^{(\lvert\mathrm{MdLQ}\rvert)} - 1)$	Y	Order	Percentage; robust and resistant

$$\mathrm{MdLQ} = M \log_e(Q_i)$$

Processes and pathways: choosing metrics

Metric	Definition	Symmetry	Scale/Order dependent	Comments
ε	$y - x$	Y	Scale	Forecast error
Q	y/x	N	Order	Ratio; complement of forecast relative error
MPE	$\frac{100}{n}\sum_{i=1}^{n}\frac{\varepsilon_i}{x_i}$	N	Order	Percentage; penalizes overprediction more heavily
SSPB	$100\ \mathrm{sgn}(\mathrm{MdLQ})\ (e^{(\lvert \mathrm{MdLQ} \rvert)} - 1)$	Y	Order	Percentage; robust and resistant

$$\mathrm{MdLQ} = M \log_e(Q_i)$$

Mean Percentage Error

Processes and pathways: choosing metrics

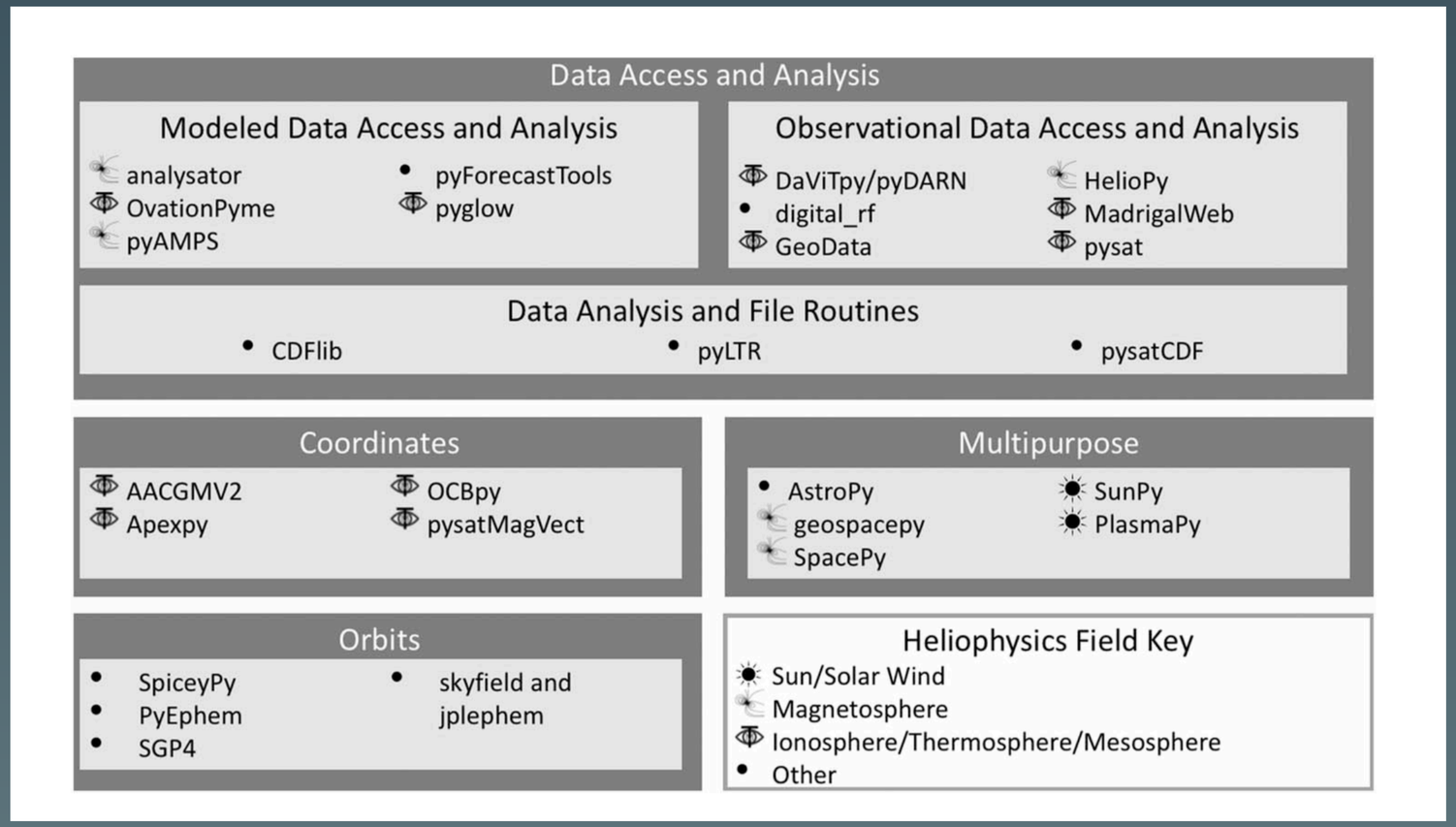
Metric	Definition	Symmetry	Scale/Order dependent	Comments
ε	$y - x$	Y	Scale	Forecast error
Q	y/x	N	Order	Ratio; complement of forecast relative error
MPE	$\frac{100}{n}\sum_{i=1}^{n}\frac{\varepsilon_i}{x_i}$	N	Order	Percentage; penalizes overprediction more heavily
SSPB	$100\ \mathrm{sgn}(\mathrm{MdLQ})\ (e^{(\lvert \mathrm{MdLQ} \rvert)} - 1)$	Y	Order	Percentage; robust and resistant

$$\mathrm{MdLQ} = M \log_e(Q_i)$$

Signed Symmetric Percentage Bias

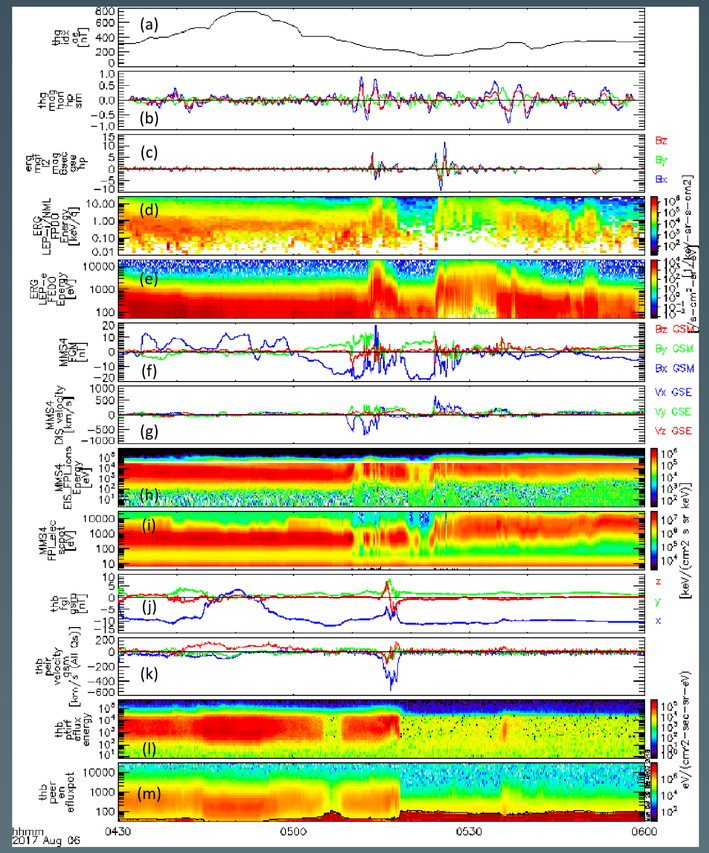
Validation tools

Overview of Python Packages



Burrell et al., 2018

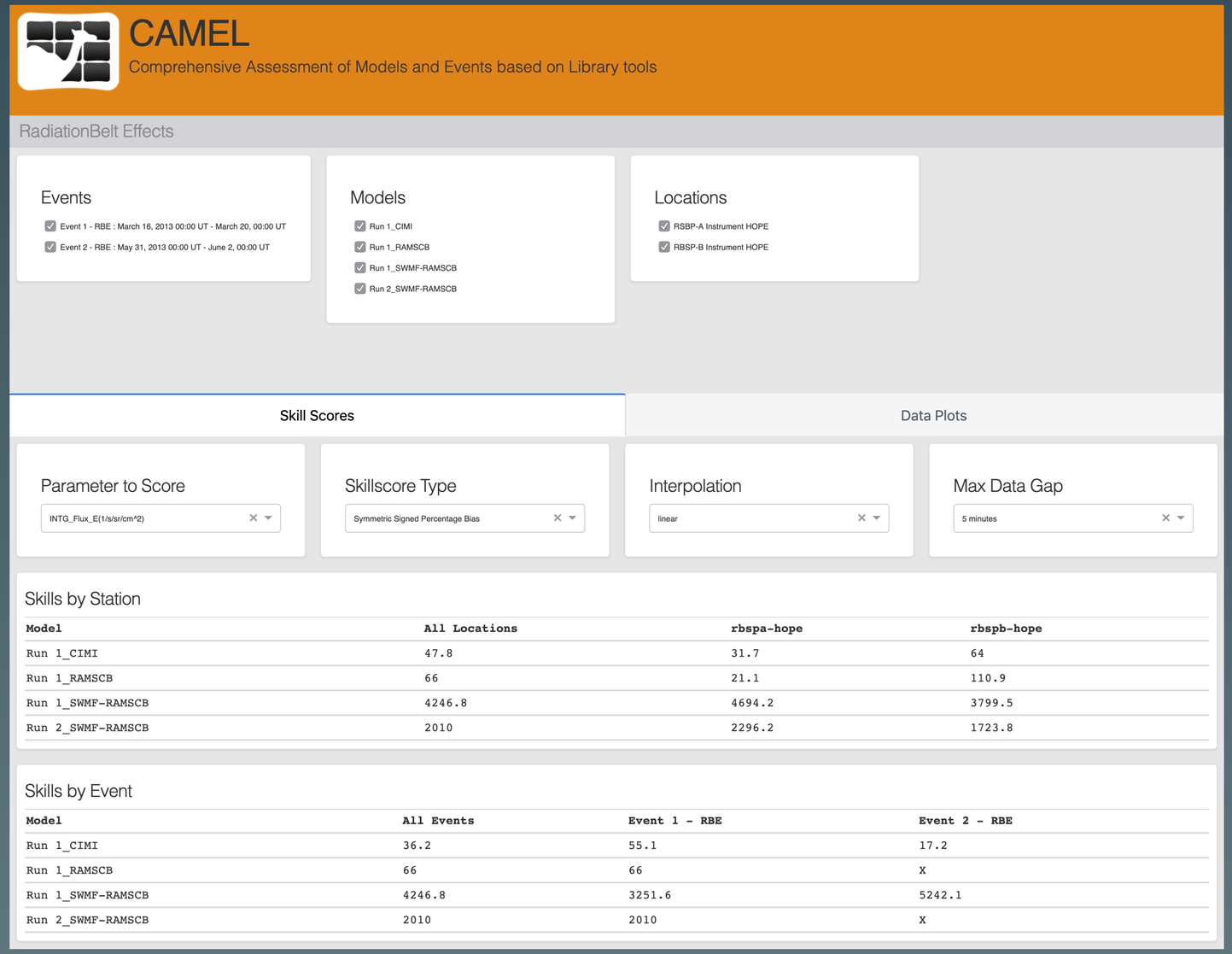
SPEDAS (spedas.org)



Angelopoulos et al., 2019

Validation tools

CAMEL computes skill scores from comparison of data with simulations in the CCMC database



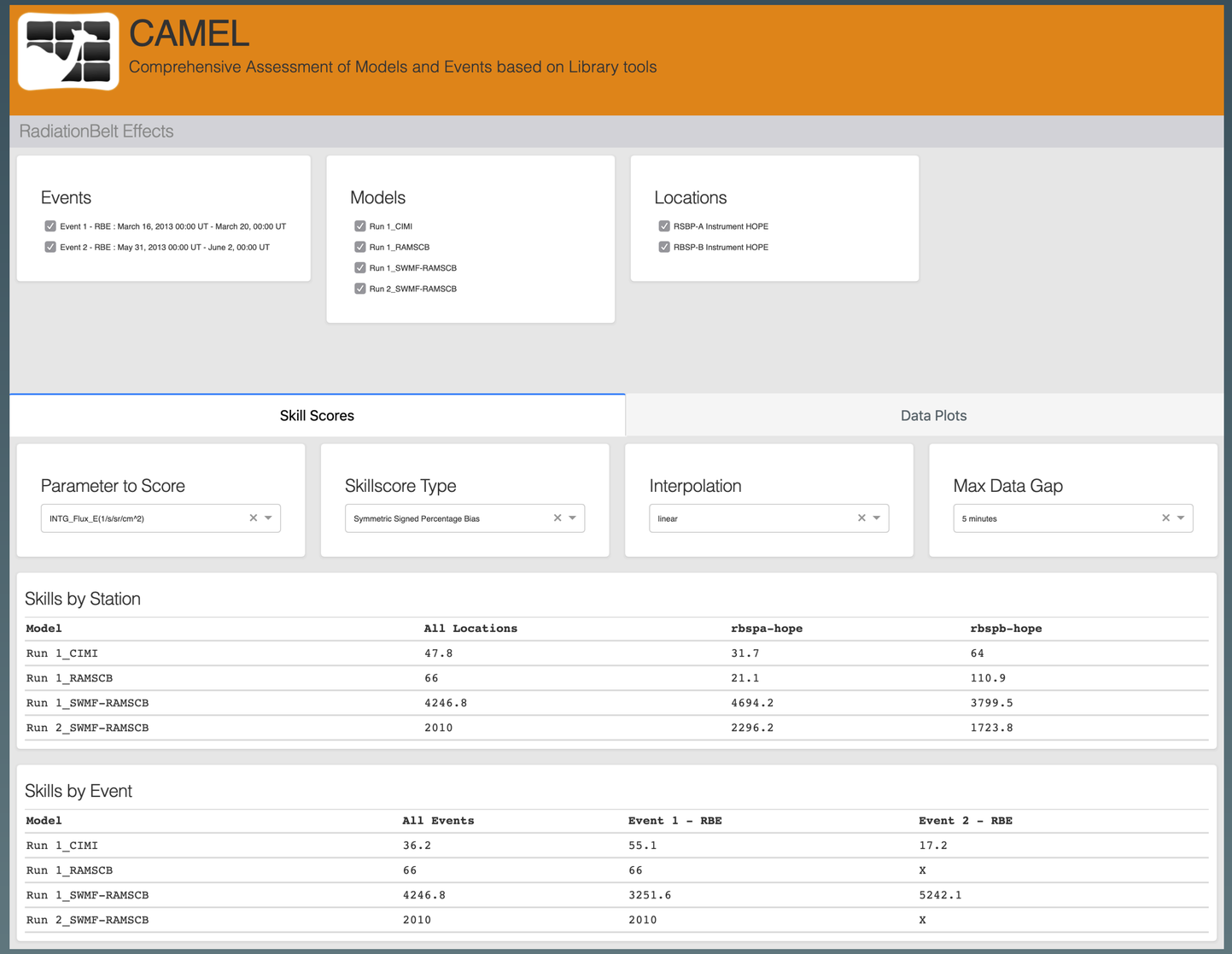
Model	All Locations	rbspa-hope	rbspb-hope
Run 1_CIMI	47.8	31.7	64
Run 1_RAMSCB	66	21.1	110.9
Run 1_SWMF-RAMSCB	4246.8	4694.2	3799.5
Run 2_SWMF-RAMSCB	2010	2296.2	1723.8

Model	All Events	Event 1 - RBE	Event 2 - RBE
Run 1_CIMI	36.2	55.1	17.2
Run 1_RAMSCB	66	66	X
Run 1_SWMF-RAMSCB	4246.8	3251.6	5242.1
Run 2_SWMF-RAMSCB	2010	2010	X

Validation tools

CAMEL is accepting recommendations for metrics, events, and data. It is planned to be open source.

ccmc.gsfc.nasa.gov/camel



Model	All Locations	rbspa-hope	rbspb-hope
Run 1_CIMI	47.8	31.7	64
Run 1_RAMSCB	66	21.1	110.9
Run 1_SWMF-RAMSCB	4246.8	4694.2	3799.5
Run 2_SWMF-RAMSCB	2010	2296.2	1723.8

Model	All Events	Event 1 - RBE	Event 2 - RBE
Run 1_CIMI	36.2	55.1	17.2
Run 1_RAMSCB	66	66	X
Run 1_SWMF-RAMSCB	4246.8	3251.6	5242.1
Run 2_SWMF-RAMSCB	2010	2010	X

Outline

Photo by Martin Adams on Unsplash

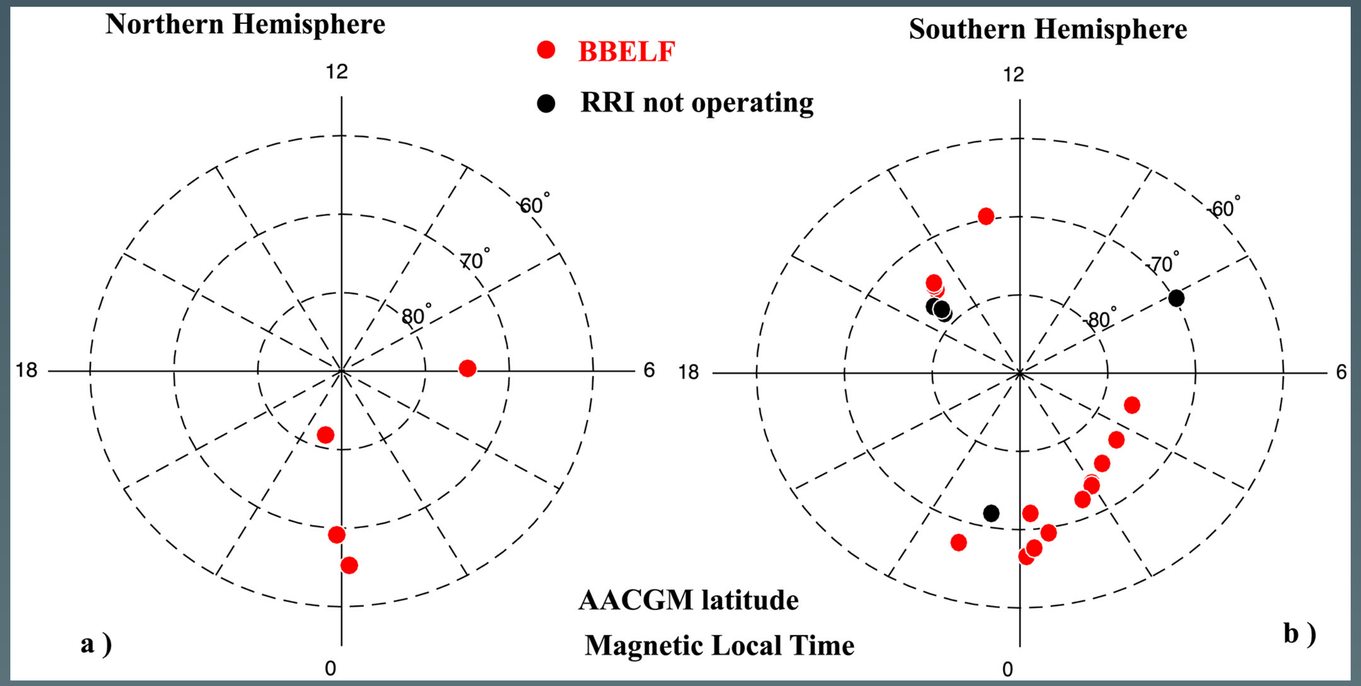
- Applications in our field
- Asking the validation question
- Validation processes and pathways
- Validation tools
- Validation needs, challenges, and opportunities
- Modeling Methods and Validation Focus Group

Validation needs, challenges, and opportunities: Statistics and Climatology

Upcoming Presentation in M3-I2:
e-POP Ion Composition Observations of Topside Ion Upflows: Storm Challenge Events 1 and 2 and Statistical Study of Other Storms in 2014-2017

by Andrew Yau

Ion Heating Observed with BBELF Waves

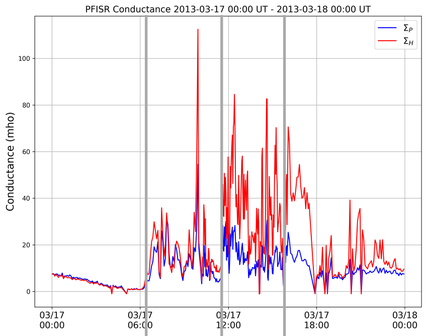


Shen et al., 2018

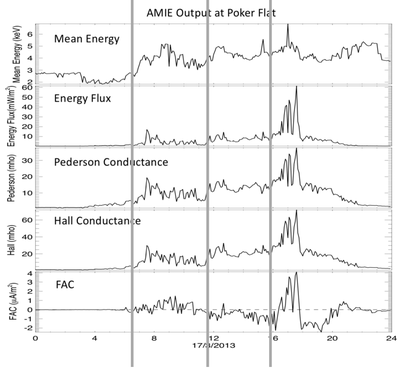
Validation needs, challenges, and opportunities: Quantifying Uncertainties during Storm Times ***and Quiet Times***

Conductance Challenge

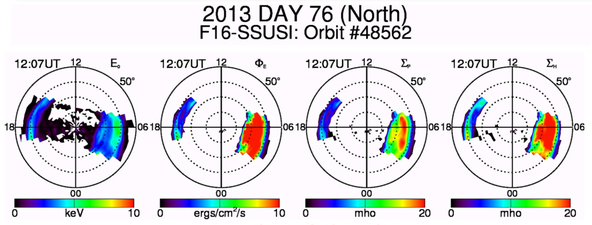
PFISR conductances Kaeppler et al.



AMIE output Gang Lu



How do we quantify uncertainties?

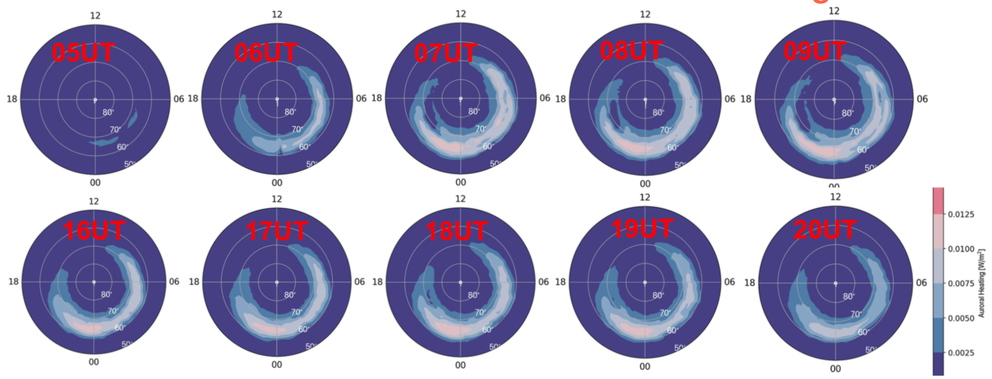


DMSP SSUSI Zhang and Paxton

Small-scale variability vs. global models

Height-integrated Auroral Heating

GITM Meng et al.

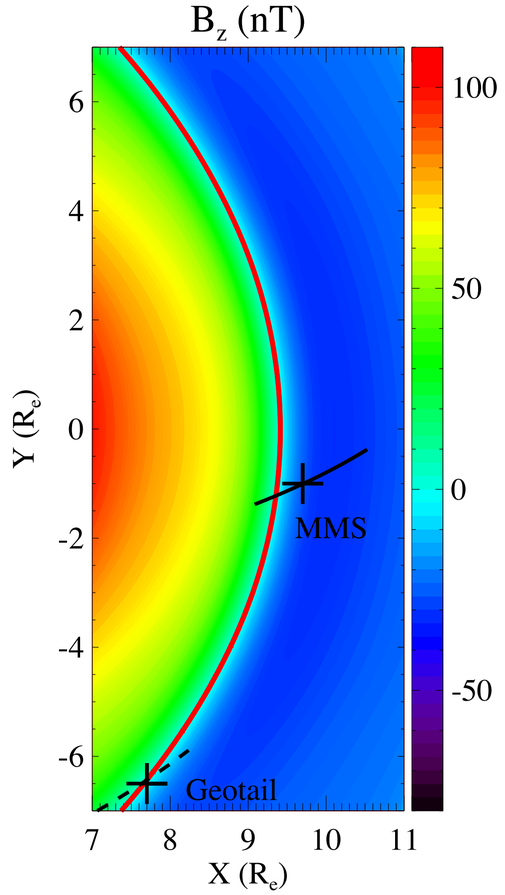


- Enhanced northern hemispheric auroral heating during both intervals

Validation needs, challenges, and opportunities: Validation in 2D and 3D

Dayside Kinetics FG Southward IMF Challenge

Compare in situ observations, remote observations, kinetic simulations, empirical models,…



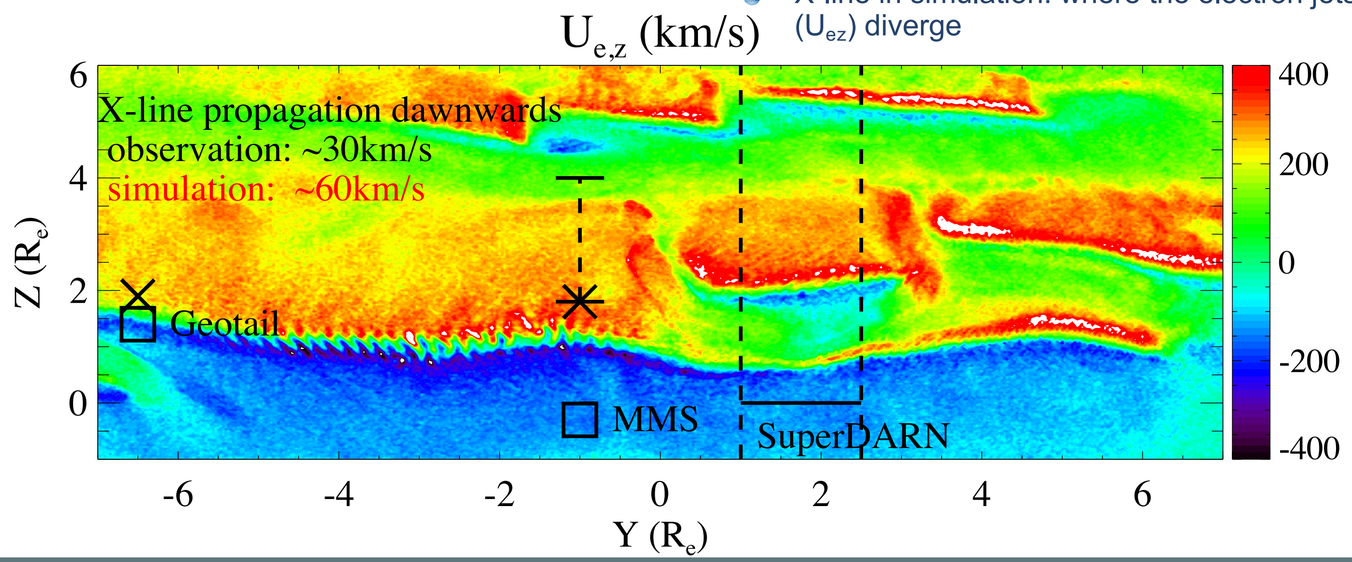
Highlight: MMS, Geotail, SuperDARN, and MHD-EPIC simulation
(Y. Chen, S. Vines, Y. Nishimura, G. Toth, H. Hietala, et al.)

Magnetopause (MP) location

- Black plus (+): MMS3 and Geotail observed MP
- Red curve: MP in simulation in the z=0 plane

X-line location

- Black square (□): MMS3 and Geotail location
- Black cross (X): X-line location estimated by MMS and Geotail
- X-line in simulation: where the electron jets (U_{ez}) diverge



Validation needs, challenges, and opportunities: How to get involved

- ***GEM Challenges!!!***
- Whole Heliosphere and Planetary Interval https://whpi.hao.ucar.edu/
- International Space Weather Action Teams (COSPAR Panel on Space Weather) http://ccmc.gsfc.nasa.gov/iswat
- iCCMC/LWS teams https://ccmc.gsfc.nasa.gov/assessment/

Outline

Photo by Martin Adams on Unsplash

- Applications in our field
- Asking the validation question
- Validation processes and pathways
- Validation tools
- Validation needs, challenges, and opportunities
- Modeling Methods and Validation Focus Group

GEM Validation Efforts

Focus Group	Resource Group
• Started in 2005 by Masha Kuznetsova and Aaron Ridley	• Presentation of proposal to Steering Committee Mini-GEM 2018 with transition in 2019?
• 5 year limit	• Permanent part of GEM
• 4 Co-Chairs	• 4 rotating Co-Chairs
• One of many focus groups to carry out GEM challenges	• Provide resources for focus groups to carry out GEM challenges

The Curiosity Paradox

Curiosity is unruly. It doesn’t like rules or, at least it assumes that all rules are provisional… It disdains the approved pathways, preferring diversions, unplanned excursions, impulsive left turns. In short, curiosity is deviant.

— from *Curious: The Desire to Know and Why Your Future Depends on It* by Ian Leslie

Validation is at the heart of GEM

Validation points to where the science is

We have to test the rules to find where they're broken

Find your validation question

How do I know whether this model, theory, or data is *valiд* for this application?

And bring validation to ***your*** focus groups