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 observational modes, each focusing on particular aspects of the geospace environment.”

-from GEM wiki, emphasis mine
What is GEM?

GEM involves *communication between different regimes* for *predictive understanding*. 
What is GEM?

GEM involves communication between different regimes for predictive understanding.

How? Validation!
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

- Pure basic research
- Use-inspired basic research
- Pure applied research
- Tinkering

Relevance for generalized knowledge vs. Relevance for immediate applications

Bohr
Pasteur
Tinkering
Edison
An alternate view
Outline

• 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 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

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 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
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 valid 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?

- Can this data be applied to refine my model uncertainties?
- How can my theory be applied to the needs of modelers?
- What model can be applied to provide context for my data?
Outline

• 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$^3$

Conductance model matters!
Grid resolution upstream of bow shock matters!

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

Lutz Rastaetter, et al. (2019 in prep.)
Mid-tail modeling challenge

Mid-Tail (X ~ −60 Re) 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 Re) 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 \(\rightarrow\) high N plasma penetrate deep to lobe \(\rightarrow\) reduce B

flow vortices N perturbations B perturbations
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
• 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 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.

Space Weather Capabilities Assessment

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

Table 1
A Summary of Key Metrics

<table>
<thead>
<tr>
<th>Metric</th>
<th>Definition</th>
<th>Symmetry</th>
<th>Scale/Order dependent</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\epsilon$</td>
<td>$y - x$</td>
<td>Y</td>
<td>Scale</td>
<td>Forecast error</td>
</tr>
<tr>
<td>$Q$</td>
<td>$y/x$</td>
<td>N</td>
<td>Order</td>
<td>Ratio; complement of forecast relative error</td>
</tr>
<tr>
<td>MSE</td>
<td>$\frac{1}{n} \sum_{i=1}^{n} (y_i - x_i)^2$</td>
<td>Y</td>
<td>Scale</td>
<td>Different units/scale; quadratic penalty</td>
</tr>
<tr>
<td>RMSE</td>
<td>$\sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - x_i)^2}$</td>
<td>Y</td>
<td>Scale</td>
<td>Same units as $x$, $y$; quadratic penalty</td>
</tr>
<tr>
<td>MAE</td>
<td>$\frac{1}{n} \sum_{i=1}^{n}</td>
<td>y_i - x_i</td>
<td>$</td>
<td>Y</td>
</tr>
<tr>
<td>MAPE</td>
<td>$\frac{1}{n} \sum_{i=1}^{n} \frac{y_i - x_i}{y_i}$</td>
<td>Y</td>
<td>Order</td>
<td>Percentage; penalizes overprediction more heavily</td>
</tr>
<tr>
<td>sMAPE</td>
<td>$\frac{100}{n} \sum_{i=1}^{n} \left</td>
<td>\frac{y_i - x_i}{y_i} \right</td>
<td>$</td>
<td>Y</td>
</tr>
<tr>
<td>ME</td>
<td>$\frac{1}{n} \sum_{i=1}^{n} (y_i - x_i)$</td>
<td>Y</td>
<td>Scale</td>
<td>Same units as $x$, $y$</td>
</tr>
<tr>
<td>MPE</td>
<td>$\frac{100}{n} \sum_{i=1}^{n} \frac{y_i}{y_i}$</td>
<td>N</td>
<td>Order</td>
<td>Percentage; penalizes overprediction more heavily</td>
</tr>
<tr>
<td>MAGE</td>
<td>$\frac{1}{n} \sum_{i=1}^{n} \log(</td>
<td>y_i</td>
<td>) - \log(</td>
<td>x_i</td>
</tr>
<tr>
<td>SEPB</td>
<td>$100 \frac{\operatorname{sgn}(Q) \log(\frac{x}{Q})}{\log(\frac{Q}{y})}$</td>
<td>Y</td>
<td>Order</td>
<td>Percentage; robust and resistant</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Metric</th>
<th>Definition</th>
<th>Symmetry</th>
<th>Scale/Order dependent</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\varepsilon$</td>
<td>$y - x$</td>
<td>Y</td>
<td>Scale</td>
<td>Forecast error</td>
</tr>
<tr>
<td>Q</td>
<td>$y/x$</td>
<td>N</td>
<td>Order</td>
<td>Ratio; complement of forecast relative error</td>
</tr>
<tr>
<td>MAPE</td>
<td>$\frac{100}{n} \sum_{i=1}^{n} \left</td>
<td>\frac{y_i}{x_i} \right</td>
<td>$</td>
<td>N</td>
</tr>
<tr>
<td>$\zeta$</td>
<td>$100 \left( e^{\left( M \left( \log_{e}(Q) \right) \right)} - 1 \right)$</td>
<td>Y</td>
<td>Order</td>
<td>Percentage; robust and resistant</td>
</tr>
</tbody>
</table>
Processes and pathways: choosing metrics

Mean Absolute Percentage Error

<table>
<thead>
<tr>
<th>Metric</th>
<th>Definition</th>
<th>Symmetry</th>
<th>Scale/Order dependent</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>$e$</td>
<td>$y - x$</td>
<td>Y</td>
<td>Scale</td>
<td>Forecast error</td>
</tr>
<tr>
<td>$Q$</td>
<td>$y/x$</td>
<td>N</td>
<td>Order</td>
<td>Ratio; complement of forecast relative error</td>
</tr>
<tr>
<td>MAPE</td>
<td>$\frac{100}{n} \sum_{i=1}^{n} \frac{</td>
<td>y_i - x_i</td>
<td>}{</td>
<td>y_i</td>
</tr>
<tr>
<td>$\xi$</td>
<td>$100 \left( e^{M[\log(\xi)]} - 1 \right)$</td>
<td>Y</td>
<td>Order</td>
<td>Percentage; robust and resistant</td>
</tr>
</tbody>
</table>
Processes and pathways: choosing metrics

<table>
<thead>
<tr>
<th>Metric</th>
<th>Definition</th>
<th>Symmetry</th>
<th>Scale/Order dependent</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\varepsilon$</td>
<td>$y - x$</td>
<td>Y</td>
<td>Scale</td>
<td>Forecast error</td>
</tr>
<tr>
<td>Q</td>
<td>$y/x$</td>
<td>N</td>
<td>Order</td>
<td>Ratio; complement of forecast relative error</td>
</tr>
<tr>
<td>MAPE</td>
<td>$100 \frac{1}{n} \sum_{i=1}^{n} \left</td>
<td>\frac{e(y_i)}{Q_i} \right</td>
<td>$</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>$100 \left(e^{\left(\log(e(Q_i))\right)} - 1\right)$</td>
<td>Y</td>
<td>Order</td>
<td>Percentage; robust and resistant</td>
</tr>
</tbody>
</table>

Median Symmetric Accuracy
## Processes and pathways: choosing metrics

<table>
<thead>
<tr>
<th>Metric</th>
<th>Definition</th>
<th>Symmetry</th>
<th>Scale/Order dependent</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \varepsilon )</td>
<td>( y - x )</td>
<td>Y</td>
<td>Scale</td>
<td>Forecast error</td>
</tr>
<tr>
<td>Q</td>
<td>( y/x )</td>
<td>N</td>
<td>Order</td>
<td>Ratio; complement of forecast relative error</td>
</tr>
<tr>
<td>MPE</td>
<td>( \frac{100}{n} \sum_{i=1}^{n} \frac{\hat{y}_i - y_i}{y_i} )</td>
<td>N</td>
<td>Order</td>
<td>Percentage; penalizes overprediction more heavily</td>
</tr>
<tr>
<td>SSPB</td>
<td>( 100 \text{ sgn}(\text{MdLQ}) (e^{\text{MdLQ}} - 1) )</td>
<td>Y</td>
<td>Order</td>
<td>Percentage; robust and resistant</td>
</tr>
</tbody>
</table>

\[
\text{MdLQ} = M \log_e(Q_i)
\]
Processes and pathways: choosing metrics

<table>
<thead>
<tr>
<th>Metric</th>
<th>Definition</th>
<th>Symmetry</th>
<th>Scale/Order dependent</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>ε</td>
<td>y – x</td>
<td>Y</td>
<td>Scale</td>
<td>Forecast error</td>
</tr>
<tr>
<td>O</td>
<td>y/x</td>
<td>N</td>
<td>Order</td>
<td>Ratio; complement of forecast relative error</td>
</tr>
<tr>
<td>MPE</td>
<td>$\frac{100}{n} \sum_{i=1}^{n} \frac{e^{x_i}}{x_i}$</td>
<td>N</td>
<td>Order</td>
<td>Percentage; penalizes overprediction more heavily</td>
</tr>
<tr>
<td>SSPB</td>
<td>$100 \text{sgn}(\text{MdLQ}) (e^{\text{MdLQ}} - 1)$</td>
<td>Y</td>
<td>Order</td>
<td>Percentage; robust and resistant</td>
</tr>
</tbody>
</table>

MdLQ \( = M \log_e(Q_i) \)

Mean Percentage Error
Processes and pathways: choosing metrics

<table>
<thead>
<tr>
<th>Metric</th>
<th>Definition</th>
<th>Symmetry</th>
<th>Scale/Order dependent</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\varepsilon$</td>
<td>$y - x$</td>
<td>Y</td>
<td>Scale</td>
<td>Forecast error</td>
</tr>
<tr>
<td>Q</td>
<td>$y/x$</td>
<td>N</td>
<td>Order</td>
<td>Ratio; complement of forecast relative error</td>
</tr>
<tr>
<td>MRE</td>
<td>$\frac{100}{n} \sum_{i=1}^{n} \frac{e_i}{x_i}$</td>
<td>N</td>
<td>Order</td>
<td>Percentage; penalizes overprediction more heavily</td>
</tr>
<tr>
<td>SSPB</td>
<td>$100 \operatorname{sgn}(\text{MdLQ}) (e^{\text{MdLQ}} - 1)$</td>
<td>Y</td>
<td>Order</td>
<td>Percentage; robust and resistant</td>
</tr>
</tbody>
</table>

MdLQ = $M \log_{e}(Q_i)$

Signed Symmetric Percentage Bias
Validation tools

Overview of Python Packages

<table>
<thead>
<tr>
<th>Data Access and Analysis</th>
<th>Observational Data Access and Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modeled Data Access and Analysis</td>
<td>DaViTpy/pyDARN</td>
</tr>
<tr>
<td>analytsator</td>
<td>digital_rf</td>
</tr>
<tr>
<td>OvationPyme</td>
<td>GeoData</td>
</tr>
<tr>
<td>pyAMPS</td>
<td>HelioPy</td>
</tr>
</tbody>
</table>

- CDFlib
- pyLRT
- pysatCDF

Data Analysis and File Routines

<table>
<thead>
<tr>
<th>Coordinates</th>
<th>Multipurpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>AACGMV2</td>
<td>AstroPy</td>
</tr>
<tr>
<td>Apexpy</td>
<td>geospacepy</td>
</tr>
<tr>
<td>OCBpy</td>
<td>SpacePy</td>
</tr>
<tr>
<td>pysatMagVect</td>
<td>SunPy</td>
</tr>
</tbody>
</table>

- SpicyPy
- PyEphem
- SGP4
- skyfield and jplephem

Heliophysics Field Key

- Sun/Solar Wind
- Magnetosphere
- Ionosphere/Thermosphere/Mesosphere
- Other

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.
Validation tools

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

ccmc.gsfc.nasa.gov/camel
Outline

• 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

How do we quantify uncertainties?

Small-scale variability vs. global models

PFISR conductances
Kaeppler et al.

AMIE output
Gang Lu

DMSP SSUSI
Zhang and Paxton

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

- X-line propagation dawnwards
  - Observation: ~30km/s
  - Simulation: ~60km/s

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

• 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

<table>
<thead>
<tr>
<th>Focus Group</th>
<th>Resource Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Started in 2005 by Masha Kuznetsova and Aaron Ridley</td>
<td>• Presentation of proposal to Steering Committee Mini-GEM 2018 with transition in 2019?</td>
</tr>
<tr>
<td>• 5 year limit</td>
<td>• Permanent part of GEM</td>
</tr>
<tr>
<td>• 4 Co-Chairs</td>
<td>• 4 rotating Co-Chairs</td>
</tr>
<tr>
<td>• One of many focus groups to carry out GEM challenges</td>
<td>• Provide resources for focus groups to carry out GEM challenges</td>
</tr>
</tbody>
</table>
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 valid for this application?

And bring validation to your focus groups