Global MHD

Jimmy Raeder
University of New Hampshire

with contributions from Kai Germaschewski, Doug Cramer, Bashi Ferdousi, Beket Tulegenov (UNH), Naomi Maruyama (NOAA/CU)

Outline

- Magneto-hydro-dynamic equations: what they are, what not, and what the limitations are.
- How do we solve them: nothing is perfect.
- The most fundamental process: magnetic reconnection.
- The global system: boundaries and sub-models.
- Why do we do this?
- Proof is in the pudding: some examples.
- What next?

MHD equations

- The MHD equations are the conservation equations for mass, momentum, energy, and magnetic flux.
- Even though they include simplifications, any more sophisticated plasma model would still have to obey the MHD equations, except for terms related to the displacement current.
- The elementary solutions to the MHD equations are frequently found in space plasmas: Alfven waves, magnetosonic waves, tangential discontinuities, rotational discontinuities, shocks, surface waves (KH).
- The MHD equations are derived from the Vlasov (or Boltzmann) equation and Maxwell's equations.
- The former equations are the foundation of kinetic theory and include temporal and spatial scales that are completely impractical for global simulations.
- The goal is to simplify the equations such that they still capture all relevant physical processes at well defined macro- and meso-scales.

Derivation I

- Not many details here, those are in textbooks. But we need to understand the limitations.
- In Maxwell's equations the displacement current is dropped. That eliminates the time/spatial scale of light waves. It also modifies the Alfven speed (which has no limit in this approximation). This affects the near Earth region where the Alfven speed can get very large (aka "Alfven resonator") but that region is generally excluded from global simulations. However, the displacement current is often inserted back in the numerical solution to limit the time step (aka "Boris correction").
- The plasma equations are the first 3 moment equations of the Vlasov/Boltzmann equation (see Barakat and Schunk, 1982.)
- Ion and electron moments are combined under the assumption that me/mi (=1/1836) is small and that Ne=~Ni (quasi-neutrality). Those assumption eliminate all processes that explicitly depend on charge separation, like plasma oscillations, i.e., plasma frequency and Debye length.

Derivation III

- The continuity and the momentum equation are straight forward:

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot (\rho \mathbf{v})$$

$$\frac{\partial \rho \mathbf{v}}{\partial t} = -\nabla \cdot \left\{ \rho \mathbf{v}\mathbf{v} + p\underline{\underline{\mathbf{I}}} - (\mathbf{B}\mathbf{B} - \frac{1}{2}B^2\underline{\underline{\mathbf{I}}}) \right\}$$

- The equations are written here in divergence form. This is important for their numerical treatment.
- Pressure is assumed to be isotropic. The next approximation would be the CGL approximation with anisotropic pressure.
- There are also no source terms. In much of the magnetosphere there are no significant sources or sinks, but they can exist close to Earth due to charge exchange, for example.

Derivation IV

- The energy equation is more involved:

$$\frac{\partial U}{\partial t} = -\nabla\cdot\{(U + p + B^2)\mathbf{v} - (\mathbf{v}\cdot\mathbf{B})\mathbf{B} + \mathbf{j}\times\mathbf{B} + \mathbf{q}\}$$

with $U = p/(\gamma - 1) + \rho v^2 + B^2/2$ (total energy density), or:

$$\frac{\partial e}{\partial t} = -\nabla\cdot(\{e + p\}\mathbf{v} + \mathbf{q}) + \mathbf{j}\cdot\mathbf{E}$$

with $e = \rho v^2/2 + p/(\gamma - 1)$ (plasma energy density.)

- Both forms are mathematically equivalent, but the latter behaves better numerically.
- The vector $\mathbf{q}$ is the heat flux. It arises because the moment equations are an infinite hierarchy.
- The simplest *closure* is usually applied, that is zero heat flux.
- Zero heat flux is consistent with Maxwellian particle distributions.

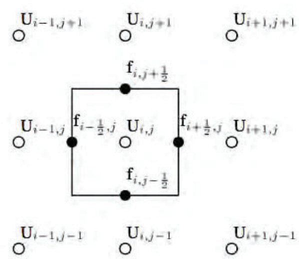
Derivation V

Limitations summarized:

- The MHD spatial scale limits are the ion gyro radius and the ion skin depth (c/w_pi). These are typically less than 100 km in the magnetosphere.
- The MHD temporal scale limit is the gyro period.
- Closure requires zero heat flux, also equivalent with the conservation of specific entropy s=p/rho^gamma. That is consistent with symmetric distribution functions f(v).
- A static electric field does not appear in the equations. That does not mean it doesn't exist. The charge density can be computed from Gauss' law (the *quasi* in quasi-neutrality.)
- The magnetic field does not do any work on particles. Therefore **B** does not appear explicitly in the energy equation, consistent with the Poynting theorem.
- The ideal MHD equation do not permit reconnection. However, there is no stable numerical method that solves the ideal MHD equations without introducing numerical resistivity.

Numerical Solution I

- Typical solution methods use Finite Difference or Finite Volume methods. Grids vary a bit between major codes: Stretched Cartesian (OpenGGCM), Block-Adaptive (BATSRUS), and matched spherical/cylindrical (LFM/GAMERA).
- The conservation character of the equations must be carried over to the numerical solution: Mass, momentum and energy are conserved exactly by placing variables at cell centers and fluxes at face centers:



$$\frac{\partial U}{\partial t} = -\left(f_{i+\frac{1}{2},j}(U) - f_{i-\frac{1}{2},j}(U)\right)/\Delta x - \left(f_{i,j+\frac{1}{2}}(U) - f_{i,j-\frac{1}{2}}(U)\right)/\Delta y$$

- The fluxes are “limited” to provide non-oscillatory stable solutions at discontinuities.

Numerical Solution II

- Ultimately, the finite difference error terms introduce both diffusion and dispersion:

$$\Delta x \frac{\partial U}{\partial t} = -(f_{i+\frac{1}{2}} - f_{i-\frac{1}{2}}) + a_1(\Delta x)^2 \frac{\partial^2}{\partial x^2} F(U) + b_1(\Delta x)^3 \frac{\partial^3}{\partial x^3} F(U) + a_2(\Delta x)^4 \frac{\partial^4}{\partial x^4} F(U) + b_2(\Delta x)^5 \frac{\partial^5}{\partial x^5} F(U) + \dots$$

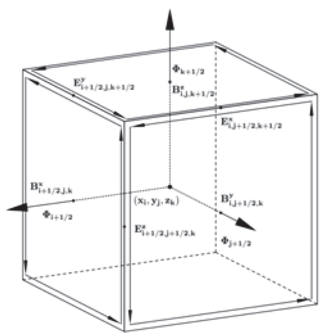
- Calculation of the fluxes is an art. Requires balance between accuracy, stability, dispersion, and diffusion. Diffusion free schemes are possible, but suffer strong dispersion (oscillation at discontinuities).
- Most schemes are second order accurate, that is the leading error term is diffusive and ~ $(\Delta x)^2$.
- Most codes employ an explicit time integration scheme that limits the stable time step to be proportional to the smallest grid size. Thus, the computational cost is inverse to $(\Delta x)^4$: half the grid size costs 16 times to compute.

Numerical Solution III

- Conservation of magnetic flux: Faraday's law and Maxwell's law require that magnetic flux is conserved and the magnetic field be divergence free:

$$\nabla \cdot \mathbf{B} = 0 \qquad\qquad \nabla \cdot \frac{\partial \mathbf{B}}{\partial t} = \frac{\partial (\nabla \cdot \mathbf{B})}{\partial t} = -\nabla \cdot \nabla \times \mathbf{E} = 0$$

- This can be achieved with the "Constrained Transport" algorithm and a staggered grid (Yee grid):



$$\begin{aligned}
&\frac{\partial}{\partial t}\int\int_{cell} \Phi df = \Delta y \Delta z (\frac{\partial B_x}{\partial t})_{i-\frac{1}{2}} \\
&+\Delta y \Delta z (\frac{\partial B_x}{\partial t})_{i+\frac{1}{2}} + \Delta x \Delta z (\frac{\partial B_y}{\partial t})_{j+\frac{1}{2}} + \dots \\
&= \{((E_y)_{i+\frac{1}{2},j,k+\frac{1}{2}} - (E_y)_{i+\frac{1}{2},j,k+\frac{1}{2}}) + \\
&((E_y)_{i+\frac{1}{2},j,k-\frac{1}{2}} - (E_y)_{i+\frac{1}{2},j,k-\frac{1}{2}}) + \dots\} \Delta x \Delta y \Delta z = 0
\end{aligned}$$

- CT (or variations thereof) is employed by many, but not all codes.

Numerical Solution IV

- Reconnection is one of the most important processes in the magnetosphere and many other settings (corona, heliosphere,..), but impossible in ideal MHD. Reconnection requires the existence of a localized parallel electric field. Such a field can be caused by resistivity, electron inertia, or a divergence of the electron pressure tensor.
- In MHD codes only resistivity can produce reconnection. Since every practical MHD code has numerical resistivity all such codes produce reconnection. Sometimes extra (anomalous) resistivity terms are added.
- The questions thus are:
  - Does it matter what produces the parallel electric field? → Not really. The MHD scales are much larger than the e- scales. So, it only matters if the parallel field has the correct magnitude and extent. Theory tells us that the dependence on resistivity (and thus the parallel field) is weak (~log or ~ some small power). Is it the right location? In MHD the parallel field is strongest where the currents are the strongest, but that is also where the kinetic effects maximize. Check.
  - Are the MHD reconnection rates correct? → Theory would indicate so. But in the magnetosphere the global reconnection rate is quite well measured (= CPCP) and global codes are well within the ballpark. Check
  - Does reconnection in MHD codes occur at the right locations and at the right times? → Much harder to assess. In MHD codes thin current sheets begin to reconnect quickly. In nature there may be delays (CDAW9 event). Not clear yet.
- Reconnection seems to be similar to boundary layer physics in viscous flows: An airplane can't fly without viscosity, but the value of viscosity matters little.

Numerical Solution V

- The 2001 GEM Reconnection Challenge paper (Birn et al., JGR) seems to indicate that MHD simulations cannot produce fast reconnection rates. The figure in that paper is misleading, because it shows only the result from a MHD simulation with uniform resistivity.
- The A. Otto paper (same issue) showed that non-uniform resistivity produces reconnection rates that are essentially the same as for those models that include the Hall term.
- In global MHD codes resistivity is always (with very few deliberate exceptions) current dependent and thus localized.

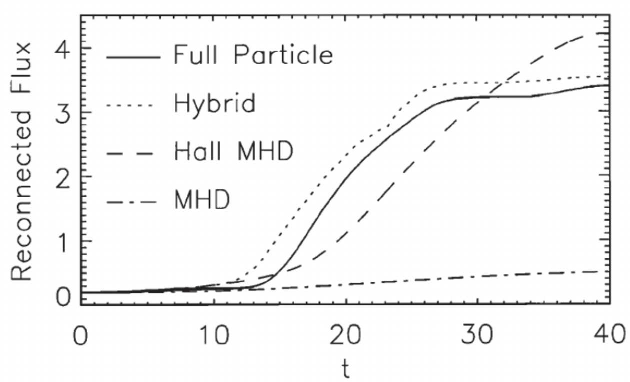
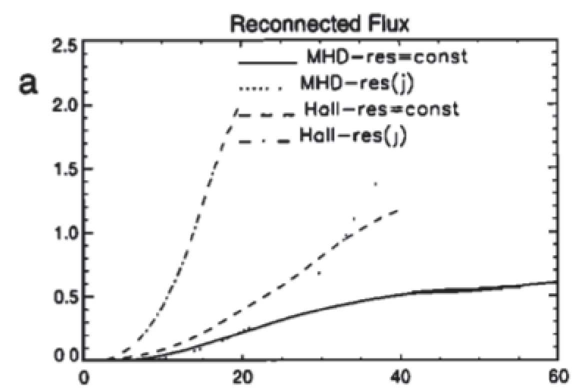
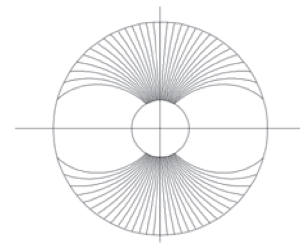


Figure 1. The reconnected magnetic flux versus time from a variety of simulation models: full particle, hybrid, Hall MHD, and MHD (for resistivity $\eta = 0.005$).



Going Global

- There are many versatile 3d MHD codes out there: ZEUS, ATHENA, FLASH, PLUTO, …
- However, modeling the magnetosphere requires much more, in particular coupling to the ionosphere. Processes on field lines close to Earth are not well described by MHD (huge Alfven speed, non-Maxwellian f(v), parallel electric fields), so MHD equations are only solved to with 1.5 – 2.5 RE from Earth and quantities are mapped in between.



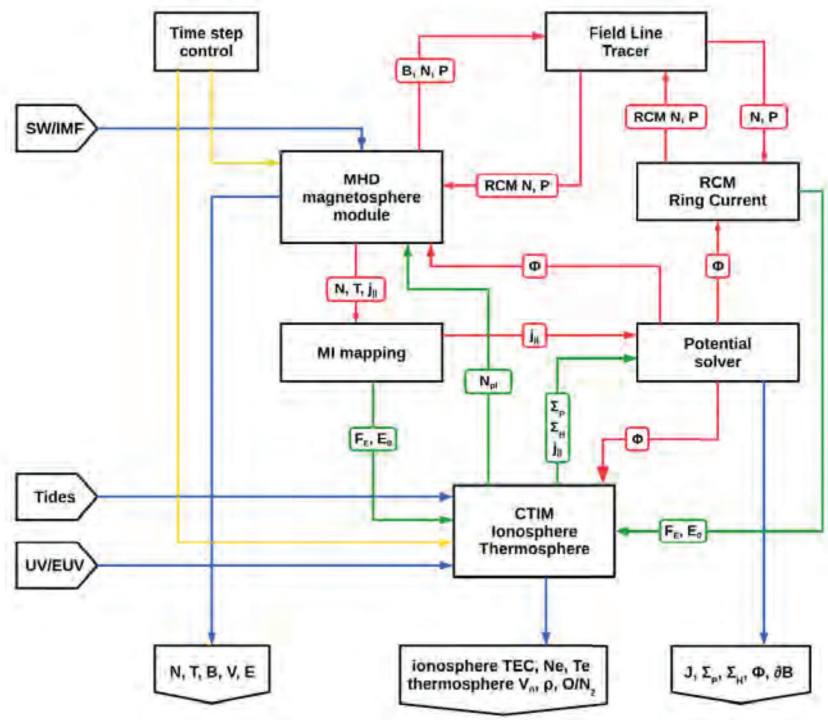
- At a minimum, an equation for the ionosphere potential is required:

$$\nabla \cdot \underline{\underline{\Sigma}} \cdot \nabla \Phi = -j_{||} \sin I$$

- The FAC comes from the ionosphere, and the conductance tensor is either an empirical model or a full blown ionosphere – thermosphere model (in OpenGGCM currently CTIM or WAM-IPE).
- The latter also requires a sub model for e- precipitation, typically some parameterization based on the Knight relation or similar.
- Feedback to the magnetosphere is through the electric field, setting a boundary condition for convection. Some codes also include ion outflow.

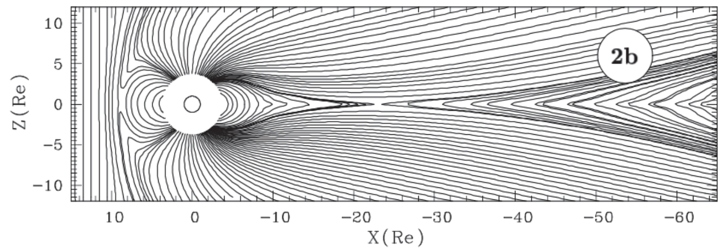
Putting it all together

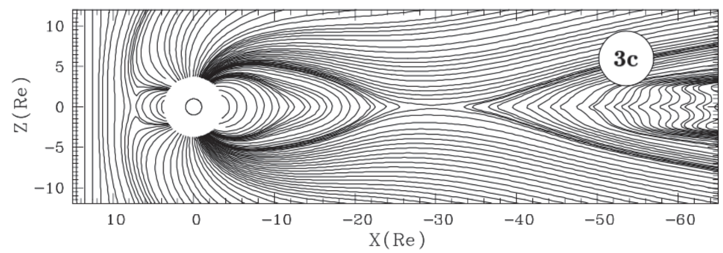
- Speaking for OpenGGCM now.
- I'm skipping the discussion of the inner magnetosphere here, where the RCM rules (and soon CIMI as well.) Obviously, the drift physics is not that well handled by MHD.
- It's all MPI parallelized and runs on 8-32,000 cores with grids that range from 10^6 to 10^9 cells. Runs with 10^7 cells and ~500 cores are faster than real time. Stretched Cartesian allows for resolutions down to 50 km.
- V5.0 is available at CCMC for runs on demand.
- Presently adding CIMI as an option instead of RCM (adding pitch angles), and replacing CTIM with WAM-IPE (better low-latitude ionosphere, plasmasphere, and all the way to the ground.)



Global I: Ionosphere feedback

- Very old simulation.
- Increasing ionosphere conductance.
- Top: Self-consistent ionosphere conductance.
- Bottom: Artificially high conductance (50 S).
- High conductance 'chokes' convection and reconnected flux from the tail takes longer to convect back to the dayside.
- In the limit of infinite conductance convection would be completely suppressed.
- In the opposite limit (Mercury) convection proceeds freely.
- The Earth is somewhere in between.
- → ionosphere feedback is critical.

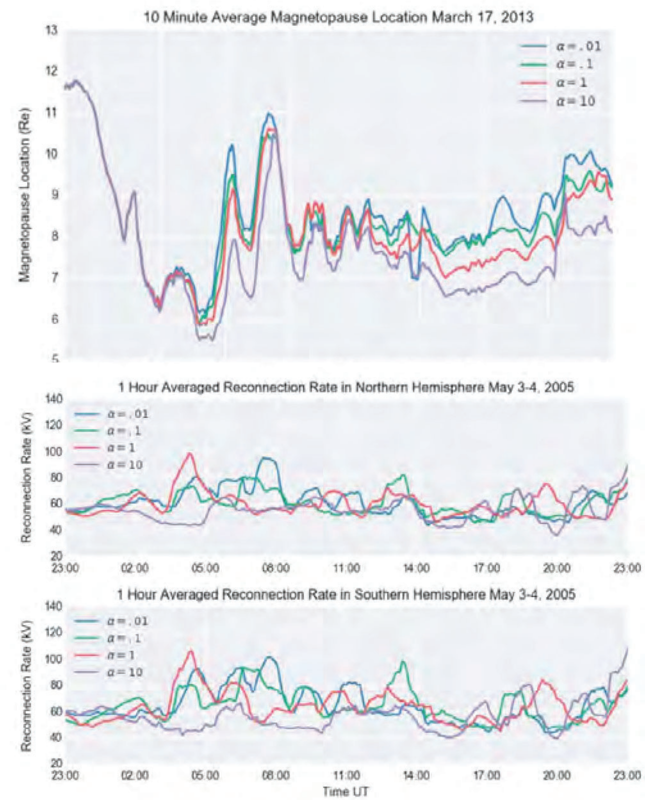




Raeder et al., in: *The Physics of Space Plasmas*, editor(s): T. Chang, and J.R. Jasperse, 14, 403, Publisher: MIT Center for Theoretical Geo/Cosmo Plasma Physics

Global II: Ionosphere feedback

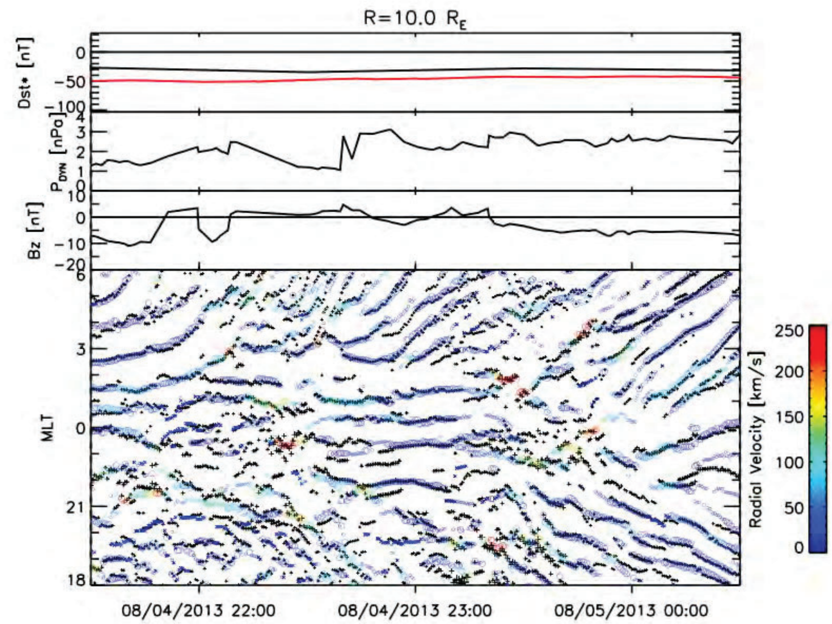
- More subtle ionosphere feedback.
- Changing parameters that affect the electron precipitation. Higher alpha → larger conductance.
- Top: Magnetopause stabdoff distance decreases as conductance increases.
- Bottom: Reconnection rate decreases as conductance increases.
- Lower conductance produces more substorm-like spikes in the reconnection rate. → High conductance may suppress substorms.
- → again, ionosphere feedback is critical.



Jensen et al., JGR, 2017

Global III: Injections

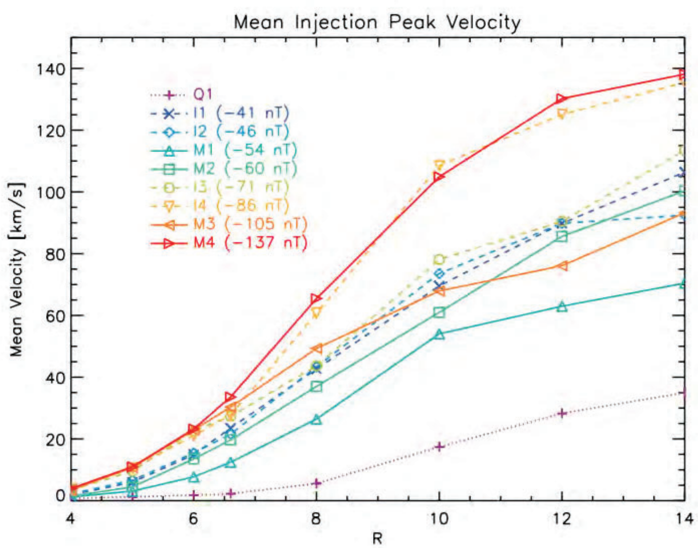
- How do BBFs/Bubbles inject plasma into the inner magnetosphere?
- Requires fully coupled model with RCM.
- Snake plot shows at R=10 RE as a function of time: speed maxima (color coded) and minima of flux tube entropy.
- Shows that BBFs are, for the most part, low entropy bubbles.
- Bubbles move from local midnight towards the flanks.

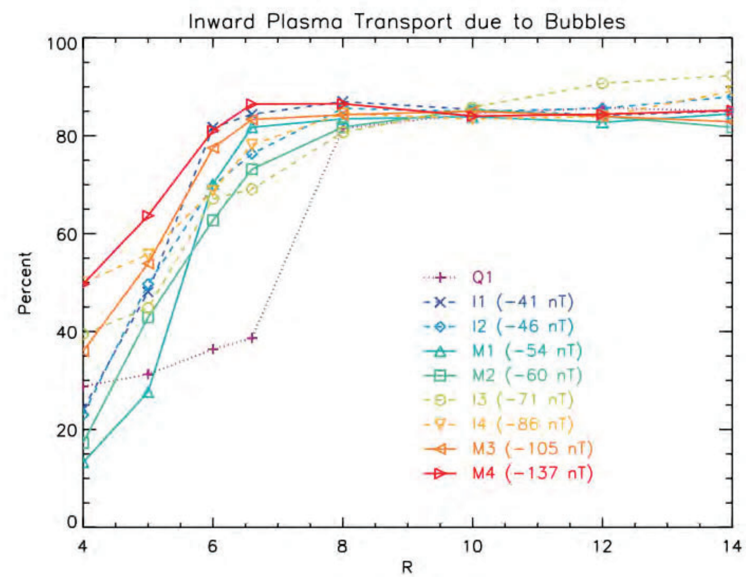


Cramer et al., JGR, 2017

Global IV: Injections

- 80% of plasma transport is due to bubbles. Closer to Earth the fraction decreases.
- The breaking point where bubbles become less important is farther away for quiet time.
- The injection velocity continually decreases closer to Earth.

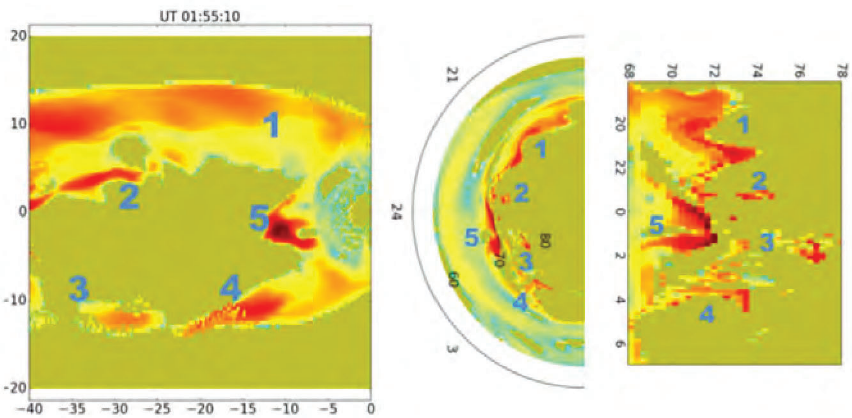


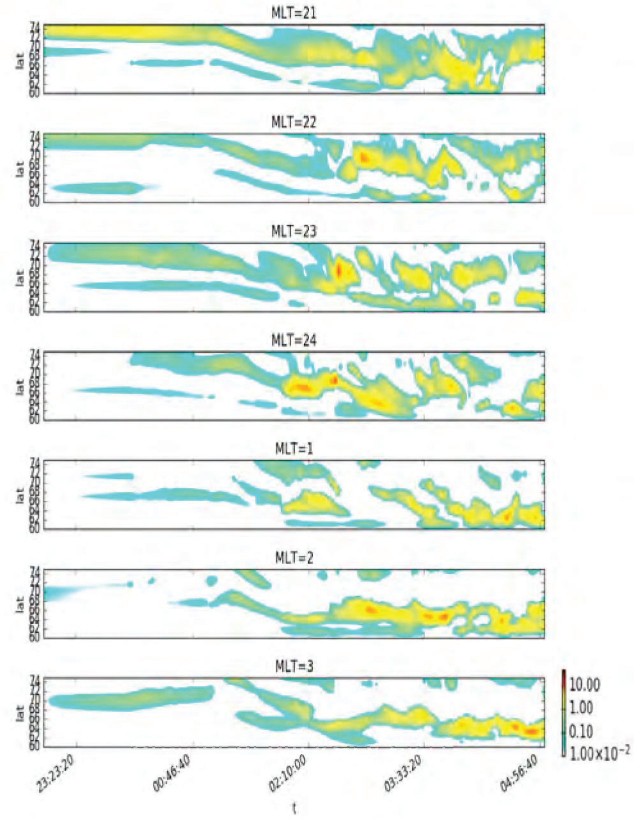


Cramer et al., JGR, 2017

Global V: BBFs and PBI/streamers

- There has been some evidence that BBFs are the source of Poleward Boundary Intensifications (PBIs) and/or auroral streamers.
- Here, we use OpenGGCM to map BBFs into the ionosphere and show that the resulting features resemble PBIs/streamers.
- Because field lines converge quickly from the tail to the ionosphere, the ionosphere quantities such as FAC don't resolve streamers, but we can show them by simple mapping.
- Synthetic keograms show poloar cap expansion and substorm onset.





Global VI: BBFs and PBI/streamers

- This event starts out as quiet, followed by a substorm (essentially one big BBF), which does not recover but becomes a SMC.
- BBFs are present all the time, but increase substantially with the substorm and remain at an elevated level during the SMC.

Ferdousi et al., JGR, in revision

What next?

- Add CIMI (Mei-Ching Fok's inner magnetosphere model) with pitch angle resolution and radiation belt physics.
- Replace CTIM with WAM-IPE (Whole Atmosphere Model – Ionosphere plasmasphere Electrodynamics) from NOAA.
- Data Assimilation, for example with DART (Data Assimilation Research Testbed) ensemble Kalman filter (EnKF) to optimize parameters.
- Investigate ML techniques to estimate parameters.
- We need a magnetosphere constellation mission with 200+ small s/c, Mag+Plasma.
- More ground magnetometers, AMPERE, radars would also help.

Summary and final thoughts

- Global MHD modeling of the magnetosphere has come a long way. Computational constraints are not much of an issue any more.
- Coupling with the ionosphere-thermosphere system is critical. Comparable to global atmosphere models that would not work without ocean/cryosphere/biosphere models.
- M-I coupling is still primitive and needs a lot of improvements.
- Global MHD models get a lot of things right. In particular, reconnection in MHD models fits observations quite well.
- BUT, results need to be interpreted with caution due to limited physics, limited resolution, and numerical effects.
- Numerical diffusion has always been considered a bad thing. However, experience with the inherently diffusive models may indicate that the 'collisionless' plasma of the magnetosphere is not quite that ideal, possibly due to turbulence. In some sense MHD models may be in a "golden zone" where numerical imperfections mimic real dissipative and non-ideal processes.