SWMF Global Magnetosphere Simulations of January 2005: Geomagnetic Indices and Cross-Polar Cap Potential

John D. Haiducek[1], Daniel T. Welling[1], Natalia Y. Ganushkina[1,2], Steven K. Morley[3], and Dogacan Su Öztürk[1]

[1]Climate and Space Sciences, University of Michigan, Ann Arbor, MI, USA.

[2]Finnish Meteorological Institute, Helsinki, Finland

[3]Space Science and Applications Group (ISR-1), Los Alamos National Laboratory, Los Alamos, NM, USA.

Key Points:

- Increasing grid resolution from that used by SWPC improves AL index prediction during disturbances, but has little effect on Kp, Sym-H, or CPCP
- The model does an excellent job at predicting Sym-H, but less well in predicting AL.
- SWMF tends to over-predict Kp and CPCP during quiet times, but predicts those quantities better during active times.

Corresponding author: John Haiducek, jhaiduce@umich.edu

Abstract

We simulated the entire month of January, 2005 using the Space Weather Modeling Framework (SWMF) with observed solar wind data as input. We conducted this simulation with and without an inner magnetosphere model, and tested two different grid resolutions. We evaluated the model's accuracy in predicting Kp, Sym-H, AL, and cross polar cap potential (CPCP). We find that the model does an excellent job of predicting the Sym-H index, with an RMSE of 17-18 nT. Kp is predicted well during storm-time conditions, but over-predicted during quiet times by a margin of 1 to 1.7 Kp units. AL is predicted reasonably well on average, with an RMSE of 230-270 nT. However, the model reaches the largest negative AL values significantly less often than the observations. The model tended to over-predict CPCP, with RMSE values on the order of 46-48 kV. We found the results to be insensitive to grid resoution, with the exception of the rate of occurrence for strongly negative AL values. The use of the inner magnetosphere component, however, affected results significantly, with all quantities except CPCP improved notably when the inner magnetosphere model was on.

1 Introduction

Magnetohydrodynamic (MHD) models [e.g. *De Zeeuw et al.*, 2000; *Lyon et al.*, 2004], coupled with inner magnetosphere and ionosphere models [e.g. *Pembroke et al.*, 2012; *Glocer et al.*, 2012; *Yu et al.*, 2014; *Cramer et al.*, 2017], are a powerful tool for understanding the dynamics of the Earth's magnetosphere [e.g. *Crooker et al.*, 1998; *Zhang et al.*, 2007]. By solving a subset of Maxwell's equations, an MHD solver provides magnetic fields and current systems throughout its computational domain. Coupling the MHD model to an inner magnetosphere and ionosphere model produces a system that accounts for ring currents and ionospheric currents as well. This results in a detailed representation of magnetospheric dynamics that is applicable under a wide variety of conditions.

These capabilities naturally make the coupled global MHD and ring current approach attractive for forecasting applications. In 2016 the NOAA Space Weather Prediction Center (SWPC) added a geospace modeling capability based on the Space Weather Modeling Framework (SWMF) [*Tóth et al.*, 2005; *Tóth et al.*, 2012] to their suite of operational forecasting tools (http://clasp.engin.umich.edu/articles/view/715). This was the result of a community validation effort focusing on six storm events, in which three MHD models and two empirical models were evaluated with respect to their ability to predict

$\frac{dB}{dt}$ at several ground-based magnetometer stations. The validation effort is described in *Pulkkinen et al.* [2013], and builds from *Pulkkinen et al.* [2010] and *Rastätter et al.* [2011]. *Pulkkinen et al.* [2013] found that the SWMF achieved the best predictive skill of the models evaluated, but with the caveat that the predictions delivered by SWMF may not be adequate for some operational uses. A number of follow-up papers have examined the results of this effort further. *Glocer et al.* [2016] evaluated the models' ability to reproduce the local K index, finding that the SWMF performed especially well in predicting local K. *Welling et al.* [2017] showed that the SWPC events exceeded the range of validity for the empirical ionospheric conductance models used in the participating MHD codes, and that all of the models tended to underpredict surface $\frac{dB}{dt}$, though SWMF less so than the others. *Anderson et al.* [2017] compared the field-aligned currents from the models with those obtained using the Active Magnetosphere and Planetary Electrodynamics Response Experiment (AMPERE).

Though unique in its rigorous comparison of multiple models, the scope of *Pulkkinen et al.* [2013] was limited to a small number of storm events. This has been common practice within the MHD modeling community in recent years. Simulations of single storm events constitute a majority of existing MHD papers. Some representative examples include *Raeder et al.* [2001], which simulated the 14-16 July 2001 "Bastille Day" storm, *Palmroth et al.* [2003], which simulated a major storm from April 6-7 2000, *Lopez et al.* [2001] which simulated a March 1995 substorm and a January 1997 storm, and *Kress et al.* [2007] which shows MHD and particle tracing results for the 29 October 2003 storm. MHD models have also been used to study hypothetical extreme events to better understand the possible effects of such events. For instance, *Groth et al.* [2000] simulated a coronal mass ejection (CME) from the sun and the resulting effects on Earth, *Ngwira et al.* [2013] simulated the effects of a hypothetical "Carrington-type" space weather event, and *Ngwira et al.* [2014] presented simulations aimed at predicting the effects of the 23 July 2012 CME if it had been directed Earthward.

MHD models have been used to study quiet-time conditions as well. Early work such as *Wu et al.* [1981] and *Ogino et al.* [1992] simulated steady solar wind conditions, while *Raeder et al.* [1998] modeled time-dependent quiet-time conditions. Some more recent work such as *Welling and Ridley* [2010] has included quiet time periods, although that paper focused primarily on storms. However, these constitute a minority of papers in recent years, and like the storm papers, they tend to cover short periods of time.

Only a few papers to date describe MHD simulations more than a few days in duration. *Guild et al.* [2008] compared in situ plasma sheet observations with MHD output from a 2-month simulation, finding the model generally able to reproduce the gross features of the plasma sheet in a statistical sense. *Zhang et al.* [2011] analyzed the field-aligned current structures and polar cap potentials from the *Guild et al.* [2008] simulations, finding a significant under-prediction of current strength and over-prediction of CPCP. *Huang et al.* [2010] found an MHD code to be capable of reproducing the statistics of ULF waves in geosynchronous orbit over a 27-day simulation. *Juusola et al.* [2014] compared MHD derived CPCP and auroral index predictions with observations for a 1-year period using *Facskó et al.* [2016]'s 1-year global MHD simulation. That work was accomplished using a large number of short simulations run independently of each other, because the Grand Unified Magnetosphere-Ionosphere Coupling Simulation (GUMICS-4) developed by *Janhunen et al.* [2012] is a single core code. This way the simulation state was effectively re-initialized approximately every 5 hours. *Facskó et al.* [2016]'s simulations were unsuccessful at reproducing a number of aspects of the auroral oval structures, and obtained ground magnetic field perturbations that were weaker than observed by at least a factor of 5 [*Juusola et al.*, 2014]. *Facskó et al.* [2016] derived the magnetic footprints by magnetic field mapping from the Cluster SC3 using the GUMICS simulation and also using the Tsyganenko (T96) model in order to compare two methods. The study showed that the footprints determined using the GUMICS simulation agreed relatively well with the T96 empirical model, however the footprints agreed better in the northern hemisphere than the southern one during quiet conditions. *Wiltberger et al.* [2017] covers a period of nearly a month (March 20 to April 16, 2008), which was chosen because it contains a wide variety of solar wind conditions but no major geomagnetic storms. The results presented in *Wiltberger et al.* [2017] focused on field-aligned currents and cross-polar cap potential (CPCP), finding that the simulations reproduced the statistical features of the observed field-aligned current patterns but tended to produce weaker field-aligned currents and higher potentials than the Weimer05 empirical model.

Some focus on storms is no doubt appropriate due to the hazards posed by such events. However, the approach of manually selecting storm events to validate a model can be problematic. Manual selection of storm events can introduce biases since the particular storms chosen may not be representative examples. Furthermore, undue focus of validation efforts on strong storm events could result in a model that is optimized for such events at the expense of moderately disturbed or quiet conditions. This can potentially undermine the model's usefulness as a forecasting tool, since a model designed only to model storms could over-predict or under-predict activity in weakly or moderately disturbed conditions. In the case of over-prediction, this could lead to an elevated false alarm rate for storm conditions. In the case of under-prediction, it could lead to potentially significant activity being missed. In either case, it could erode confidence in the model on the part of forecasters and customers if the model appears to be useful only during times of strong activity.

If a model performs poorly during quiet time conditions, this could be symptomatic of problems that persist during disturbed periods as well. Small deficiencies in a model may in some cases be apparent during quiet time but be difficult to notice during storm time. In addition, quiet-time conditions just prior to a storm may subtly affect the dynamics of the storm itself. Therefore, improvements to a model's representation of the quiet-time magnetosphere are likely to improve its representation of storm-time dynamics as well.

In the present work, we investigate the capability of the SWMF to deliver accurate predictions of geomagnetic indices and cross-polar cap potential. We include a realistic mix of quiet and disturbed conditions by studying the entire one-month period of January, 2005, rather than a set of selected events. In addition, the use of a single continuous time period for validation reduces any errors caused by a poor initial condition (provided those errors dissipate over time). Finally, use of a single continuous run is more representative of operational forecasting usage, in which a continuous stream of real-time data is fed into the model.

We drive three different configurations of the SWMF (the details of which are described in Section 2.1) with solar wind data observed by the Advanced Composition Explorer (ACE) spacecraft. The model's input data is described in more detail in Section 2.2. The model provides magnetic field values at a number of ground stations. From these we calculate values of the geomagnetic indices Sym-H, Kp, and AL, as well as CPCP. Sym-H is the longitudinally symmetric northward component of six low-latitude magnetometers, typically regarded as a measure of ring current and other current systems. Kp (planetarische Kenziffer) is an index computed from a number of mostly mid-latitude magnetometers and is typically regarded as a general measure of global geomagnetic activity.

AL (auroral lower) is computed from the most negative northward component of a set of auroral magnetometers, and is regarded as a measure of auroral zone currents, primarily the westward electrojet. Cross polar cap potential (CPCP) is the difference between the minimum and maximum electrostatic potential over the polar cap, and provides an indication of the coupling strength between the solar wind and the magnetosphere. Details on each of these quantities are given in Section 2.3.

After obtaining observed values for the indices and calculating equivalent values from the model, we calculate metrics to measure each model configuration's ability to predict each geomagnetic index, and from these identify strengths and weaknesses of each model configuration. The specific metrics are described in Section 2.4. Results for each geomagnetic index are presented and discussed in Section 3, and conclusions given in Section 5.

2 Methodology

2.1 Model description

Figure 1. Illustration of the models (components within SWMF) and couplings in use. Arrows denote the information that is passed between the components.

The model we use consists of the BATS-R-US (Block-Adaptive Tree Solar-Wind, Roe-Type Upwind Scheme), coupled to the Rice Convection Model (RCM) and the Ridley Ionosphere Model (RIM). A schematic of the coupling is shown in Figure 1. BATS-R-US, described in *Powell et al.* [1999] and *De Zeeuw et al.* [2000], is an adaptive mesh MHD solver which solves the ideal MHD equations throughout the magnetosphere. RCM [*Wolf et al.*, 1982; *Sazykin*, 2000; *Toffoletto et al.*, 2003] models the inner magnetosphere, and RIM [*Ridley et al.*, 2003; *Ridley et al.*, 2004a] simulates ionospheric electrodynamics. Coupling is accomplished using SWMF. Couplings between the models are identified by arrows in 1, which point in the direction of information flow and are labeled with the quantities passed between the models. The couplings are as follows:

- BATS-R-US MHD delivers magnetic field and plasma moments to RCM
- RCM provides plasma density and pressure to BATS-R-US
- BATS-R-US sends current density to RIM

Name	Grid	RCM	Composition model
SWPC	SWPC	Y	Fixed
Hi-res w/ RCM	Hi-res	Y	*Young et al.* [1982a]
Hi-res w/o RCM	Hi-res	N	Fixed

Table 1. Summary of the model configurations used.

- RIM delivers electric field to BATS-R-US
- RIM delivers electric potential to RCM

This combination of models and couplings is currently being used for operational forecasting of $\frac{dB}{dt}$, Dst, and Kp at the Space Weather Prediction Center (SWPC).

We run the model in three different configurations, summarized in Table 1. The SWPC configuration is nearly identical to that used operationally by SWPC (the main differences, besides the input data being historical rather than real-time, being in what output files are written during the run). The other configurations are similar, but use a higher resolution grid and other modifications. The two grids that are used are described in detail in Section A.0.1. The switch to the higher resolution grid necessitated other modifications in order to maintain the model's performance with respect to Sym-H. First, the plasma sheet O/H mass density ratio (used in coupling between BATS-R-US and RCM) is determined adaptively based on the current values of F10.7 flux and Kp index using the empirical model from *Young et al.* [1982b], rather than using a fixed ratio as is used in the SWPC configuration. Second, a boundary condition parameter that controls how much the inner boundary density increases as cross-polar cap potential increases [described in *Pulkkinen et al.*, 2013] was reduced from 0.1 to 0.08. These changes result in Sym-H predictions that are similar to the SWPC configuration, and have minimal effect on the other quantities analyzed in this paper. Details of the model configuration, including settings for each component, are described in Appendix A.

2.2 Model execution

In order to create a dataset for statistical evaluation of the model, we ran the model for the entire month of January, 2005. We repeated this for each of the three configu-

	Min	25th percentile	Median	75th percentile	Max
IMF B_z (nT)	-27.97	-1.7	0.28	2.83	30.92
Solar wind u_x (km/s)	318	468	570	672	1055
Solar wind dynamic pressure (nPa)	0.0859	1.53	2.07	3.03	80.62
Kp	0.0	2.0	3.0	4.0	8.0
Sym-H (nT)	-112	-29	-17	-7	57
AL (nT)	-4418	-279	-123	-40	10
CPCP (kV)	6.67	27.0	63.2	77.5	1460

Table 2. Minumum, 25th percentile, median, 75th percentile, and maximum for a number of observed quantities characterizing the solar wind conditions and (observed) geomagnetic conditions during the month of January, 2005. Components of IMF and solar wind velocity are given in GSM coordinates.

rations described in section 2 of this paper. This time period was selected to support a project currently in progress to evaluate the model's capability to predict magnetospheric substorms. Sequences of substorms in January, 2005 were previously studied in *Morley* [2007] and *Morley et al.* [2009], and the period was identified as having a sufficiently large number of substorms to allow statistical analysis with regard to substorm predictions. The month was in the late declining phase of solar cycle 23. Minima, maxima, and medians of observed quantities characterizing the month are shown in Table 2. The month includes three geomagnetic storms. The first, on January 7, was the result of a coronal mass ejection (CME) indicated by a small velocity change but a large spike in proton density. The January 7 storm reached a minimum Sym-H of -112 nT. The second storm, on January 16, was the result of a CME indicated by a solar wind velocity increase from 600 to 800 km/s and a large density spike. An additional CME arrived on January 18th, before the completion of recovery from the January 16 storm. The January 16 storm reached a minimum Sym-H of -107 nT. The third storm was on January 21. The January 21 storm was the result of a CME which resulted in a solar wind speed increase from 600 to 900 km/s and a large density spike. The January 21 storm reached a minimum Sym-H of -101 nT. A final CME arrived on 31 January but did not result in a geomagnetic storm.

To simulate this month, we drive the model using solar wind velocity, magnetic field, density, and temperature, which are used to construct the upstream boundary condition of BATS-R-US. The only other input parameter is F10.7 flux, which is used by RIM in computing ionospheric conductivity [*Ridley et al.*, 2004b; *Moen and Brekke*, 1993]. In the high-resolution configuration with RCM, F10.7 is also used to compute the oxygen to hydrogen ratio via the *Young et al.* [1982a] empirical model.

Solar wind parameters are obtained from the 1-minute OMNI dataset provided by the NASA Goddard Spaceflight Center (GSFC). This is a combined dataset which includes data from multiple spacecraft, although during the time period in question the data came primarily from the ACE spacecraft. The OMNI date is provided "time shifted" to the bow shock nose using the techniques described in *Weimer and King* [2008]. We obtain F10.7 observations from http://lasp.colorado.edu/lisird/tss/noaa_radio_flux.html, which combines the historical archive available through the National Centers for Environmental Information (NCEI) with modern measurements managed by NOAA SWPC. The flux values are the 1 AU adjusted flux observed at Penticton, BC [*Tapping*, 2013].

The solar wind data receives some additional processing before being input to the model. In addition to the OMNI data, we use temperatures from the ACE spacecraft, time-shifted by 45 minutes. To simplify some of the post-processing and analysis, only the x component of velocity was used and the y and z components were set to zero. This reduces the motion of the magnetotail so that it remains near the x axis of the grid. Although the y and z components can significantly affect the orientation of the magnetotail, we expect they would have relatively little impact on the geomagnetic indices that are the focus of the present work [see e.g. *Borovsky*, 2012]. The x component of the interplanetary magnetic field (IMF) was also set to zero in order to reduce the divergence of the magnetic field in the simulation.

Gaps of less than 1 hour in the OMNI data are filled by linear interpolation. Three gaps of longer duration had to be filled in from other sources. The first of these was on 18 January from 06:11 to 13:52 UT, the second was from 7:14 UT on 20 January to 21:44 on 21 January, and the third was from 01:04 to 09:13 UT on 22 January. These were due to instrument problems that occurred with the Solar Wind Electron, Proton, and Alpha Monitor (SWEPAM) instrument on the ACE satellite in its default mode, which attempts to track the solar wind peak in energy. SWEPAM operates in a second mode approximately once every 1/2 hour, which samples most of the instrument's energy range rather than just the peak [*McComas et al.*, 1998]. The data from this secondary mode was used for solar wind density, temperature, and velocity during the gaps in the OMNI dataset. Magnetic fields for the gap periods were available at a 1-minute cadence from the ACE Level 2 data.

Since the ACE spacecraft is located well beyond the upstream boundary of the model, it must be propagated to the upstream boundary in some way. The data obtained from OMNI are provided already time-shifted to the bow shock nose and were used as-is (see https://omniweb.gsfc.nasa.gov/html/HROdocum.html for a description of the time shifting algorithm). The ACE SWEPAM data used to fill the gaps on 18-22 January were propagated to the upstream boundary by solving a system of 1-D advection equations:

$$\frac{\partial q_i}{\partial t} = u_x \frac{\partial q_i}{\partial x}. \tag{1}$$

Here, q_i denotes one of the solar parameters, and u_x denotes the solar wind velocity in the x direction. The "time shifting" method used to create the OMNI dataset [similar techniques are described in a number of papers such as *Weimer et al.*, 2003; *Weimer*, 2004; *Cash et al.*, 2016] is equivalent to solving Equation 1 using the method of characteristics.

In the present work we solve the advection equation using a second-order finite volume method with a minmod limiter and explicit Euler time integration on an evenly spaced 1000-point grid. The time step is adjusted dynamically to maintain a maximum Courant-Friedrichs-Lewy (CFL) number of 0.5. The particulars of this class of numerical schemes are described in a number of references such as *Hirsch* [2007].

Once the runs are completed, we evaluate the model configurations with regard to their ability to predict Kp, Sym-H, AL, and CPCP. Observational data for the Kp index provided by the NOAA National Geophysical Data Center (NGDC) and was obtained through the NASA/GSFC 1-hour OMNI dataset. Observational data for the Sym-H index provided by World Data Center Kyoto was obtained through the NASA/GSFC 1-minute OMNI dataset. Magnetic fields at ground-based magnetometer stations were obtained from SuperMAG [http://supermag.jhuapl.edu/ *Gjerloev*, 2012] and used to calculate the AL index as described in Section 2.3. Since no direct observation of CPCP is available, we instead use the Assimilative Mapping of Ionospheric Electrodynamics (AMIE) model, which estimates CPCP based on a number of observational datasets [*Richmond and Kamide*, 1988; *Richmond*, 1992]. The Spacepy python library [*Morley et al.*, 2011; *Morley et al.*, 2014] was used for a number of tasks including reading the MHD output and some of the observational datasets.

2.3 Predicted quantities assessed

The observed quantities assessed in this paper are all derived from ground-based magnetometers. In order to reproduce these observations with the MHD model, the magnetic fields resulting from magnetospheric and ionospheric currents are calculated at various points on the Earth's surface. This is accomplished using a Biot-Savart integral over the entire MHD domain, as well as the height-integrated Hall and Pedersen currents computed by RIM [*Yu and Ridley*, 2008; *Yu et al.*, 2010]. From these magnetic fields we obtain equivalents to the geomagnetic indices Kp, Sym-H, and AL.

The Kp index is a measure of general geomagnetic activity, and is particularly sensitive to magnetospheric convection and to the latitude of the auroral currents [*Thomsen*, 2004]. Kp is calculated from 13 magnetometer stations whose geomagnetic latitudes range from 54 to 63 degrees [*Rostoker*, 1972]. Kp is obtained from the local K (Kenziffer) index which is calculated individually for each magnetometer. The procedure for calculating local K is described in *Bartels et al.* [1939], and the procedure for calculating the planetary Kp from local K is given in *Mayaud* [1980]. Kp has historically been reported with fractional values denoted with "+" and "-" symbols, with e.g. 4+ indicating $4\frac{1}{3}$ and 4- indicating $3\frac{2}{3}$. Since the "+" and "-" notation would complicate presentation and analysis, we follow the convention used in the OMNI dataset where the fractional components are rounded to the nearest tenth, i.e., "4-"=3.7, "4+"=4.3, etc.

Although the model Kp could be computed using the model output for the 13 stations used observationally, we instead use a different set of locations. These consist of an evenly spaced ring of 24 points having a constant latitude of 60 degrees. For each of the 24 points, the local K value is calculated using the procedure described in *Bartels et al.* [1939]. The K-scale mapping for the magnetometer station Niemegk [also given in *Bartels et al.*, 1939] is applied to all stations. This choice of mapping was found by trial and error to produce the best Kp predictions. Having obtained the local K values for each of the 24 points, the Kp index is then computed as the mean of these local K values, rounded to the nearest one-third. Rather than calculating the model Kp every 3 hours as is done in the observations, the model Kp is calculated using a rolling 3-hour window, and values are output every minute. This rolling 3-hour window ends at the time of each output, so that at the time of the observations the model's rolling window coincides with the period used to calculate the observed Kp.

The AL index, introduced in *Davis and Sugiura* [1966], provides a measure of the effect of the westward electrojet on the surface magnetic field. While *Davis and Sugiura* [1966] used a set of 10 magnetometer stations, we calculate the AL index from an alternate set of magnetometers, the complete list of which is provided in the supplemental data. An identical set of magnetometer locations is used in both the model and observations. Since the Biot-Savart integrals used in the model explicitly exclude the intrinsic field of the Earth, the baseline removal step described in *Davis and Sugiura* [1966] is not necessary for the model output. For the observational data, we use data from SuperMAG which has the baseline signal removed according to the procedures described in *Gjerloev* [2012]. The remainder of the AL calculation procedure (following baseline removal) is the same for both model and observations and is implemented as described in *Davis and Sugiura* [1966].

The Sym-H index is intended to measure the strength of currents circling the Earth around the dipole axis. It is calculated from a set of near-equatorial magnetometers according to procedures described *Iyemori* [1990] and http://wdc.kugi.kyoto-u.ac.jp/aeasy/asy.pdf. Sym-H is often described as a measure of the symmetric ring current. However, it was shown [see the review by *Maltsev*, 2004, and references therein] that it contains contributions from many other current systems (magnetopause currents, cross-tail current, partial ring current, substorm current wedge) and their contributions can be signicant or even dominant during disturbed conditions [e.g. *Ohtani et al.*, 2001; *Liemohn et al.*, 2001; *Ganushkina et al.*, 2004; *Kalegaev et al.*, 2005; *Dubyagin et al.*, 2014]. Sym-H is very similar to the Dst index, differing primarily in that Sym-H uses a larger number of magnetometer stations and is calculated at a higher time resolution. *Wanliss and Showalter* [2006] showed that despite the differences in how Sym-H and Dst are calculated, Sym-H can effectively be used as a high-resolution substitute for Dst. *Katus and Liemohn* [2013] found that the difference (measured in RMSE) between Sym-H and Dst was 9.1 nT during the period 1985-2005. During the same interval, the RMSE difference between Sym-H and USGS Dst [a 1-minute cadence Dst implementation provided by the U.S. Geological Survey, described in *Gannon and Love*, 2011] was 11.0 nT. Since these very similar indices differ from each other on the scale of 9-11 nT, one could consider model predictions of Sym-H with errors less than 9-11 nT to be indistinguishable from observations.

As with Kp, SWMF provides output for Sym-H. Rather than calculating Sym-H using the set of surface magnetometers used in the observations, SWMF calculates the magnetic perturbation in the direction of the magnetic pole via a Biot-Savart integration of all currents within the MHD domain about a point at the center of the Earth. Since the magnetic field is calculated at the center of the Earth, the step of averaging in longitude described in *Iyemori* [1990] is not needed. This methodology was validated against storm-time observations in *Rastätter et al.* [2011].

Cross-polar cap potential (CPCP) is the difference between the maximum and minimum electric potential over the polar cap. It is dependent on the solar wind electric field, the size of the open flux region connecting the polar cap to the magnetopause, and the magnetospheric dynamics that determine the strength of the coupling between those two regions [*Bristow et al.*, 2004; *Lockwood and Morley*, 2004; *Milan*, 2004]. Observationally, CPCP must be obtained indirectly, and for the present work we used output from the AMIE model [*Richmond and Kamide*, 1988; *Richmond*, 1992], which computes a potential pattern through an expansion of basis functions chosen by fitting to observations from magnetometers, radar, and spacecraft. CPCP in the model is obtained from the potentials computed by the RIM ionosphere model.

2.4 Assessing prediction quality

To give an overall picture of the model's agreement with the observations we calculate accuracy and bias metrics for the entire month, as well as probability distributions, for each predicted quantity. Given a set of observations x_i and corresponding predictions y_i, the error is given by

$$\epsilon_i = y_i - x_i. \tag{2}$$

Mean error is defined as

$$\bar{\epsilon} = \frac{1}{n}\sum_{i=1}^{n}\epsilon_i. \tag{3}$$

$\bar{\epsilon}$ is a measure of bias; a positive value indicates that the model over-predicts on average, while a negative value indicates that the model under-predicts on average. An unbiased prediction will be indicated by $\bar{\epsilon}$ at or near zero.

The root mean squared error (RMSE),

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\epsilon_i^2}, \tag{4}$$

provides a measure of the average discrepancy between predictions and observations, independent of the sign of the error. RMSE is always positive and, like $\bar{\epsilon}$, has the same units as the input data. A smaller value for RMSE indicates a more accurate prediction.

Both mean error and RMSE are computed from a mean, and hence their uncertainty can be computed using the formula for computing the uncertainty of a mean:

$$\sigma_{mean} = \frac{\sigma}{\sqrt{n}}, \tag{5}$$

where n is the number of points, and σ is sample standard deviation of the points from which the mean is computed [*Taylor*, 1997]. Taking σ as the standard deviation of all the points ($std(x)$), the uncertainty of RMSE is estimated by

$$\sigma_{RMSE} = \sqrt{\frac{std(\epsilon^2)}{\sqrt{n}}} \tag{6}$$

and the uncertainty of mean error is estimated by

$$\sigma_{\bar{\epsilon}} = \frac{std(\epsilon)}{\sqrt{n}}. \tag{7}$$

All of the above metrics require a set of observations x_i and corresponding predictions y_i. Since the model is configured to produce output at specific times that may or may not coincide with the observations, linear interpolation of the model output is used to obtain values that correspond to the exact time of the observations. In the case of Kp, the model produces output at a much higher time resolution than the available observations, and this process results in a set of Kp predictions which correspond with the observations in terms of the number of values and in terms of the time range of the magnetometer data from which those values are derived.

Summarizing bias or accuracy with a single number provides a useful summary of a model's capabilities, but this single number can be misleading, particularly if the quantity being predicted has an asymmetric distribution. In the case of Kp, the pseudo-logarithmic scale complicates interpretation further. To get a more detailed picture of the model's predictive ability than is possible using mean error and RMSE, we compute probability density functions (PDFs) or distribution functions for each predicted quantity and its error. A PDF (or distribution function) of a quantity is a function that gives the relative likelihood that the variable will have a given value. Ideally, the distribution of the model values for a predicted quantity should be identical to the distribution of the observations for that quantity. Systematically biased predictions will result in a curve that is shifted right or left relative to the observations. When the shape of the PDF differs, this may indicate a tendency to over-predict or under-predict under a specific set of conditions. For the distribution of an error, the ideal case is a narrow, symmetric peak centered at zero. Bias in the model results in an off-center or asymmetric peak in the error distribution. An inaccurate prediction is indicated by a broad peak.

For this paper we approximate PDFs using kernel density estimation [*Parzen*, 1962]. This approximates the underlying PDF from a finite set of observations by smoothing with a kernel function, in this case a Gaussian. The bandwidth (the width of the Gaussian kernels) is determined for each PDF using Scott's Rule [*Scott*, 2015]. The specific implementation for the kernel density estimates is that of the Scipy software library [*Jones et al.*, 2001, updated frequently].

3 Results

The mean error and RMSE of several predicted quantities were calculated for the entire month for each model configuration; these and their associated uncertainties are shown in Table 3. In addition to mean error and RMSE, we also give a normalized RMSE for each predicted quantity, which is computed by dividing the RMSE by the standard deviation of the observed values. By normalizing the RMSE values by the spread of the observational data, we obtain a unitless accuracy metric. This provides a means to compare between RMSE values for disparate quantities. The normalized RMSE values seem to suggest that the model predicts Kp better than any other quantity. However, this is likely due to the fact that Kp is based on a 3-hour maximum of magnetic field variations, and is therefore insensitive to variations of shorter duration or magnitude. The other predicted

Metric	Model configuration: SWPC	Hi-res w/ RCM	Hi-res w/o RCM
	Kp metrics		
Mean error	0.68 ± 0.05	0.84 ± 0.06	−0.17 ± 0.07
RMSE	1.1 ± 0.3	1.3 ± 0.3	1.1 ± 0.4
Normalized RMSE	0.6 ± 0.2	0.8 ± 0.2	0.6 ± 0.2
	Sym-H metrics		
Mean error (nT)	−7.36 ± 0.07	−3.99 ± 0.08	21.54 ± 0.09
RMSE (nT)	17 ± 2	18 ± 2	29 ± 3
Normalized RMSE	0.77 ± 0.09	0.86 ± 0.09	1.4 ± 0.1
	AL metrics		
Mean error (nT)	71 ± 1	15 ± 1	123 ± 1
RMSE (nT)	250 ± 40	230 ± 40	270 ± 40
Normalized RMSE	0.9 ± 0.1	0.8 ± 0.1	1.0 ± 0.1
	CPCP metrics		
Mean error (kV)	2.5 ± 0.2	14.9 ± 0.2	14.5 ± 0.2
RMSE (kV)	46 ± 10	47 ± 9	48 ± 9
Normalized RMSE	0.8 ± 0.2	0.8 ± 0.1	0.8 ± 0.1

Table 3. Metrics for all quantities and all model configurations, given as the value ± one standard error.

quantities have 1-minute time resolutions, so the prediction quality metrics for those quantities reflect errors in predicting high-frequency oscillations that are removed in the calculation of Kp. Note all of the metrics in Table 3 are calculated for the entire month, and as a result are likely dominated by the quiet-time tendencies for each quantity.

The results are discussed in detail for each predicted quantity in sections 3.1-3.4, and differences between quiet and active periods are addressed where appropriate. The figures in the following sections use a common color scheme to identify results from the different model configurations. The SWPC configuration is shown in red, the high-resolution grid with RCM is shown in orange, and the high-resolution grid without RCM is shown in blue. Observations, where applicable, are shown as a thick, light blue curve.

3.1 Kp

The mean error and RMSE metrics for Kp are shown in Table 3. These values represent deviations on the pseudo-logarithmic Kp scale, and hence are dimensionless. Kp predictions from the high-resolution configuration without RCM have the smallest RMSE (1.1), which indicates that these predictions have on average the best accuracy of the three model configurations, but the uncertainties in these RMSE values are large enough that the difference may not be significant. The high-resolution configuration without RCM also has the lowest bias with respect to Kp prediction, with a mean error of -0.20, indicating a slight under-prediction. Both configurations with RCM have positive biases, indicating over-prediction, and the biases are of greater magnitude than those for the configuration without RCM. Although the metrics seem to suggest that the configuration without RCM performs the best, they are misleading in this case as will be discussed later in this section when the distributions of Kp are examined in detail.

Figure 2a shows the probability distribution of Kp error for the three model configurations. The Kp error curve for the configuration without RCM is nearly centered about zero, indicating that the errors are relatively unbiased. The half width at half max of that curve is about 1, also consistent with the RMSE of 1.1 from Table 3. The Kp error curves for the SWPC configuration and the high resolution with RCM configuration are both centered to the right of zero. This indicates that these configurations tend to over-predict Kp, consistent with the positive mean errors shown in Table 3 for those configurations.

The probability distributions of the actual Kp values are shown in Figure 2b. In addition to distributions obtained from the three model configurations, the observed distribution is shown as a thick, light blue curve. The observations have a mode at Kp = 3.3. The two models that incorporate RCM (SWPC and high-resolution with RCM) reproduce the observed distribution fairly closely, having peaks between 3 and 4 (reasonably close to the observed peak at Kp = 3.3). However, they under-predict how often Kp values less than 2 will occur compared to the observations. The model configuration without RCM reproduces the observed distribution more closely in the Kp = $0-2$ range than do the configurations with RCM. However, the Kp distribution from the without-RCM configuration also has its peak to the left of the observations, and indeed the entire distribution seems to be shifted to the left. The fact that the configuration without RCM agrees with the observations more closely in the low Kp range seems to be merely a side-effect of this leftward shift. This means that the configuration without RCM produces more realistic quiet-time Kp values, but does so at the expense of accuracy during disturbed conditions.

Figure 2. Probability density of Kp error (**a**) and Kp itself (**b**) for all model configurations during 1-31 January 2005. Distributions for the three model configurations are plotted as colored curves: SWPC in red, high-resolution with RCM in orange, and high-resolution without RCM in blue. Observations are shown as a thick, light blue curve.

Figure 3 shows distributions of Kp similar to the one in Figure 2b, but broken down into bins covering specific ranges of observed Kp. The range of observed Kp values in each bin is labeled using the notation $[\mathrm{Kp}_{min}, \mathrm{Kp}_{max})$, indicating that the observed values in the bin start with Kp_{min} and go up to but do not include Kp_{max}. For each bin, the model output is shown for the points in time corresponding to the observational data in that bin. The number of data points per bin range from 40 (in the $\mathrm{Kp} \in [6, 9)$ bin) to 200 (in the $\mathrm{Kp} \in [3, 4)$ bin). Note that the $\mathrm{Kp} \in [6, 9)$ bin covers a greater Kp range than the others; this was done to ensure the bin contains a sufficient number of points for analysis.

The binned distributions of Figure 3 provide a sense for how the model performance varies with the amount of geomagnetic activity. For the lowest Kp bins ($[0, 1)$ and $[1, 2)$), all of the models produce distributions shifted to the right compared with the observations, indicating a tendency to over-predict Kp during times of low activity. The over-prediction appears to be least severe for the no-RCM configuration, and most severe for the high-resolution grid with RCM. The high-resolution grid without RCM matches the observations fairly closely in the $\mathrm{Kp} \in [2, 3)$ bin, but tends to under-predict for all higher Kp bins. The SWPC and Hi-res with RCM configurations continue to over-predict Kp up to the $\mathrm{Kp} \in [3, 4)$ bin. For the higher Kp values these configurations seem to produce relatively unbiased predictions.

Figure 3. Probability density of Kp for observations and for all model configurations, binned by observed Kp. Tick labels on the y axis show the range of observed Kp values contained in each bin in the form $[\mathrm{Kp}_{min}, \mathrm{Kp}_{max})$. The light blue curve within each bin shows the probability density of Kp for the observations within that bin, while the colored curves show the distribution of predictions for each model corresponding to the times of the observations falling in the bin using the same color scheme as Figure 2.

Figure 4 shows the mean error for each of the Kp bins. The x axis shows the Kp bins using the same notation as Figure 3. The no-RCM configuration has positive mean error (indicating over-prediction) for low Kp, but the mean error decreases with increasing Kp, reaching zero around Kp = 2, and having negative values thereafter (indicating under-prediction). The two configurations with RCM (red and orange curves) also have a positive mean error for low Kp, with similar values to each other but greater magnitude (stronger bias) than that of the no-RCM configuration. The mean errors for these also decrease as Kp increases, but at a slower rate than the no-RCM configuration. For the configurations with RCM the mean error remains positive up to Kp = 5, but turns negative for Kp > 6.

Figure 4. Mean error for each Kp bin. The ranges for each bin are denoted in the x axis labels in the form [Kp_{min}, Kp_{max}). The color scheme follows the previous figures. All the configurations over-predict low values of Kp, and the without-RCM configuration under-predicts the higher Kp values.

These results are similar to those of *Glocer et al.* [2016], which evaluated SWMF and several other models based on their predictions of local K. *Glocer et al.* [2016] did not include bias or accuracy metrics in their results, but in their supplemental data they provided distributions of predicted K for several values of observed K. From these, an unbiased prediction is apparent for observed K = 4, a under-prediction occurs for observed K = 6, and even greater under-prediction for observed K = 8. Thus the downward trend in bias is apparent as K increases in the *Glocer et al.* [2016] results, similar to the present work. The *Glocer et al.* [2016] results do not seem to show the positive bias that we see at lower values of Kp; this difference may be due to the *Glocer et al.* [2016] results being based on a study of storm events while our results include a considerable amount of quiet periods, as well as the difference in using individual magnetometer stations in that study versus the global Kp index in the present work.

The model's ability to predict Kp during disturbed periods is notably improved with the addition of RCM, primarily during disturbed periods. This suggests that the differences between the model without RCM and those with (SWPC and Hi-res with RCM) are due primarily to differences in those current systems that are affected by the coupling with RCM, specifically the azimuthal currents that are modeled directly by RCM, and the Region 2 field-aligned currents which are driven by inner magnetosphere pressure gradients affected by the coupling.

3.2 Sym-H

From the Sym-H results in Table 3, it is apparent that the two configurations using RCM (SWPC and Hi-res with RCM) predict Sym-H more accurately than the configuration without RCM. This is indicated by the comparatively low error (measured by RMSE) and bias (mean error closer to zero) relative to the configuration without RCM. The SWPC configuration predicts Sym-H with a slightly lower RMSE but a higher mean error than the high-resolution configuration with RCM. The configuration without RCM tends to over-predict Sym-H by 21.54 nT. The two configurations with RCM under-predict, but do so with a much lower magnitude (by a factor of 3-5) than the configuration with RCM.

Comparing these values of mean error and RMSE to the difference between Sym-H and similar indices gives a sense for whether the metrics indicate a good quality prediction. As mentioned earlier, *Katus and Liemohn* [2013] found discrepancies on the order of 9-11 nT between Sym-H and two similar indices. Therefore, Sym-H predictions with an RMSE of less than about 9-11 nT might be considered to be of good quality. The predictions from all three of our model configurations exceed 11 nT, but the two configurations with RCM exceed this threshold by only 55-65%, while the configuration without RCM exceeds it by 160%.

The probability distribution of Sym-H error (Figure 5a) shows a similar tendency as the metrics with regard to bias. The two runs with RCM appear largely similar to each other. Both are centered around zero (indicating an unbiased prediction), and have a half width at half maximum of about 15 nT. The run without RCM is centered around 15 nT, indicating a clear positive bias.

The distribution of Sym-H itself is shown the Figure 5b. The underlying cause for the positive bias of Sym-H from the no-RCM configuration is clearly apparent: It tends to produce Sym-H values near zero (as indicated by the high probability density at that point), while the observed distribution peaks around -20 nT and a long tail extending to -120 nT. The two configurations with RCM, on the other hand, produce a distribution that is largely similar to the observations.

A notable exception is the part of the distribution corresponding to Sym-H greater than 10 nT, where the configuration without RCM seems to produce a more realistic Sym-H distribution than the configurations with RCM. The observed distribution shows a small but significant probability for positive values of Sym-H going as high as 15 nT on Figure 5. The configuration without RCM appears to capture the outer part of this area (5-15 nT) fairly accurately. The two configurations with RCM, on the other hand, predict positive Sym-H values at a much lower rate than occurs in the observations, as evidenced by the near-zero Sym-H probabilities between 5 and 15 nT for those configurations.

Figure 5. Probability density of Sym-H error (**a**) and Sym-H itself (**b**) for all model configurations. The color scheme follows the previous figures. The two configurations with RCM reproduce the observed Sym-H fairly well, while the one without RCM tends to produce Sym-H values near zero regardless of conditions.

Figure 6 shows time series of Sym-H during the storms on 7 and 21 January. For both of these storms, the configurations with RCM make reasonably good predictions of Sym-H, while the configuration without RCM produces very little Sym-H response except for some oscillations immediately following the initial disturbances. The two configurations with RCM, on the other hand, produce reasonably good approximations of the observed Sym-H response. These warrant further examination.

For the 7 January storm, the two configurations with RCM produce a minimum Sym-H of around -160 nT, while the observed Sym-H reached a minimum of -100 nT. Thus the model Sym-H deviates from the observations by about 50% at the time of greatest disturbance. The models recover gradually over the course of about a day, at which point they are again close to the observed Sym-H. For the 21 January storm, the configurations with RCM produce a Sym-H curve that descends more sharply than the observations and rapidly reaches a minimum of -120 nT, again stronger than the observed minimum. In this case, however, the Sym-H from the configurations with RCM recovers rapidly, with the high-resolution configuration briefly becoming less negative than the observed Sym-H (from about 22:00 UT on 21 January to about 03:00 UT on 22 January) before descending again to match the observations. For the 21 January storm it took about 2 days (until 00:00 UT on 24 January) to recover, but in this case the model output (for the configurations with RCM) followed the observations closely throughout the recovery.

Figure 6. Sym-H time series for the storms on 7 Jan (panel a) and 21 Jan (panel b). The color scheme is the same as the previous figures. The model configurations with RCM produce stronger (by 20-50%) Sym-H responses than the observations, while the configuration without RCM produces little response to the storms.

The tendency of the configurations with RCM (SWPC and Hi-Res w/ RCM) to miss positive Sym-H values previously noted in Figure 5 is apparent in both time series shown in Figure 6. In the case of the 21 January storm, a storm sudden commencement (SSC) is apparent. The configuration without RCM reproduces the observed Sym-H signature resulting from the SSC quite well, but the two configurations with RCM severely under-predict the magnitude of the SSC oscillations. A possible explanation for this is that the inner magnetosphere currents produced by RCM counteract the effects of magnetopause currents to a greater degree than occurs in reality. This reduces the influence of such currents on the surface magnetic fields and in turn the frequency and magnitude of positive Sym-H values as seen in Figure 5.

The time series plots of Sym-H show considerable improvement in Sym-H predictions over some earlier results such as *Ganushkina et al.* [2010] in which SWMF predicted Sym-H with approximately correct magnitudes but with an approximately 6-hour delay compared to the observed Sym-H. A similar improvement can be seen in other work such as *Liemohn et al.* [2013] and in some (though arguably not all) of the Dst time series plots in *Rastätter et al.* [2013].

The stark difference in Sym-H predictions with and without the RCM component highlights the importance of the inner magnetosphere model in producing realistic ring current dynamics. The inner magnetosphere model can also, through coupling with the MHD solver, affect mid-tail currents to which Sym-H is sensitive, as evidenced by increased tail stretching in MHD models when coupling to an inner magnetosphere model is used [e.g. *Welling et al.*, 2015; *Pembroke et al.*, 2012]. That SWMF predicts Dst (similar to Sym-H) better when a ring current model is used has been shown previously in *Rastätter et al.* [2013]. Changing the MHD grid resolution, on the other hand, seems to have relatively little effect on Sym-H.

3.3 AL

Table 3 shows that the mean error in AL is positive for all configurations, indicating a tendency to over-predict AL. Note that AL has negative values during times of high activity, so over-prediction of AL implies under-prediction of geomagnetic activity. Of all the model configurations, the high-resolution grid with RCM exhibits the lowest mean error for AL. The RMSE values are comparable for all three model configurations, falling within the uncertainty bounds of each other. The RMSE values for all of the models are much larger than the mean error, suggesting that random errors rather than bias are the main contributor to the RMSE values.

Figure 7. Probability density of AL error (**a**) and AL itself (**b**) for observations and for all model configurations. The color scheme follows the previous figures. The distribution is shown on a logarithmic scale due to the importance of the wings of the distribution. All three model configurations capture the overall shape of the distribution, but under-predict the probability of large negative values.

The distribution of error in AL is shown in the Figure 7a. Because the distribution is characterized by a long tail, it is plotted on a logarithmic scale. All three configurations peak around zero, but the wings of the distributions are asymmetric, with higher probabilities in the positive direction than the negative. This asymmetry is apparently responsible for the positive biases shown in the AL section of Table 3. The asymmetry is most severe for the high-resolution configuration with RCM, and least severe for the high-resolution configuration with RCM. The fact that the curves peak near zero suggests that the model produces fairly unbiased AL predictions most of the time, but the asymmetry indicates an occasional tendency toward over-prediction.

The distribution of the AL values themselves is shown in Figure 7b. All of the model configurations peak just to the left of zero, similar to the observations. At the same time, they under-predict the probibilities of the more negative AL values. The high-resolution grid with RCM under-predicts less severely than the other configurations. As a result, the high-resolution grid with RCM comes somewhat closer to reproducing the observed distribution. The under-prediction of the frequency of strongly negative values is probably the main cause of the biases apparent in the AL section of Table 3 and Figure 7a. It's worth noting that positive AL values are under-predicted by all of the models, and less severely by the configuration without RCM. This may be related to the results for Kp, where the no-RCM configuration performed better than the others during times of low activity.

The fact that the error curves peak near zero (Figure 7a) suggests that the model configurations all tend to produce realistic quiet-time conditions. The wings in the error distributions suggest less accurate predictions during times of higher activity. At the same time, all the model configurations under-predict how often the strongest negative AL values will occur (Figure 7b). This implies that the model produces a weaker westward electrojet current during disturbed periods than occurs in the observations. Since the westward electrojet is often associated with substorms [*Akasofu and Yoshida*, 1966], this suggests that the model under-predicts the magnitude of substorm-related field aligned currents.

3.4 CPCP

The errors for CPCP are calculated relative to the AMIE model [*Richmond and Kamide*, 1988; *Richmond*, 1992]. In the CPCP section of Table 3, all three SWMF configurations show positive mean error for CPCP compared to AMIE, indicating over-prediction. The SWPC configuration over-predicts only slightly, while the two high-resolution configurations over-predict more significantly. All three configurations have an RMSE that well exceeds the mean error, indicating that the errors in CPCP are not dominated by a systematic bias in one particular direction.

Probability distributions of CPCP error are shown in Figure 8a. All of the error distributions have peaks to the right of zero (around 20-30 kV), consistent with the positive mean errors reported for CPCP in Table 3. The peaks are centered 5-15 kV higher than the mean errors shown in Table 3, perhaps due to the long, thin tail of negative errors found in all three distributions.

Figure 8. Probability densities of CPCP error relative to the AMIE model (**a**) and of CPCP itself (**b**) for all model configurations. The color scheme is the same as the previous figures. These plots show that all of the model configurations over-predict CPCP.

The distribution of CPCP itself is shown in Figure 8b. The probability density of AMIE outputs (thick, light blue curve) peaks around 25 kV, while the model configurations all peak around 50-60 kV. This results in the models overestimating CPCP on average, as was seen in Table 3. The CPCP distributions obtained from all three models have half widths at half max of around 45 kV, slightly greater than the width of the observed distribution.

Figure 9. Probability density of CPCP for observations and for all model configurations, binned by observed CPCP. Tick labels on the y axis shown the range of observed CPCP values contained in each bin in the form $[CPCP_{min}, CPCP_{max})$. Probability distributions corresponding to each bin are plotted following the same color scheme used in previous figures. The model tends to over-predict CPCP during quiet times, but under-predict during the most active times.

Figure 9 shows distributions of CPCP, binned by observed CPCP. The range of observed CPCP values in each bin is labeled using the notation $[CPCP_{min}, CPCP_{max})$, much like Figure 3. From these it is immediately clear that all three models over-predict CPCP during quiet times, but under-predict during active times. This pattern is similar to what occurred for Kp, except that the configuration without RCM no longer stands out from the others.

Discrepancies between modeled and observed CPCP could be attributed to a number of possible underlying causes, including strength and location of field-aligned currents, ionospheric conductivity, and ionospheric outflow. The field-aligned current structure and conductivity both affect the potential through Ohm's Law, $\mathbf{J} = \sigma\mathbf{E}$, where the potential is proportional to the current and inversely proportional to conductivity. Thus, over-prediction of the potential (which occurs primarily during quiet time) indicates either over-prediction of field-aligned current strength, or under-prediction of the conductivity. Conversely, under-prediction of the potential (which occurs primarily during active times) indicates either under-prediction of the field-aligned current strengths or over-prediction of the conductivity.

The conductivity connection may also indicate a discrepancy in rate of outflow fom the ionospheric boundary. CPCP has been shown to decrease as heavy ion outflow from the ionosphere increases [*Winglee et al.*, 2002; *Welling and Zaharia*, 2012], so the fact that the models over-predict CPCP could be an indication that the model is under-predicting such outflow. This could be addressed through tuning of the inner boundary condition parameters, but such tuning is complicated by the fact that the outflow is itself dependent on CPCP [*Winglee*, 2000; *Welling and Liemohn*, 2014] and is likely to affect other aspects of the model such as tail dynamics, ring current, and the Sym-H values that are predicted [*Kronberg et al.*, 2014; *Welling and Liemohn*, 2016]. First-principles models of ionospheric outflow provide an alternative, but at present they are too computationally expensive for long-period runs such as those described in the present work.

4 Discussion

The relatively good accuracy achieved by the model implies a reasonably good model of the magnetospheric currents that affect the various observed quantities, including the dependency of those currents on solar wind driving and other aspects of the dynamics. Furthermore, the similarities between the results for the two highest resolution runs suggests that the model configuration is near grid convergence with regard to the predicted quantities examined in this paper. A notable exception is the AL index, where a larger difference can be seen. This could be due to the high-latitude current structures to which AL is sensitive, which may require a higher resolution in order to be fully resolved.

It's worth noting that the high-resolution configuration with RCM differs from the SWPC configuration not only in the grid but also its use of the *Young et al.* [1982a] empirical composition model in the coupling between BATS-R-US and RCM. This means that we cannot definitively attribute differences in predictions from those two configurations to the difference in grid resolution. Another limitation of these results is that the data come from a single one-month period, so any dependence of the results on season, such as those found by [*Juusola et al.*, 2014], or solar cycle will not be apparent.

The fact that Sym-H is predicted more accurately when RCM is used is expected because RCM simulates current systems to which Sym-H sensitive. These same current systems are likely responsible for improving the Kp distribution as well. Kp can be directly influenced by the current systems that affect Sym-H, particularly during times when the strength of the currents are rapidly changing. At the same time, the Region 2 field-aligned currents, to which Kp is also sensitive, are driven in part by the kinds of inner magnetosphere currents that are modeled by RCM. This has been shown theoretically by *Vasyliunas* [1970] and demonstrated using an inner magnetosphere model by *Zheng et al.* [2006] and *Zheng et al.* [2008]. The mean error and RMSE metrics for Kp seem to suggest a detrimental effect of RCM, but this is due to the quiet-time overprediction Kp being masked by an overall reduction in the magnitude of Kp due to the lack of a ring current.

Since the model over-predicts both Kp and CPCP during quiet times, it seems that there may be a common cause (or causes) behind the discrepancies in those quantities. Both Kp and CPCP are sensitive to middle and high latitude ionospheric state and dynamics (particle precipitation, conductivity, and currents). One possible underlying cause of these discrepancies is the model of ionospheric conductivity, which directly affects CPCP and affects Kp through the current structure. In the present model, the ionospheric conductivity is obtained from a number of empirical relationships. The range of validity for these empirical relationships can easily be exceeded during execution of an MHD model under realistic conditions, and in fact were exceeded during the month in question. *Welling et al.* [2017] identifies the range of validity for these models in terms of solar wind electric field to be from -1.84 mV/m to 2.30 mV/m. Solar wind electric field is defined in that paper as $u_x B_z$, where u_x is the solar wind velocity in GSM coordinates and B_z is the IMF magnetic field in the GSM z direction. $u_x B_z$ for January, 2005 ranged from -28.6 mV/m to 25.2 mV/m, roughly an order of magnitude greater than the valid range listed in *Welling et al.* [2017]. The observational data used to construct the empirical conductivity model used in RIM came from solar flux observations from 1985-1990 and magnetometer data from a one-month period of January, 1997 [*Ridley et al.*, 2004b; *Moen and Brekke*, 1993]. Construction of a more comprehensive empirical model by including more recent data would certainly be possible. Such an improved conductance model might result in better representation of auroral current systems and, in turn, indices and other observable quantities that are sensitive to them.

Like the present paper, *Wiltberger et al.* [2017] found $\frac{1}{4}$ R_e to be sufficient resolution for resolving certain aspects of magnetospheric dynamics. They compared field aligned currents for a one-month run of the Lyon-Fedder-Mobarry (LFM) MHD model, and compared the results with the *Weimer* [2005] empirical model. They presented results using three different grid resolutions, the finest of which had cell sizes between $\frac{1}{4}$ and $\frac{1}{2}$ R_e in the inner magnetosphere, similar to the SWPC grid used in the present work. They found that the relationship between field-aligned currents and CPCP was very similar between the two highest resolution grids, and concluded that the model was approaching a common solution at those resolutions. However, the results they reported were based on time averages for the entire run, so under-resolved transient features might not affect the results significantly. The indications in the present work are that the greatest magnitudes of the AL index are under-predicted, and these correspond with transient phenomena.

Wiltberger et al. [2017] also found that LFM under-predicted field-aligned current strength and over-predicted CPCP compared to the *Weimer* [2005] model. This could be explained by an under-prediction of ionospheric conductivity in that model. Analyzing field-aligned current strength in SWMF might shed some light on the problem of ionospheric conductivity, but such an analysis is beyond the scope of the present paper. Nonetheless, the results of the present work, like *Wiltberger et al.* [2017], suggest that ionospheric conductivity is an area for improvement.

5 Conclusions

This work shows the strengths and limitations of the SWMF with regard to prediction of geomagnetic indices and CPCP. By testing a one-month period with three different model configurations, we have accumulated a sufficient quantity of data to make statistical comparisons with observations under a variety of conditions.

We find that the model does an excellent job of predicting the Sym-H index. With RCM turned on, the model predicts Sym-H with RMSE values of 17-18 nT, only 50-60% larger than the observational uncertainty for that index. The model predicts the Kp index well during storm conditions, with absolute mean errors of less than one for Kp values above 3. During quiet time though, it consistently over-predicts Kp, with all configurations over-predicting by at least 1 Kp unit on average. An over-prediction of quiet-time activity is also apparent in the model's prediction of CPCP, with mean errors between 2.5 and 14.9 kV. The model tends to under-predict the magnitude of the AL index, with mean errors between 15 and 230 nT.

Of the quantities assessed in this paper, the model performs best at predicting Sym-H, and least well at predicting AL. That the model predicts Sym-H poorly without RCM is an expected exception to this. The model's relatively poor performance in predicting AL indicates problems in capturing the structure of auroral-zone currents. A better model of ionospheric conductivity would probably be the most effective way to improve these in the near term, although better predictions of dynamics affecting the field-aligned current structure are needed if the auroral-zone observations are to be predicted to a high degree of accuracy. Depending on what changes are made, such improvements may also reduce the problem of over-predicting Kp during quiet time as well, since Kp is also sensitive to auroral-zone dynamics.

Increasing the grid resolution compared with the SWPC grid had relatively little effect on prediction quality. For all four predicted quantities, the model's predictive accuracy, measured by RMSE, changed by insignificant amounts, as indicated by the error bounds of each RMSE value. There are some indications that the increased grid resolution may have improved the model's prediction of the more extreme values attained by the AL index, however. This implies that the auroral currents during disturbed periods are improved by the increased grid resolution.

Unlike the grid resolution, the presence or absence of an inner magnetosphere model has a dramatic effect on the Sym-H results, with the distribution of Sym-H taking a notably different shape and width when RCM was turned off, and a resulting change in RMSE that far exceeded the uncertainty bound (29 nT without RCM versus 18 nT with). The Kp and AL indices are also affected by the use of RCM, though to a lesser degree than the Sym-H index. Like the Sym-H index, the predictive skill for the AL index was improved by the use of RCM, with RMSE increasing from 230 nT to 270 nT when RCM was turned off. RMSE proved to be somewhat misleading as a measure of accuracy for Kp. RMSE decreased notably when RCM was turned off, which ordinarily would indicate better accuracy. However, a careful examination of the dataset reveals that the accuracy only improved during relatively quiet periods (Kp $\leq$ 2), while the accuracy during the most disturbed intervals was noticeably worse. CPCP was the only quantity not affected significantly by the use of the inner magnetosphere model, with only a very small change in RMSE when RCM was turned off.

The datasets produced for this paper can be utilized for a number of possible follow-on projects. The MHD solution can be used to reproduce spacecraft observations, which will enable an assessment of the model's ability to predict magnetic fields in the inner magnetosphere, and locations of the bow shock and magnetopause. As mentioned in the previous section, the field-aligned current structure can be analyzed in detail in order to determine what aspects of the field-aligned currents the model is able to capture. Finally, the model output can be analyzed to identify signatures of substorms, in order to assess how well the model reproduces their timing and dynamics.

It may be useful to conduct additional work like this covering other time periods. This would make it possible to assess variations depending on season or solar cycle. The resulting datasets could also be analyzed in combination, which would produce results with increased statistical significance and enable more detailed statistical analysis.

A: Model configuration details

A.0.1 MHD solver

For all of the runs in this paper we use BATS-R-US [*Powell et al.*, 1999] to solve the ideal MHD equations. The flux scheme is Sokolov's Local Artificial Wind flux [see *Sokolov et al.*, 2002], and a Koren's third order limiter [*Koren*, 1993] with beta=1.2. Cross-sections of the two MHD grids are shown in Figure A.1. These cross-sections are in the X-Z plane through the origin; the grids are symmetric such that Y-Z cuts through the origin would look identical. Both are Cartesian grids in GSM coordinates, with the cell size varied using adaptive mesh refinement (AMR). The outer boundaries form a cube 256 Earth radii (R_e) in width. The grids are offset in the x direction so that they extends 32 R_e sunward of the Earth and 224 R_e tailward. In the y and z directions the grids are centered around the Earth, extending 128 R_e from the Earth along each of those axes. An inflow boundary condition populated with time-dependent solar wind data is used on the boundary located at x=32 R_e, while the opposite face (at x=-224 R_e) uses an outflow boundary condition. The remaining outer boundaries use a zero-gradient boundary condition.

Figure A.1. X-Z cuts showing cell sizes in the two MHD grids. Left panel shows the grid used for the SWPC configuration (minimum cell size of 1/4 R_e, while the right panel shows the higher resolution grid used for the other two runs (minimum cell size of 1/8 R_e).

While the two grids are identical in their overall extent, their resolutions differ significantly. The SWPC grid (left panel of Figure A.1) has cell sizes ranging from 8 R_e at the outflow boundaries to 1/4 R_e within a 16 R_e diameter cube surrounding the Earth. The cell size of the high-resolution grid (right panel of Figure A.1) varies from 8 R_e at the outflow boundaries to 1/8 R_e near the Earth. The refined regions are the same as those used in *Welling and Ridley* [2010]. A 1 R_e cell size is used in a region around the x axis extending from the inflow boundary to 112 R_e down-tail, while the near tail region from 8 to 20 R_e down-tail is resolved to 1/4 R_e. The minimum cell size occurs within an 8 R_e wide cube surrounding the Earth, from which a 2.5 R_e sphere is excluded from the MHD grid; this region is modeled through coupling to the ionospheric model described in the next section. The SWPC grid contains around 1 million cells, while the high-resolution grid contains 1.9 million cells.

A.0.2 Inner magnetosphere

In the inner magnetosphere, transport by gradient and curvature drift becomes more important to the plasma motion, making the ideal MHD approximation inaccurate there [*Heinemann and Wolf*, 2001]. We model this region using the Rice Convection Model (RCM). By averaging out the gyro and bounce motion, this model treats the inner magnetosphere plasma as a fluid that drifts across field lines.

Unlike the MHD solver, the RCM breaks the plasma population into bins according to an energy invariant, and each energy invariant is treated as a separate fluid. In addition, oxygen, hydrogen, and electrons are treated as separate species. Since the MHD solver is being run in single-fluid mode, the coupling between the two codes must divide the MHD fluid into hydrogen and oxygen. The operational model used by SWPC accomplishes this by using a fixed ratio of 10% oxygen and 90% hydrogen by number density. However, we found that with the higher resolution grid this configuration resulted in poorer quality Sym-H predictions than with the lower-resolution grid. We were able to address this problem by replacing the fixed oxygen to hydrogen ratio with one computed using the empirical plasma sheet composition model from *Young et al.* [1982b]. The *Young et al.* [1982b] model gives relative quantities of oxygen and hydrogen as a function of F10.7 and Kp. In our implementation, F10.7 values are provided through an input file, and Kp is obtained from the MHD solver. The results presented in this paper use the fixed ratios of 10% oxygen and 90% hydrogen for the SWPC configuration, and the *Young et al.* [1982b] model for the high-resolution with RCM configuration.

A.0.3 Ionospheric electrodynamics

The Ridley Ionosphere Model (RIM) models calculates ionospheric parameters on a height-integrated basis. This model is described in *Ridley and Liemohn* [2002] and *Ridley et al.* [2004a]. It receives field-aligned current values from the MHD solver, and from these calculates conductance and electric potential. The potential values are then passed back to the inner magnetosphere and MHD models, where they are used to determine the velocity tangent to the inner boundary (the velocity normal to the boundary is set to zero) [*Welling and Liemohn*, 2014]. As discussed in *Welling and Liemohn* [2016], the ionospheric boundary is of crucial importance to the overall dynamics of the magnetospheric dynamics. While more sophisticated models exist to model the interaction through this boundary, most are either too computationally costly [such as the Polar Wind Outflow Model *Glocer et al.*, 2007], or lack a fully tested coupling to an MHD model.

Acknowledgments

Thanks to Ruth Skoug of Los Alamos National Laboratory for providing solar wind data from the Advanced Composition Explorer (ACE) satellite to cover gaps in the publicly available Level 2 datasets.

Thanks to World Data Center Kyoto for providing observed values for the Sym-H index, the NOAA NGDC for providing observed values for the Kp index, and the NASA Goddard Spaceflight Center for their OMNI and CDAWeb tools that provide access to a variety of datasets.

For the ground magnetometer data we gratefully acknowledge: Intermagnet; USGS, Jeffrey J. Love; CARISMA, PI Ian Mann; CANMOS; The S-RAMP Database, PI K. Yumoto and Dr. K. Shiokawa; The SPIDR database; AARI, PI Oleg Troshichev; The MACCS program, PI M. Engebretson, Geomagnetism Unit of the Geological Survey of Canada; GIMA; MEASURE, UCLA IGPP and Florida Institute of Technology; SAMBA, PI Eftyhia Zesta; 210 Chain, PI K. Yumoto; SAMNET, PI Farideh Honary; The institutes who maintain the IMAGE magnetometer array, PI Liisa Juusola; PENGUIN; AUTUMN, PI Martin Connors; DTU Space, PI Dr. Rico Behlke; South Pole and McMurdo Magnetometer, PI's Louis J. Lanzarotti and Alan T. Weatherwax; ICESTAR; RAPIDMAG; PENGUIn; British Artarctic Survey; McMac, PI Dr. Peter Chi; BGS, PI Dr. Susan Macmillan; Pushkov Institute of Terrestrial Magnetism, Ionosphere and Radio Wave Propagation (IZMIRAN); GFZ, PI Dr. Juergen Matzka; MFGI, PI B. Heilig; IGFPAS, PI J. Reda; University of L'Aquila, PI M. Vellante; SuperMAG, PI Jesper W. Gjerloev.

This research was conducted as part of the Space Hazards Induced near Earth by Large, Dynamic Storms (SHIELDS) project, funded by the U.S. Department of Energy through the LANL/LDRD-DR Program under contract DE-AC52-06NA25396.

S. Morley was supported by the U.S. Department of Energy Laboratory Directed Research and Development (LDRD) project award 20170047DR.

N. Ganushkina thanks the International Space Science Institute (ISSI) in Bern, Switzerland for the support of the international teams on “Analysis of Cluster Inner Magnetosphere Campaign Data, in Application to the Dynamics of Waves and Wave-Particle Interaction within the Outer Radiation Belt,” and “Ring Current Modeling: Uncommon Assumptions and Common Misconceptions.” The research of N. Ganushkina leading to these results was partly funded by the European Union’s Horizon 2020 research and innovation program under grant agreement No. 637302 PROGRESS. Support for N. Ganushkina at Michigan was provided by research grants from NASA (NNX14AF34G, NNX17AI48G).

Full output from the three model runs presented in this paper can be found at the following URLs:

http://vmr.engin.umich.edu/Model/_swmf_mag/plot?run=Jan2005_SWPC

http://vmr.engin.umich.edu/Model/_swmf_mag/plot?run=Jan2005_Hi-res_w_RCM

http://vmr.engin.umich.edu/Model/_swmf_mag/plot?run=Jan2005_Hi-res_wo_RCM

References

Akasofu, S.-I., and S. Yoshida (1966), Growth and decay of the ring current and the polar electrojets, *Journal of Geophysical Research*, *71*(1), 231–240, doi: 10.1029/JZ071i001p00231.

Anderson, B. J., H. Korth, D. T. Welling, V. G. Merkin, M. J. Wiltberger, J. Raeder, R. J. Barnes, C. L. Waters, A. A. Pulkkinen, and L. Rastätter (2017), Comparison of predictive estimates of high-latitude electrodynamics with observations of global-scale birkeland currents, *Space Weather*, *15*(2), 352–373, doi:10.1002/2016SW001529, 2016SW001529.

Bartels, J., N. H. Heck, and H. F. Johnston (1939), The three-hour-range index measuring geomagnetic activity, *Terrestrial Magnetism and Atmospheric Electricity*, *44*(4), 411–454, doi:10.1029/TE044i004p00411.

Borovsky, J. E. (2012), The effect of sudden wind shear on the Earth’s magnetosphere: Statistics of wind shear events and CCMC simulations of magnetotail disconnections, *Journal of Geophysical Research: Space Physics*, *117*(A6), doi:10.1029/2012JA017623.

Bristow, W. A., R. A. Greenwald, S. G. Shepherd, and J. M. Hughes (2004), On the observed variability of the cross-polar cap potential, *Journal of Geophysical Research: Space Physics*, *109*(A2), doi:10.1029/2003JA010206.

Cash, M. D., S. Witters Hicks, D. A. Biesecker, A. A. Reinard, C. A. de Koning, and D. R. Weimer (2016), Validation of an operational product to determine l1 to earth propagation time delays, *Space Weather*, *14*(2), 93–112, doi:10.1002/2015SW001321, 2015SW001321.

Cramer, W. D., J. Raeder, F. R. Toffoletto, M. Gilson, and B. Hu (2017), Plasma sheet injections into the inner magnetosphere: Two-way coupled openggcm-rcm model results, *Journal of Geophysical Research: Space Physics*, doi:10.1002/2017JA024104, 2017JA024104.

Crooker, N. U., J. G. Lyon, and J. A. Fedder (1998), MHD model merging with IMF By: Lobe cells, sunward polar cap convection, and overdraped lobes, *Journal of Geophysical Research: Space Physics*, *103*(A5), 9143–9151, doi:10.1029/97JA03393.

Davis, T. N., and M. Sugiura (1966), Auroral electrojet activity index ae and its universal time variations, *Journal of Geophysical Research*, *71*(3), 785–801, doi:10.1029/JZ071i003p00785.

De Zeeuw, D. L., T. I. Gombosi, C. P. T. Groth, K. G. Powell, and Q. F. Stout (2000), An adaptive MHD method for global space weather simulations, *IEEE Trans. Plasma Sci.*, *28*, 1956.

Dubyagin, S., N. Ganushkina, M. Kubyshkina, and M. Liemohn (2014), Contribution from different current systems to SYM and ASY midlatitude indices, *Journal of Geophysical Research: Space Physics*, *119*(9), 7243–7263, doi:10.1002/2014JA020122.

Facskó, G., I. Honkonen, T. Živković, L. Palin, E. Kallio, K. Ågren, H. Opgenoorth, E. I. Tanskanen, and S. Milan (2016), One year in the Earth's magnetosphere: A global MHD simulation and spacecraft measurements, *Space Weather*, *14*(5), 351–367, doi:10.1002/2015SW001355.

Gannon, J., and J. Love (2011), USGS 1-min Dst index, *Journal of Atmospheric and Solar-Terrestrial Physics*, *73*(2-3), 323–334, doi:10.1016/j.jastp.2010.02.013.

Ganushkina, N. Y., T. I. Pulkkinen, M. V. Kubyshkina, H. J. Singer, and C. T. Russell (2004), Long-term evolution of magnetospheric current systems during storms, *Annales Geophysicae European Geosciences Union*, *22*(4), 1317–1334.

Ganushkina, N. Y., M. W. Liemohn, M. V. Kubyshkina, R. Ilie, and H. J. Singer (2010), Distortions of the magnetic field by storm-time current systems in earth's magnetosphere, *Annales Geophysicae*, *28*(1), 123–140, doi:10.5194/angeo-28-123-2010.

Gjerloev, J. W. (2012), The SuperMAG data processing technique, *Journal of Geophysical Research: Space Physics*, *117*(A9), doi:10.1029/2012JA017683, a09213.

Glocer, A., T. I. Gombosi, G. Tóth, K. C. Hansen, A. J. Ridley, and A. Nagy (2007), Polar wind outflow model: Saturn results, *J. Geophys. Res.*, *112*, doi:10.1029/2006JA011755.

Glocer, A., M. Fok, X. Meng, G. Tóth, N. Buzulukova, S. Chen, and K. Lin (2012), CRCM + BATS-R-US two way coupling, *J. Geophys. Res.*, *118*, doi:10.1002/jgra.50221.

Glocer, A., L. Rastätter, M. Kuznetsova, A. Pulkkinen, H. J. Singer, C. Balch, D. Weimer, D. Welling, M. Wiltberger, J. Raeder, R. S. Weigel, J. McCollough, and S. Wing (2016), Community-wide validation of geospace model local k-index predictions to support model transition to operations, *Space Weather*, *14*(7), 469–480, doi:10.1002/2016SW001387, 2016SW001387.

Groth, C., D. L. De Zeeuw, T. Gombosi, and K. Powell (2000), Global 3D MHD simulation of a space weather event: CME formation, interplanetary propagation, and interaction with the magnetosphere, *J. Geophys. Res.*, *105*, 25,053.

Guild, T. B., H. E. Spence, E. L. Kepko, V. Merkin, J. G. Lyon, M. Wiltberger, and C. C. Goodrich (2008), Geotail and LFM comparisons of plasma sheet climatology: 1. Average values, *Journal of Geophysical Research: Space Physics*, *113*(A4), doi:10.1029/2007JA012611.

Heinemann, M., and R. A. Wolf (2001), Relationships of models of the inner magnetosphere to the rice convection model, *J. Geophys. Res.*, *106*(A8), 15,545 – 15,554.

Hirsch, C. (2007), *Numerical Computation of Internal and External Flows: The Fundamentals of Computational Fluid Dynamics*, Elsevier Science.

Huang, C.-L., H. E. Spence, H. J. Singer, and W. J. Hughes (2010), Modeling radiation belt radial diffusion in ULF wave fields: 1. Quantifying ULF wave power at geosynchronous orbit in observations and in global MHD model, *Journal of Geophysical Research: Space Physics*, *115*(A6), doi:10.1029/2009JA014917.

Iyemori, T. (1990), Storm-time magnetospheric currents inferred from mid-latitude geomagnetic field variations., *Journal of geomagnetism and geoelectricity*, *42*(11), 1249–1265, doi:10.5636/jgg.42.1249.

Janhunen, P., M. Palmroth, T. Laitinen, I. Honkonen, L. Juusola, G. Facskó, and T. I. Pulkkinen (2012), The GUMICS-4 global {MHD} magnetosphere-ionosphere coupling simulation, *Journal of Atmospheric and Solar-Terrestrial Physics*, *80*(0), 48–59, doi:http://dx.doi.org/10.1016/j.jastp.2012.03.006.

Jones, E., T. Oliphant, P. Peterson, et al. (2001, updated frequently), SciPy: Open source scientific tools for Python, [Online; accessed 2017-03-06].

Juusola, L., G. Facskó, I. Honkonen, P. Janhunen, H. Vanhamäki, K. Kauristie, T. V. Laitinen, S. E. Milan, M. Palmroth, E. I. Tanskanen, and A. Viljanen (2014), Statistical comparison of seasonal variations in the GUMICS-4 global MHD model ionosphere and measurements, *Space Weather*, *12*(10), 582–600, doi:10.1002/2014SW001082.

Kalegaev, V. V., N. Y. Ganushkina, T. I. Pulkkinen, M. V. Kubyshkina, H. J. Singer, and C. T. Russell (2005), Relation between the ring current and the tail current during magnetic storms, *Annales Geophysicae*, *23*(2), 523–533.

Katus, R. M., and M. W. Liemohn (2013), Similarities and differences in low- to middle-latitude geomagnetic indices, *Journal of Geophysical Research: Space Physics*, *118*(8), 5149–5156, doi:10.1002/jgra.50501.

Koren, B. (1993), A robust upwind discretisation method for advection, diffusion and source terms, in *Numerical Methods for Advection-Diffusion Problems*, edited by C. Vreugdenhil and B.Koren, p. 117, Vieweg, Braunschweig.

Kress, B. T., M. K. Hudson, M. D. Looper, J. Albert, J. G. Lyon, and C. C. Goodrich (2007), Global mhd test particle simulations of >10 mev radiation belt electrons during storm sudden commencement, *Journal of Geophysical Research: Space Physics*, *112*(A9), doi:10.1029/2006JA012218, a09215.

Kronberg, E. A., M. A.-a. Iannis, D. C. Delcourt, E. E. Grigorenko, L. M. Kistler, I. V. Kuzichev, J. Liao, R. Maggiolo, H. V. Malova, K. G. Orlova, V. Peroomian, D. R. Shklyar, Y. Y. Shprits, D. T. Welling, and L. M. Zelenyi (2014), Circulation of Heavy Ions and Their Dynamical Effects in the Magnetosphere : Recent Observations and Models Charge Energy Mass experiment Extreme Ultraviolet radiation, *Space Science Reviews*, *184*, 173–235, doi:10.1007/s11214-014-0104-0.

Liemohn, M. W., J. U. Kozyra, M. F. Thomsen, J. L. Roeder, G. Lu, J. E. Borovsky, and T. E. Cayton (2001), Dominant role of the asymmetric ring current in producing the stormtime Dst*, *Journal of Geophysical Research: Space Physics*, *106*(A6), 10,883–10,904, doi:10.1029/2000JA000326.

Liemohn, M. W., D. L. De Zeeuw, N. Y. Ganushkina, J. U. Kozyra, and D. T. Welling (2013), Magnetospheric cross-field currents during the January 6-7, 2011 high-speed stream-driven interval, *Journal of Atmospheric and Solar-Terrestrial Physics*, *99*, 78–84, doi:10.1016/j.jastp.2012.09.007.

Lockwood, M., and S. K. Morley (2004), A numerical model of the ionospheric signatures of time-varying magneticreconnection: I. ionospheric convection, *Annales Geophysicae*, *22*(1), 73–91.

Lopez, R., J. Lyon, M. Wiltberger, and C. Goodrich (2001), Comparison of global mhd simulation results with actual storm and substorm events, *Advances in Space Research*, *28*(12), 1701 – 1706, doi:http://dx.doi.org/10.1016/S0273-1177(01)00535-X.

Lyon, J., J. Fedder, and C. Mobarry (2004), The Lyon-Fedder-Mobarry (LFM) global MHD magnetospheric simulation code, *J. Atmos. Sol-Terr. Phys.*, *66*, 1333.

Maltsev, Y. (2004), Points of controversy in the study of magnetic storms, *Space Science Reviews*, *110*(3/4), 227–277, doi:10.1023/B:SPAC.0000023410.77752.30.

Mayaud, P. N. (1980), *Derivation, Meaning, and Use of Geomagnetic Indices*, American Geophysical Union, Washington, DC, doi:10.1002/9781118663837.

McComas, D. J., S. J. Bame, P. Barker, W. C. Feldman, J. L. Phillips, P. Riley, and J. W. Griffee (1998), *Solar Wind Electron Proton Alpha Monitor (SWEPAM) for the Advanced Composition Explorer*, pp. 563–612, Springer Netherlands, Dordrecht, doi:10.1007/978-94-011-4762-0_20.

Milan, S. E. (2004), Dayside and nightside contributions to the cross polar cap potential: placing an upper limit on a viscous-like interaction, *Annales Geophysicae*, *22*(10), 3771–3777, doi:10.5194/angeo-22-3771-2004.

Moen, J., and A. Brekke (1993), The solar flux influence of quiet-time conductances in the auroral ionosphere, *Geophys. Res. Lett.*, *20*, 971.

Morley, S., J. Koller, D. Welling, B. Larsen, and J. Niehof (2014), SpacePy: Python-Based Tools for the Space Science Community, Astrophysics Source Code Library.

Morley, S. K. (2007), 7th Australian Space Science Conference Proceedings, pp. 118–129, National Space Society of Australia Ltd.

Morley, S. K., A. P. Rouillard, and M. P. Freeman (2009), Recurrent substorm activity during the passage of a corotating interaction region, *Journal of Atmospheric and Solar-Terrestrial Physics*, *71*(10), 1073–1081, doi:10.1016/j.jastp.2008.11.009.

Morley, S. K., D. T. Welling, J. Koller, B. A. Larsen, M. G. Henderson, and J. Niehof (2011), SpacePy - A Python-based Library of Tools for the Space Sciences, *Proceedings of the 9th Python in Science Conference*, pp. 39–45.

Ngwira, C. M., A. Pulkkinen, M. Leila Mays, M. M. Kuznetsova, A. B. Galvin, K. Simunac, D. N. Baker, X. Li, Y. Zheng, and A. Glocer (2013), Simulation of the 23 july 2012 extreme space weather event: What if this extremely rare cme was earth directed?, *Space Weather*, *11*(12), 671–679, doi:10.1002/2013SW000990.

Ngwira, C. M., A. Pulkkinen, M. M. Kuznetsova, and A. Glocer (2014), Modeling extreme “carrington-type” space weather events using three-dimensional global mhd simulations, *Journal of Geophysical Research: Space Physics*, *119*(6), 4456–4474, doi:10.1002/2013JA019661.

Ogino, T., R. J. Walker, and M. Ashour-Abdalla (1992), A global magnetohydrodynamic simulation of the magnetosheath and magnetosphere when the interplanetary magnetic field is northward, *IEEE Transactions on Plasma Science*, *20*(6), 817–828, doi:10.1109/27.199534.

Ohtani, S., M. Nosé, G. Rostoker, H. Singer, and A. Lui (2001), Storm-substorm relationship: Contribution of the tail current, *J. Geophys.*

Palmroth, M., T. I. Pulkkinen, P. Janhunen, and C.-C. Wu (2003), Stormtime energy transfer in global mhd simulation, *Journal of Geophysical Research: Space Physics*, *108*(A1), doi:10.1029/2002JA009446, 1048.

Parzen, E. (1962), On estimation of a probability density function and mode, *Ann. Math. Statist.*, *33*(3), 1065–1076, doi:10.1214/aoms/1177704472.

Pembroke, A., F. Toffoletto, S. Sazykin, M. Wiltberger, J. Lyon, V. Merkin, and P. Schmitt (2012), Initial results from a dynamic coupled magnetosphere-ionosphere-ring current model, *Journal of Geophysical Research: Space Physics*, *117*(A2), doi:10.1029/2011JA016979, a02211.

Powell, K., P. Roe, T. Linde, T. Gombosi, and D. L. De Zeeuw (1999), A solution-adaptive upwind scheme for ideal magnetohydrodynamics, *J. Comp. Phys.*, *154*, 284–309.

Pulkkinen, A., L. Rastätter, M. Kuznetsova, M. Hesse, A. Ridley, J. Raeder, H. J. Singer, and A. Chulaki (2010), Systematic evaluation of ground and geostationary magnetic field predictions generated by global magnetohydrodynamic models, *Journal of Geophysical Research: Space Physics*, *115*(A3), doi:10.1029/2009JA014537, a03206.

Pulkkinen, A., L. Rastätter, M. Kuznetsova, H. Singer, C. Balch, D. Weimer, G. Tóth, A. Ridley, T. Gombosi, M. Wiltberger, J. Raeder, and R. Weigel (2013), Community-wide validation of geospace model ground magnetic field perturbation predictions to support model transition to operations, *Space Weather*, *11*(6), 369–385, doi: 10.1002/swe.20056.

Raeder, J., J. Berchem, and M. Ashour-Abdalla (1998), The geospace environment modeling grand challenge: Results from a global geospace circulation model, *Journal of Geophysical Research: Space Physics*, *103*(A7), 14,787–14,797, doi:10.1029/98JA00014.

Raeder, J., R. McPherron, L. Frank, S. Kokubun, G. Lu, T. Mukai, W. Paterson, J. Sigwarth, H. Singer, and J. Slavin (2001), Global simulation of the Geospace Environment Modeling substorm challenge event, *J. Geophys. Res.*, *106*, 281.

Rastätter, L., M. M. Kuznetsova, A. Vapirev, A. Ridley, M. Wiltberger, A. Pulkkinen, M. Hesse, and H. J. Singer (2011), Geospace environment modeling 2008-2009 challenge: Geosynchronous magnetic field, *Space Weather*, *9*(4), doi: 10.1029/2010SW000617, s04005.

Rastätter, L., M. M. Kuznetsova, A. Glocer, D. Welling, X. Meng, J. Raeder, M. Wiltberger, V. K. Jordanova, Y. Yu, S. Zaharia, R. S. Weigel, S. Sazykin, R. Boynton, H. Wei, V. Eccles, W. Horton, M. L. Mays, and J. Gannon (2013), Geospace environment modeling 2008-2009 challenge: Dst index, *Space Weather*, *11*(4), 187–205, doi: 10.1002/swe.20036.

Richmond, A. D. (1992), Assimilative mapping of ionospheric electrodynamics, *Adv. Space Res.*, *12*, 59.

Richmond, A. D., and Y. Kamide (1988), Mapping electrodynamic features of the high-latitude ionosphere from localized observations: Technique, *Journal of Geophysical Research*, *93*(A6), 5741, doi:10.1029/JA093iA06p05741.

Ridley, A., T. Gombosi, and D. Dezeeuw (2004a), Ionospheric control of the magnetosphere: conductance, *Annales Geophysicae*, *22*, 567–584, doi:10.5194/angeo-22-567-2004.

Ridley, A., T. Gombosi, and D. Dezeeuw (2004b), Ionospheric control of the magnetosphere: conductance, *Annales Geophysicae*, *22*, 567–584.

Ridley, A. J., and M. W. Liemohn (2002), A model-derived storm time asymmetric ring current driven electric field description, *Journal of Geophysical Research: Space Physics*, *107*(A8), SMP 2–1–SMP 2–12, doi:10.1029/2001JA000051.

Ridley, A. J., T. I. Gombosi, D. L. De Zeeuw, C. R. Clauer, and A. D. Richmond (2003), Ionospheric control of the magnetosphere: Thermospheric neutral winds, *J. Geophys. Res.*, *108*(A8), 1328, doi:10.1029/2002JA009464.

Rostoker, G. (1972), Geomagnetic indices, *Reviews of Geophysics*, *10*(4), 935, doi: 10.1029/RG010i004p00935.

Sazykin, S. Y. (2000), Theoretical studies of penetration of magnetospheric electric fields to the ionosphere, Ph.D. thesis, Utah State University, Logan, Utah.

Scott, D. W. (2015), Multivariate density estimation: theory, practice, and visualization.

Sokolov, I., E. V. Timofeev, J. Sakai, and K. Takayama (2002), Artificial wind – a new framework to construct simple and efficient upwind shock-capturing schemes, *J. Comput. Phys.*, *181*, 354–393, doi:10.1006/jcph.2002.7130.

Tapping, K. F. (2013), The 10.7 cm solar radio flux (F 10.7), *Space Weather*, *11*(7), 394–406, doi:10.1002/swe.20064.

Taylor, J. (1997), *An introduction to error analysis : the study of uncertainties in physical measurements*, University Science Books, Sausalito, Calif.

Thomsen, M. F. (2004), Why kp is such a good measure of magnetospheric convection, *Space Weather*, *2*(11), doi:10.1029/2004SW000089, s11004.

Toffoletto, F., S. Sazykin, R. Spiro, and R. Wolf (2003), Inner magnetospheric modeling with the Rice Convection Model, *Space Sci. Rev.*, *107*, 175–196.

Tóth, G., I. V. Sokolov, T. I. Gombosi, D. R. Chesney, C. Clauer, D. L. D. Zeeuw, K. C. Hansen, K. J. Kane, W. B. Manchester, K. G. Powell, A. J. Ridley, I. I. Roussev, Q. F. Stout, O. Volberg, R. A. Wolf, S. Sazykin, A. Chan, B. Yu, and J. Kóta (2005), Space Weather Modeling Framework: A new tool for the space science community, *J. Geophys. Res.*, *110*, A12,226, doi:10.1029/2005JA011126.

Tóth, G., B. van der Holst, I. V. Sokolov, D. L. De Zeeuw, T. I. Gombosi, F. Fang, W. B. Manchester, X. Meng, D. Najib, K. G. Powell, Q. F. Stout, A. Glocer, Y.-J. Ma, and M. Opher (2012), Adaptive numerical algorithms in space weather modeling, *Journal of Computational Physics*, *231*(3), 870–903, doi:10.1016/j.jcp.2011.02.006.

Vasyliunas, V. M. (1970), Mathematical Models of Magnetospheric Convection and its Coupling to the Ionosphere, pp. 60–71, Springer Netherlands, doi:10.1007/978-94-010-3284-1_6.

Wanliss, J. A., and K. M. Showalter (2006), High-resolution global storm index: Dst versus sym-h, *Journal of Geophysical Research: Space Physics*, *111*(A2), doi:

10.1029/2005JA011034, a02202.

Weimer, D. R. (2004), Correction to "Predicting interplanetary magnetic field (IMF) propagation delay times using the minimum variance technique", *Journal of Geophysical Research*, *109*(A12), A12,104, doi:10.1029/2004JA010691.

Weimer, D. R. (2005), Improved ionospheric electrodynamic models and application to calculating joule heating rates, *Journal of Geophysical Research: Space Physics*, *110*(A5), doi:10.1029/2004JA010884, a05306.

Weimer, D. R., and J. H. King (2008), Improved calculations of interplanetary magnetic field phase front angles and propagation time delays, *Journal of Geophysical Research: Space Physics*, *113*(A1), doi:10.1029/2007JA012452.

Weimer, D. R., D. M. Ober, N. C. Maynard, M. R. Collier, D. J. McComas, N. F. Ness, C. W. Smith, and J. Watermann (2003), Predicting interplanetary magnetic field (imf) propagation delay times using the minimum variance technique, *Journal of Geophysical Research: Space Physics*, *108*(A1), doi:10.1029/2002JA009405, 1026.

Welling, D. T., and M. W. Liemohn (2014), Outflow in global magnetohydrodynamics as a function of a passive inner boundary source, *Journal of Geophysical Research: Space Physics*, *119*(4), 2691–2705, doi:10.1002/2013JA019374.

Welling, D. T., and M. W. Liemohn (2016), The ionospheric source of magnetospheric plasma is not a black box input for global models, *Journal of Geophysical Research: Space Physics*, *121*(6), 5559–5565, doi:10.1002/2016JA022646, 2016JA022646.

Welling, D. T., and A. J. Ridley (2010), Validation of SWMF magnetic field and plasma, *Space Weather*, *8*, doi:10.1029/2009SW000494.

Welling, D. T., and S. G. Zaharia (2012), Ionospheric outflow and cross polar cap potential: What is the role of magnetospheric inflation?, *Geophysical Research Letters*, *39*(23), doi:10.1029/2012GL054228.

Welling, D. T., V. K. Jordanova, A. Glocer, G. Toth, M. W. Liemohn, and D. R. Weimer (2015), The two-way relationship between ionospheric outflow and the ring current, *Journal of Geophysical Research: Space Physics*, *120*(6), 4338–4353, doi: 10.1002/2015JA021231.

Welling, D. T., B. J. Anderson, G. Crowley, A. A. Pulkkinen, and L. Rastätter (2017), Exploring predictive performance: A reanalysis of the geospace model transition challenge, *Space Weather*, *15*(1), 192–203, doi:10.1002/2016SW001505, 2016SW001505.

Wiltberger, M., E. Rigler, V. Merkin, J. Lyon, B. M. Wiltberger, V. Merkin SlavaMerkin, and j. J. Lyon JohnLyon (2017), Structure of High Latitude Currents in Magnetosphere-Ionosphere Models, *Space Sci Rev*, *206*, 575–598, doi:10.1007/s11214-016-0271-2.

Winglee, R. (2000), Mapping of ionospheric outflows into the magnetosphere for varying IMF conditions., *J. Atmos. and Terr. Phys.*, *62*, 527.

Winglee, R. M., D. Chua, M. Brittnacher, G. K. Parks, and G. Lu (2002), Global impact of ionospheric outflows on the dynamics of the magnetosphere and cross-polar cap potential, *Journal of Geophysical Research: Space Physics*, *107*(A9), SMP 11–1–SMP 11–12, doi:10.1029/2001JA000214, 1237.

Wolf, R. A., M. Harel, R. W. Spiro, G. Voigt, P. H. Reiff, and C. K. Chen (1982), Computer simulation of inner magnetospheric dynamics for the magnetic storm of July 29, 1977, *J. Geophys. Res.*, *87*, 5949–5962.

Wu, C. C., R. J. Walker, and J. M. Dawson (1981), A three dimensional mhd model of the earth's magnetosphere, *Geophysical Research Letters*, *8*(5), 523–526, doi: 10.1029/GL008i005p00523.

Young, D. T., H. Balsiger, and J. Geiss (1982a), Correlations of magnetospheric ion composition with geomagnetic and solar activity, *J. Geophys. Res.*, *87*, 9077.

Young, D. T., H. Balsiger, and J. Geiss (1982b), Correlations of magnetospheric ion composition with geomagnetic and solar activity, *Journal of Geophysical Research*, *87*(A11), 9077, doi:10.1029/JA087iA11p09077.

Yu, Y., and A. Ridley (2008), Validation of the Space Weather Modeling Framework using ground-based magnetometers, *Space Weather*, *6*, doi:10.1029/2007SW000345.

Yu, Y., A. J. Ridley, D. T. Welling, and G. Tóth (2010), Including gap region field-aligned currents and magnetospheric currents in the MHD calculation of ground-based magnetic field perturbations, *Journal of Geophysical Research (Space Physics)*, *115*(A14), A08,207, doi:10.1029/2009JA014869.

Yu, Y., V. Jordanova, D. Welling, B. Larsen, S. G. Claudepierre, and C. Kletzing (2014), The role of ring current particle injections: Global simulations and van allen probes observations during 17 march 2013 storm, *Geophysical Research Letters*, *41*(4), 1126–1132, doi:10.1002/2014GL059322.

Zhang, B., W. Lotko, M. Wiltberger, O. Brambles, and P. Damiano (2011), A statistical study of magnetosphere-ionosphere coupling in the Lyon-Fedder-Mobarry global MHD model, *Journal of Atmospheric and Solar-Terrestrial Physics*, *73*(5-6), 686–702, doi:

10.1016/j.jastp.2010.09.027.

Zhang, J., M. W. Liemohn, D. L. De Zeeuw, J. E. Borovsky, A. J. Ridley, G. Toth, S. Sazykin, M. F. Thomsen, J. U. Kozyra, T. I. Gombosi, and R. A. Wolf (2007), Understanding storm-time ring current development through data-model comparisons of a moderate storm, *Journal of Geophysical Research: Space Physics*, *112*(A4), doi: 10.1029/2006JA011846.

Zheng, Y., A. T. Y. Lui, M.-C. Fok, B. J. Anderson, P. C. Brandt, T. J. Immel, and D. G. Mitchell (2006), Relationship between Region 2 field-aligned current and the ring current: Model results, *Journal of Geophysical Research*, *111*(A11), A11S06, doi: 10.1029/2006JA011603.

Zheng, Y., A. T. Lui, M.-C. Fok, B. J. Anderson, P. C. Brandt, and D. G. Mitchell (2008), Controlling factors of Region 2 field-aligned current and its relationship to the ring current: Model results, *Advances in Space Research*, *41*(8), 1234–1242, doi: 10.1016/j.asr.2007.05.084.

Figure 1.

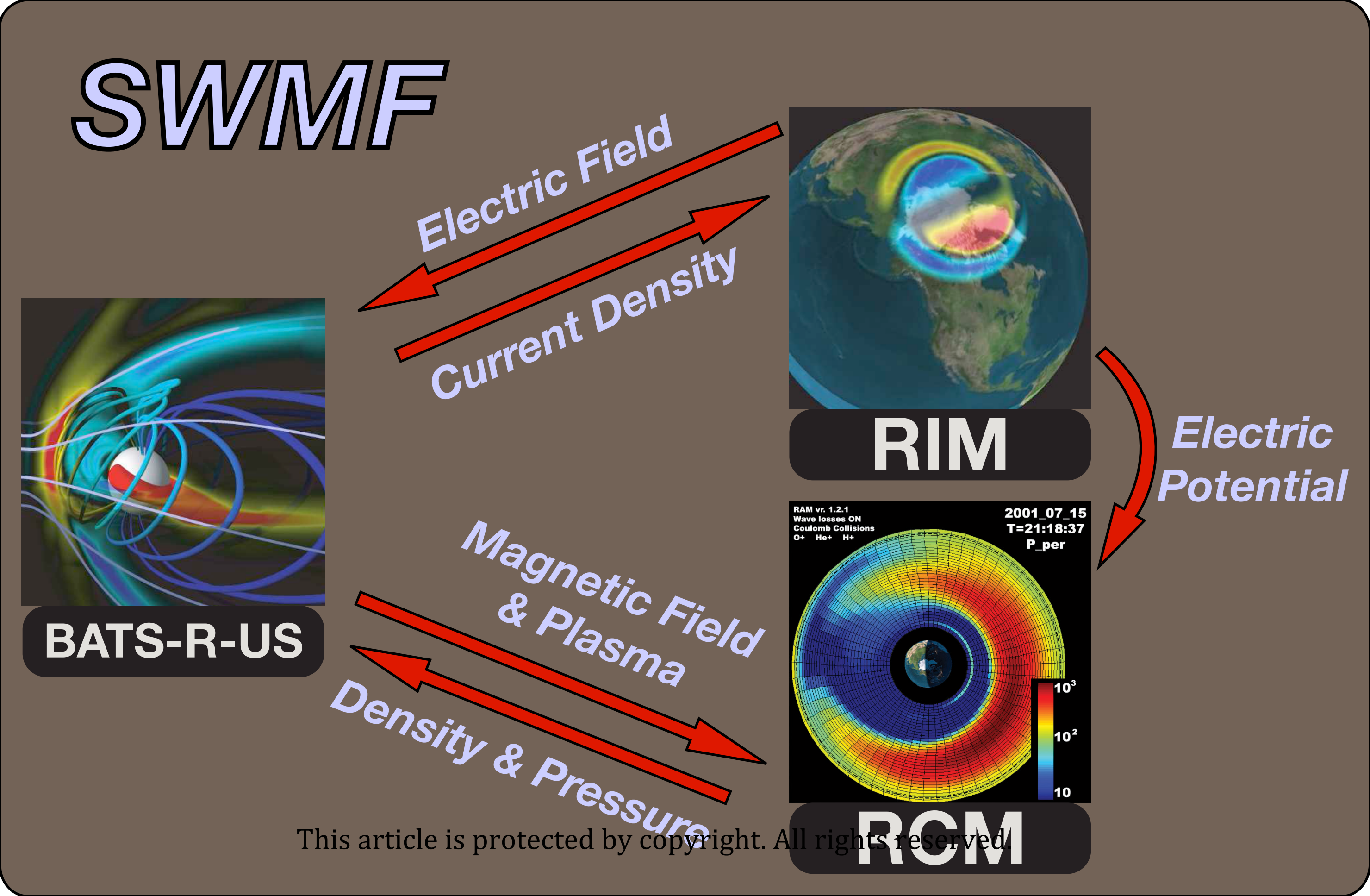

SWMF
Electric Field
Current Density
RIM
Electric Potential
BATS-R-US
Magnetic Field & Plasma
Density & Pressure
RAM vr. 1.2.1
Wave losses ON
Coulomb Collisions
O+ He+ H+
2001_07_15
T=21:18:37
P_per
10^3
10^2
10
RCM

Figure 2.

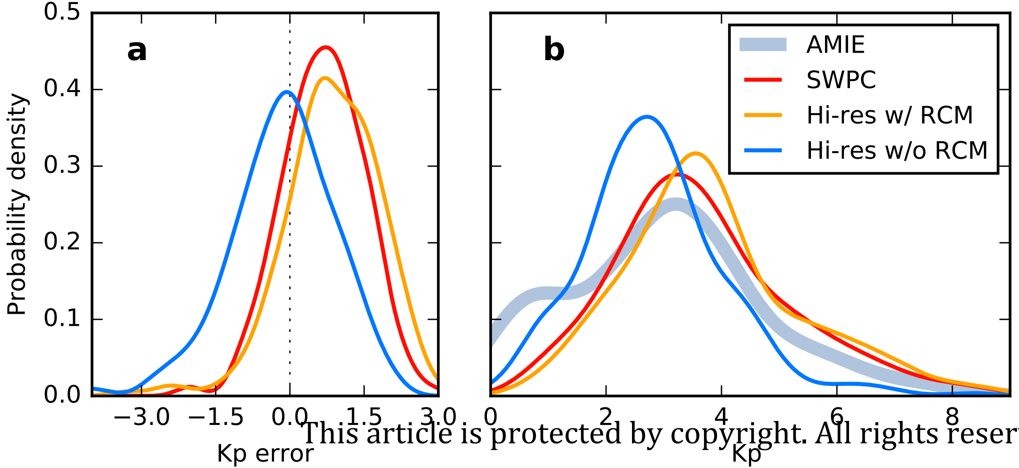
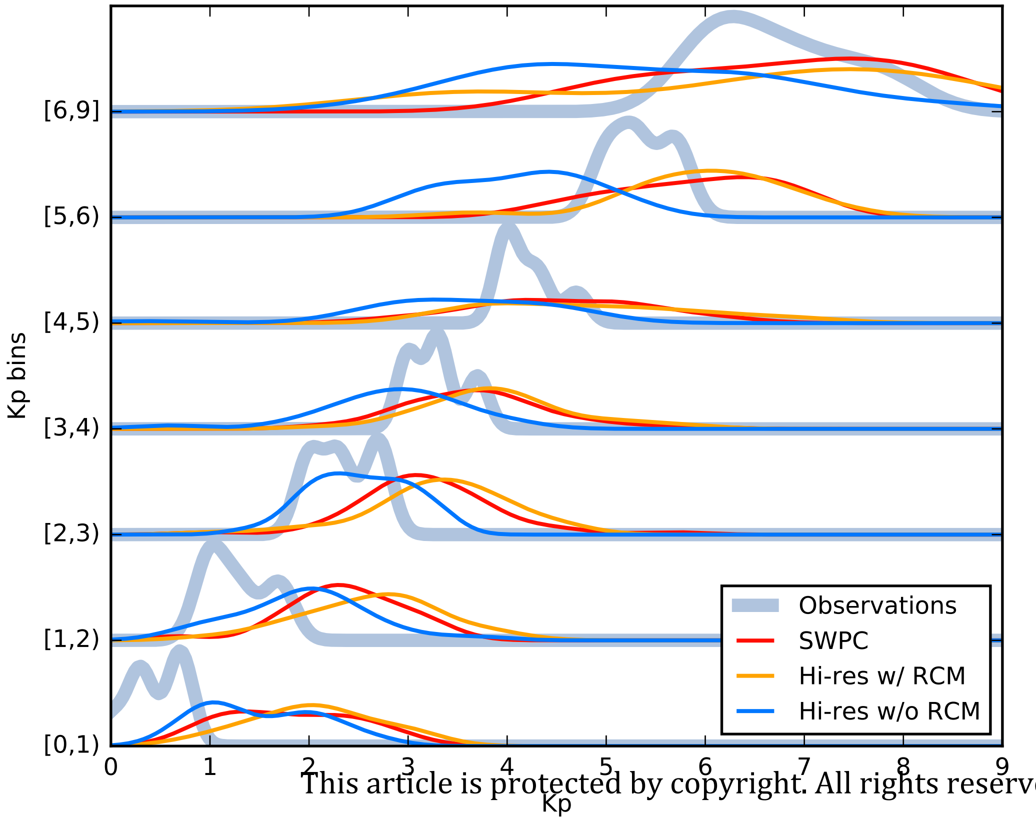

a
b
Probability density
0.0
0.1
0.2
0.3
0.4
0.5
−3.0
−1.5
0.0
1.5
3.0
Kp error
0
2
4
6
8
Kp
AMIE
SWPC
Hi-res w/ RCM
Hi-res w/o RCM

Figure 3.



[6,9]
[5,6)
[4,5)
[3,4)
[2,3)
[1,2)
[0,1)
Kp bins
0
1
2
3
4
5
6
7
8
9
Kp
Observations
SWPC
Hi-res w/ RCM
Hi-res w/o RCM

Figure 4.

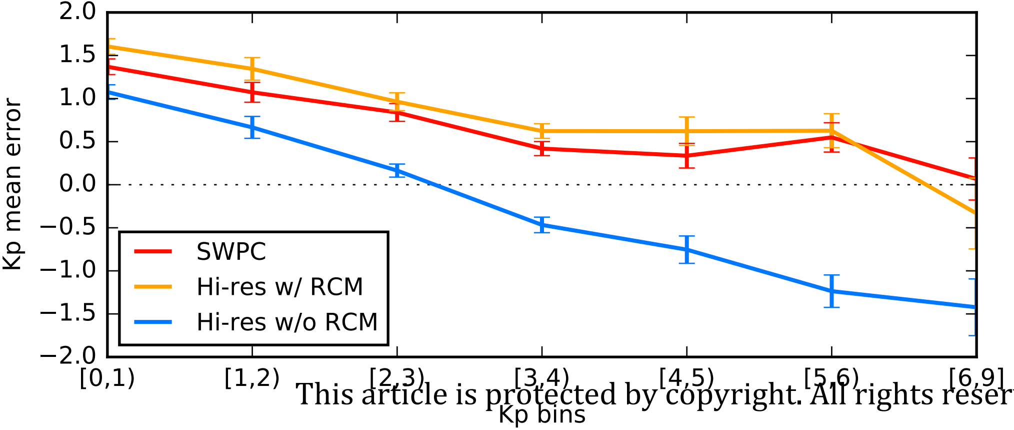
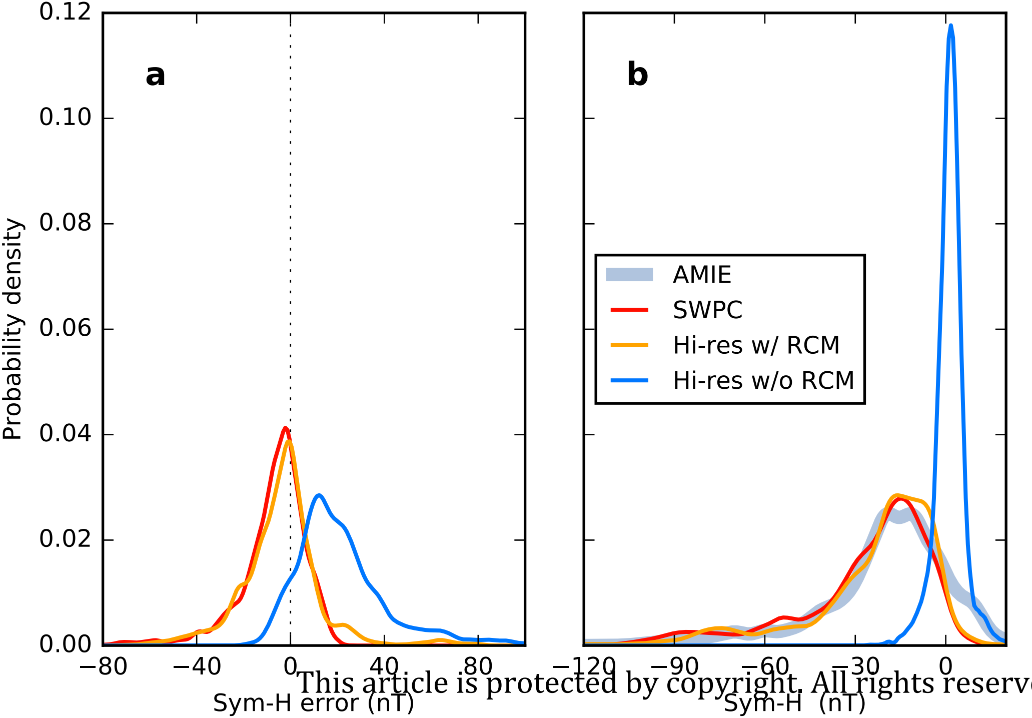

2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
Kp mean error
SWPC
Hi-res w/ RCM
Hi-res w/o RCM
[0,1)
[1,2)
[2,3)
[3,4)
[4,5)
[5,6)
[6,9]
Kp bins

Figure 5.



a
b
Probability density
0.00
0.02
0.04
0.06
0.08
0.10
0.12
−80
−40
0
40
80
−120
−90
−60
−30
0
Sym-H error (nT)
Sym-H (nT)
AMIE
SWPC
Hi-res w/ RCM
Hi-res w/o RCM

Figure 6.

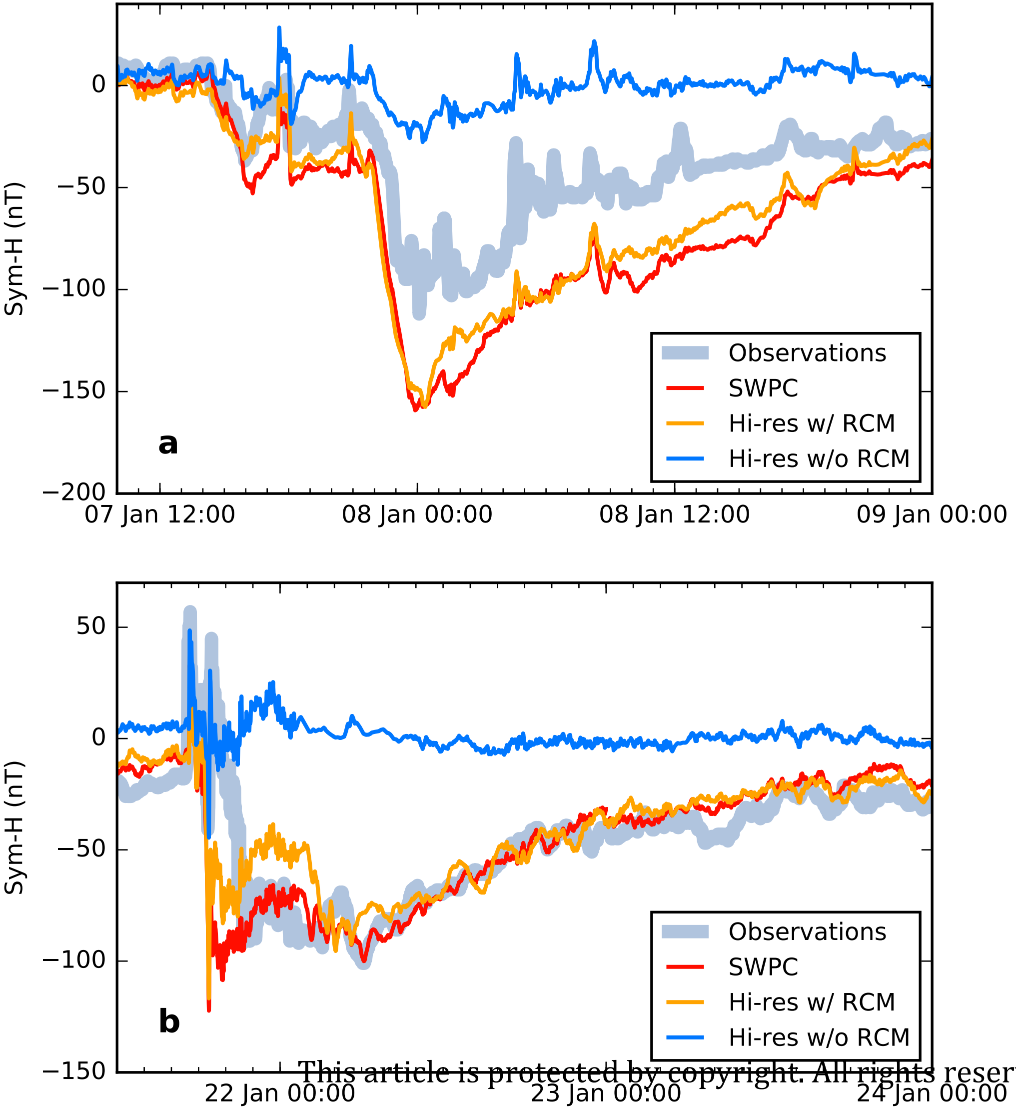

Sym-H (nT)
0
−50
−100
−150
−200
07 Jan 12:00
08 Jan 00:00
08 Jan 12:00
09 Jan 00:00
a
Observations
SWPC
Hi-res w/ RCM
Hi-res w/o RCM
Sym-H (nT)
50
0
−50
−100
−150
22 Jan 00:00
23 Jan 00:00
24 Jan 00:00
b
Observations
SWPC
Hi-res w/ RCM
Hi-res w/o RCM

Figure 7.

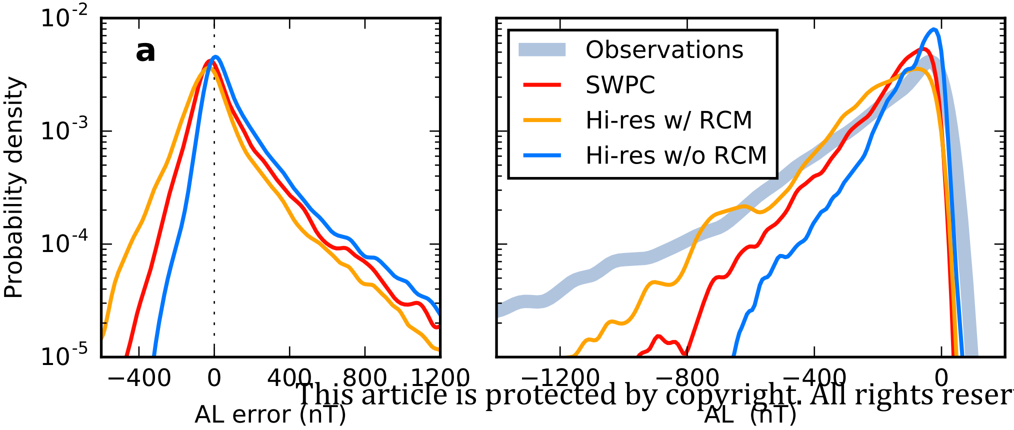

a
Probability density
10^{-2}
10^{-3}
10^{-4}
10^{-5}
−400
0
400
800
1200
AL error (nT)
Observations
SWPC
Hi-res w/ RCM
Hi-res w/o RCM
−1200
−800
−400
0
AL (nT)

Figure 8.

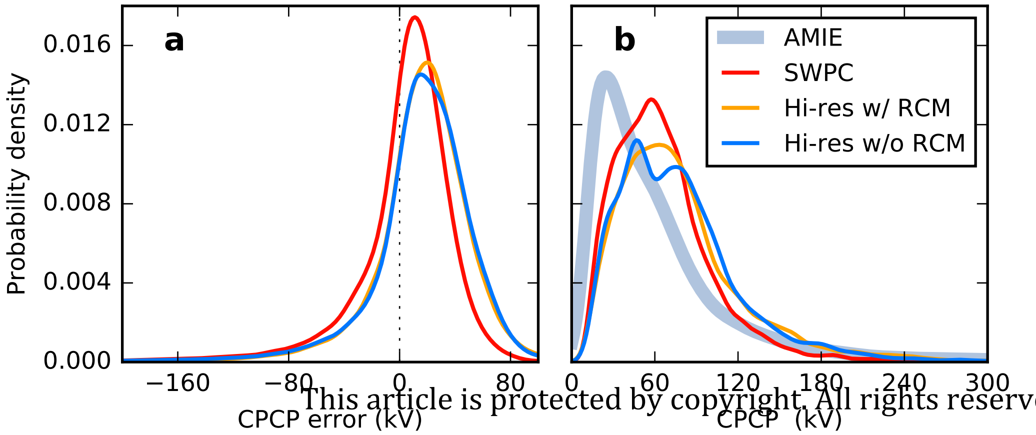
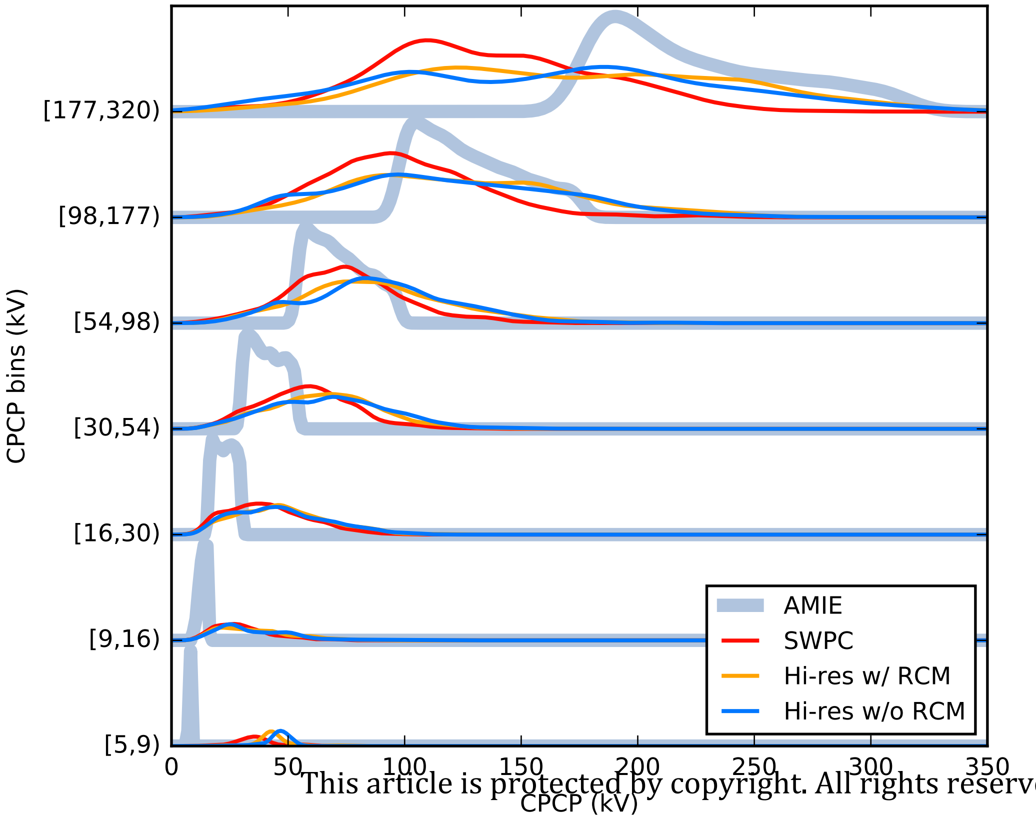
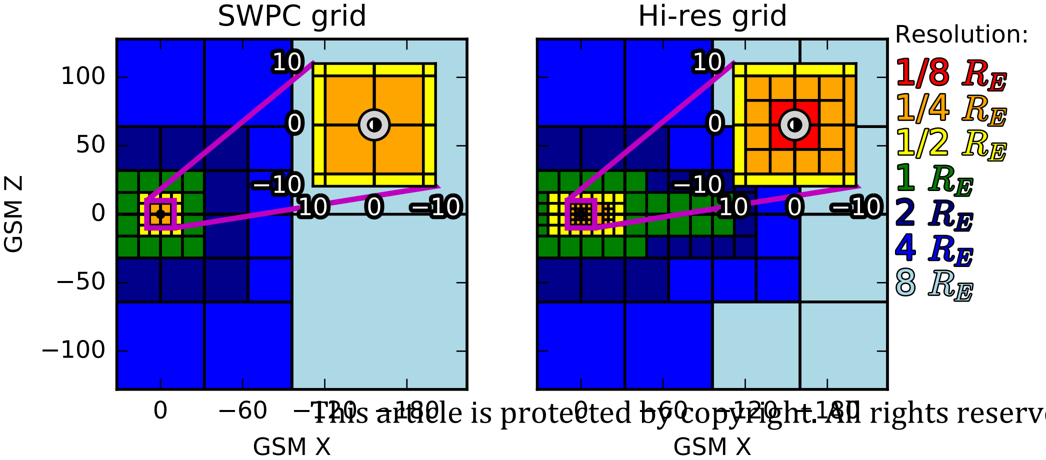

a
b
Probability density
0.000
0.004
0.008
0.012
0.016
−160
−80
0
80
CPCP error (kV)
0
60
120
180
240
300
CPCP (kV)
AMIE
SWPC
Hi-res w/ RCM
Hi-res w/o RCM

Figure 9.



[177,320)
[98,177)
[54,98)
[30,54)
[16,30)
[9,16)
[5,9)
CPCP bins (kV)
0
50
100
150
200
250
300
350
CPCP (kV)
AMIE
SWPC
Hi-res w/ RCM
Hi-res w/o RCM

Figure 10.



SWPC grid
Hi-res grid
Resolution:
1/8 R_E
1/4 R_E
1/2 R_E
1 R_E
2 R_E
4 R_E
8 R_E
GSM Z
GSM X
GSM X