JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 115, A03206, doi:10.1029/2009JA014537, 2010

Systematic evaluation of ground and geostationary magnetic field predictions generated by global magnetohydrodynamic models

A. Pulkkinen,[1,2] L. Rastätter,[2,3] M. Kuznetsova,[2] M. Hesse,[2] A. Ridley,[4] J. Raeder,[5] H. J. Singer,[6] and A. Chulaki[2]

Received 5 June 2009; revised 10 September 2009; accepted 16 October 2009; published 6 March 2010.

[1] In this work a systematic evaluation of ground and geostationary magnetic field predictions generated by a set of global magnetohydrodynamic (MHD) models is carried out. The evaluation uses four geospace storm events and ground magnetometer station and geostationary GOES data for comparisons between model output and observations. It is shown that metrics analysis of two different geospace parameters, i.e., geostationary and ground magnetic field, show surprising similarities, although the parameters reflect rather different properties of geospace. More specifically, increasing the spatial resolution and inclusion of more realistic inner magnetospheric physics successfully made the model predictions by the BATS-R-US model more accurate. Furthermore, while the OpenGGCM model had a tendency to have larger differences to observations than BATS-R-US in terms of the prediction efficiency, the model provided more accurate representation of the observed spectral characteristics of the ground and geostationary magnetic field fluctuations.

Citation: 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, *J. Geophys. Res.*, *115*, A03206, doi:10.1029/2009JA014537.

1. Introduction

[2] The ultimate goal in modeling any physical system is, given appropriate spatiotemporal boundary conditions, to accurately predict the course of the past and the future events within the system. As in meteorology, also in space physics the goal is not only challenging due to the vast complexity of the system but also of great practical interest as accurate forecasting of "weather" may enable actions that help to mitigate the potentially adverse societal impacts of the events. Although space weather forecasting is still in its infancy, significant advances made over the past decade, for example, in observational and computational space sciences have enabled realistic modeling of various parts of geospace. For example, coupled first-principles-based comprehensive geospace models [e.g., *Luhmann et al.*, 2004; *Tóth et al.*, 2005] have increased our understanding of the complex near-space phenomena and provided new unprecedented tools to tackle the space weather forecasting challenge.

[3] The advances made in observational and computational space sciences are illustrated by the rapidly increasing number of space physics model validation studies [e.g., *Raeder et al.*, 2001b; *Spence et al.*, 2004; *Huang et al.*, 2006; *Tóth et al.*, 2007; *Wang et al.*, 2008]. Also, there has been a shift from attempts to try to reproduce some very broad features of geospace such as large-scale magnetospheric plasma circulation (for one of the classical examples, see *Vasyliunas* [1970]) to attempts to try to reproduce more specific features such as behavior of individual storm time charged particle populations in the inner magnetosphere [e.g., *Fok et al.*, 2001; *Taktakishvili et al.*, 2007] and storm time ground magnetic field fluctuations [e.g., *Pulkkinen et al.*, 2007; *Yu and Ridley*, 2008]. The earlier model validation efforts have helped to better understand the limitations and the capabilities of various models and model setups and have been used to address the key model features in need improvement for better predictive capability.

[4] A majority of the earlier model validation work has been carried out by the model developers themselves and for quite limited data sets. However, the growing number of well-established space physics models and the large number of space physical quantities of interest calls for a more systematic approach to the evaluation of the model performances. More specifically, it is of interest to compare the performances of various models and model setups side by side by using quantities that reflect different aspects of geospace. Further, instead of qualitative visual impressions,

[1]Goddard Earth Sciences and Technology Center, University of Maryland, Baltimore, Maryland, USA.
[2]NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.
[3]Department of Physics, Catholic University of America, Washington, DC, USA.
[4]Department of Atmospheric, Oceanic, and Space Sciences, University of Michigan, Ann Arbor, Michigan, USA.
[5]Space Science Center and Physics Department, University of New Hampshire, Durham, New Hampshire, USA.
[6]Space Weather Prediction Center, NOAA, Boulder, Colorado, USA.

0148-0227/10/2009JA014537$09.00

Table 1. Geospace Storm Events Studied in This Work[a]

Event	Date and Time	min(Dst)	max(Kp)
1	29 Oct 2003 0600 UT to 30 Oct 0600 UT	−353 nT	9
2	14 Dec 2006 1200 UT to 16 Dec 0000 UT	−139 nT	8
3	31 Aug 2001 0000 UT to 1 Sep 0000 UT	−40 nT	4
4	31 Aug 2005 1000 UT to 1 Sep 1200 UT	−131 nT	7

[a]Min(Dst) and max(Kp) give the minimum Dst index and the maximum Kp index of the event, respectively.

the model performances need to be rigorously quantified by means of metrics that make the interpretation of the validation results unambiguous. Consequently, the purpose of this work is to start to move toward the more systematic space physics model validation process [see also *Siscoe et al.*, 2004; *Spence et al.*, 2004]. A set of global models are run for a number of geospace storm events and the model output are compared to the observations in a systematic fashion. The work described in this paper was carried out in preparation for the Geospace Environment Modeling 2008–2009 Challenge that will eventually provide more in depth comparisons for a much larger number of models and a larger set of different geospace parameters. All simulations discussed in this work were carried out at the Community Coordinated Modeling Center (CCMC) operated at NASA Goddard Space Flight Center. The simulations are publicly available for analysis via CCMC's visualization interface.

[5] In section 2 the setup up for the model evaluation is explained. Section 3 describes the models used in the analyses. In section 4 the approaches used to obtain the predicted ground and geostationary magnetic field perturbations are described and the techniques used to carry out the comparisons to the observations are explained. Furthermore, section 4 describes the results of the analyses. Section 5 provides conclusions regarding the systematic evaluation of the ground and geostationary magnetic field predictions generated by the global magnetohydrodynamic models.

2. Evaluation Setup

[6] Four geospace storm events listed in Table 1 were chosen for the study. Solar wind bulk plasma and the interplanetary magnetic field observations carried out by SWEPAM and MAG instruments on board Advanced Composition Explorer (ACE) for the events are shown in Figure 1. Note that due to problems with the SWEPAM instrument during the October 2003 event (event 1), only low temporal resolution plasma data could be constructed [*Skoug et al.*, 2004]. The events 1 and 2 are well-known CME-related major storm events while events 3 and 4 are smaller storm events associated with much more subtle changes in the solar wind driving (see Figure 1).

[7] For each event the evaluation of the model prediction performances was carried out by means of model versus data comparisons for the following geospace parameters: (1) ground magnetic field perturbations and (2) magnetic field perturbations at the geosynchronous orbit. The observations for parameter 1 were provided by ground-based magnetometer networks. 12 ground magnetometer stations given in Table 2 were selected for the evaluation process (see also Figure 2). The magnetometer stations were selected based on the global spatial and temporal coverage. The observations for the parameter 2 were in turn provided by geostationary GOES s/c. One minute magnetic field data from four different GOES, GOES 8, GOES 10, GOES 11 and GOES 12 were available for the analysis. Due to varying data coverage, subsets of two GOES were used in the analyses associated with each individual event. One of the GOES for each storm interval was located over the east coast of the United States at approximately 75 degrees west longitude (GOES 8 or GOES 12) and the other over the west coast at approximately 135 degrees west longitude (GOES 10 or GOES 11).

3. Models

[8] In this work two model suites hosted at CCMC that provide both ground and geostationary magnetic perturbations are used: (1) Space Weather Modeling Framework (SWMF), developed at the University of Michigan, and (2) OpenGGCM, developed at the University of New Hampshire.

[9] SWMF contains nine space physics components covering various regions between the Sun and the Earth [*Tóth et al.*, 2005]. Simulations presented in this paper utilize only three components of SWMF: Global Magnetosphere (GM) MHD model BATS-R-US [*Powell et al.*, 1999], Ridley Ionosphere Electrodynamics (IE) electrostatic potential solver [*Ridley et al.*, 2004], and Rice Convection Model (RCM) of the inner magnetosphere [*De Zeeuw et al.*, 2004]. The BATS-R-US simulation grid is composed of self-similar three-dimensional rectangular blocks arranged in varying degrees of spatial refinement levels. The Earth's dipole tilt is updated with time. The IE solver receives field-aligned currents mapped from the inner boundary of the GM domain, solves electrostatic current continuity equation, and delivers convection velocity at the boundary back to GM. The height-integrated conductance model used by the IE solver includes the effects of solar EUV ionization and empirical relationship between the field-aligned currents and auroral oval conductances. Two BATS-R-US versions are analyzed: version 7.73 that has been used extensively at CCMC since 2005 and a new version 8.01 recently delivered to CCMC. The version 8.01 can incorporate also the RCM. The RCM modifies the plasma pressure distribution in the inner magnetosphere of the MHD code. Since BATS-R-US solves for the magnetic field self-consistently with the pressure distribution, this modification of pressure changes the field and therefore changes the field-aligned currents into the ionosphere (which then modifies the ionospheric potential). Two different magnetospheric grids are tested: a grid with 700,000 cells used in CCMC's experimental real-time simulations, and a grid with 2 million cells used in standard CCMC runs on request.

[10] The global magnetosphere component of the OpenGGCM suite solves the MHD equations on a stretched Cartesian grid [*Raeder et al.*, 2001a]. It is coupled with an IE solver and a Coupled Thermosphere Ionosphere Model (CTIM) [*Fuller-Rowell et al.*, 1996]. CTIM is a sophisticated three-dimensional dynamical model of the thermosphere-ionosphere system that replaces the empirical ionospheric conductance model with first-principle calculations. In the

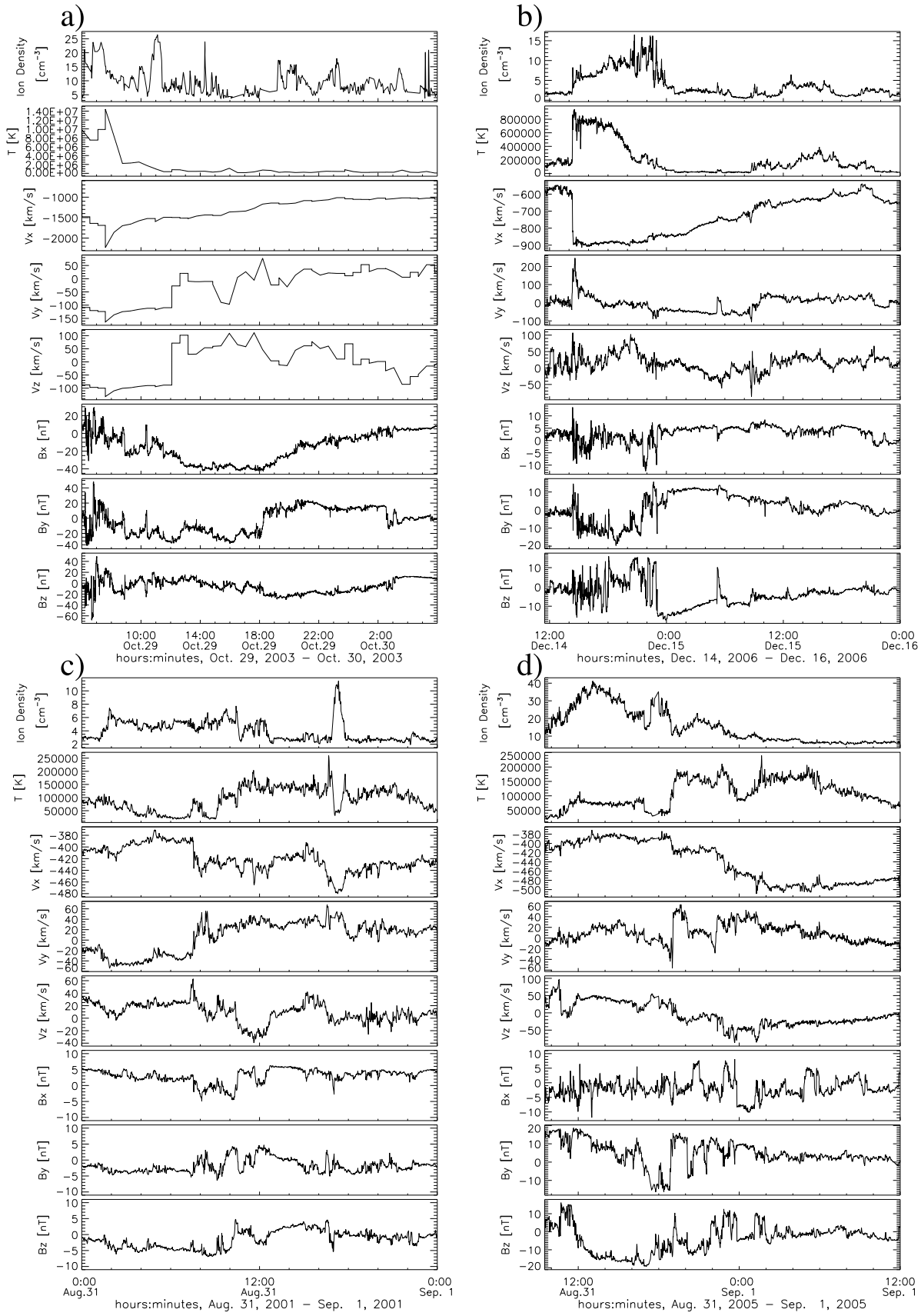

Figure 1. Solar wind bulk plasma and the interplanetary magnetic field observations for the studied storm events, (a) event 1, (b) event 2, (c) event 3, and (d) event 4, given in Table 1. See the text for details.

Table 2. Locations of the Magnetometer Stations Used in the Study

Station Code	Geomagnetic Latitude	Geomagnetic Longitude
YKC	68.9	299.4
MEA	61.6	306.2
NEW	54.9	304.7
FRN	43.5	305.3
IQA	74.0	5.2
PBQ	65.5	351.8
OTT	55.6	355.3
FRD	48.4	353.4
HRN	73.9	126.0
ABK	66.1	114.7
WNG	54.1	95.0
FUR	48.4	94.6

OpenGGCM simulations carried out in this work, the Earth's dipole tilt is not updated with time. Consequently, the dipole tilt angle in the XZ GSE plane is set to the minimum value for the simulation time interval, while the dipole tilt angle in the YZ GSE is set to the average value.

[11] For the analyses carried out in this work four different model setups were selected. These are as follows: (1) BATS-R-US with IE, version 7.73 with minimum magnetospheric cell width of 0.25 Re and 700,000 grid cells; (2) BATS-R-US with IE, version 7.73 with minimum magnetospheric cell width of 0.25 Re and 2 million grid cells; (3) BATS-R-US with RCM and IE, version 8.01 with minimum magnetospheric cell width of 0.25 Re and 2 million grid cells; and (4) OpenGGCM with IE and CTIM, version 3.1 with minimum magnetospheric cell width of 0.3 Re and 3 million model cells. The main difference between 700,000 and 2 million cell BATS-R-US grids is that in the latter the near-Earth high-resolution part of grid extends to the magnetopause and further down to the tail. Table 3 lists the models and the model configurations. All models were run for events in Table 1 using Advanced Composition Explorer (ACE) observations of the plasma and interplanetary magnetic field. ACE observations were propagated to the model inflow boundaries by means of simple convection delay. Note that the OpenGGCM setup 4 did not run successfully for event 1 (see Table 1) and thus the corresponding modeled data is not available for the analysis. The models and the model setups 1–4 are referred in this work simply as different "models" although in fact only two different global MHD models were used.

4. Analysis

4.1. Ground Magnetic Field Perturbation Modeling

[12] Ground-based magnetic field observations provide means to remotely sense the behavior of the near-space electric currents. Due to the combined spatial and temporal nature of the ground magnetic field observations, the recordings contain significant amount of information about the solar wind-magnetosphere-ionosphere system. Consequently, it is fair to state that a model that successfully reproduces the observed features of the ground magnetic field fluctuations has captured some of the most challenging elements of the physics associated with the geospace. Further, the ground magnetic field is the central quantity in the determination of the geomagnetically induced currents (GICs) that flow in long technological conductor systems [e.g., *Boteler et al.*, 1998; *Molinski*, 2002]. A model that is capable of reproducing the observed spatiotemporal properties of the magnetic field has potential to reproduce the central features of GICs.

[13] The ground magnetic field was computed from 1 min temporal resolution IE output by using the methods described by *Pulkkinen et al.* [2007]. The ground conductivity model of central Finland [*Viljanen et al.*, 1999] was used as a generic approximation applied for all stations. It is noted that the approach by *Pulkkinen et al.* [2007] takes into account also perturbations due to field-aligned currents and induced (telluric) currents flowing inside the earth. The amount of magnetic signal originating from the induced currents is dependent on number of factors such as the ground conductivity structure, the source structure and the distance from the source. However, as a very rough rule of a thumb, about 20–40% of the ground magnetic field can be of internal origin during active periods [*Tanskanen et al.*, 2001; *Pulkkinen and Engels*, 2005]. It should also be noted that the modeled ground magnetic fields are only due to the high-latitude ionospheric currents provided by global MHD; for example, the low-latitude ionospheric output from RCM in model 3 was not used in the ground magnetic field computations.

[14] Due to the computational constraints, only ionospheric currents within 1000 km radius of each location to which magnetic field is computed are taken into account. While this is acceptable when studying the horizontal components of the magnetic field, vertical component may be poorly estimated [*Yu and Ridley*, 2008]. Consequently, only horizontal components of the ground magnetic field are used below.

4.1.1. Preparation of the Data

[15] One minute temporal resolution magnetic field recordings for stations in Table 2 were downloaded via INTERMAGNET (http://www.intermagnet.org) and the data were transformed into geomagnetic coordinates. The quiet time baseline level was determined visually for each station and for each event and the baseline was removed from the magnetic field data to obtain the disturbance field. Small data gaps with length of no more than few minutes were patched by means of linear interpolation. The modeled magnetic field data were resampled by means of spline interpolation to match the time stamps of the observations.

4.1.2. Model Performance Metrics

[16] Several different methods were used to evaluate the quality of the ground magnetic field predictions. First, the observed and the modeled horizontal magnetic field data were depicted on polar plots. These plots were used as an initial visual spatiotemporal assessment of the model performances. The quantitative analysis was carried out by applying two different metrics. For clarity, it is emphasized that the term metric is not used in a strict mathematical sense in this work. Rather, the term is used more generally to refer to functions mapping two elements of a set (e.g., time series of observed and modeled ground magnetic fields) into a single real number. The first used metric is the prediction efficiency PE defined for a discrete signal $x(t_i)$ as

$$PE = 1 - \frac{\langle (x_{obs} - x_{mod})^2 \rangle_i}{\sigma^2_{obs}} \quad (1)$$

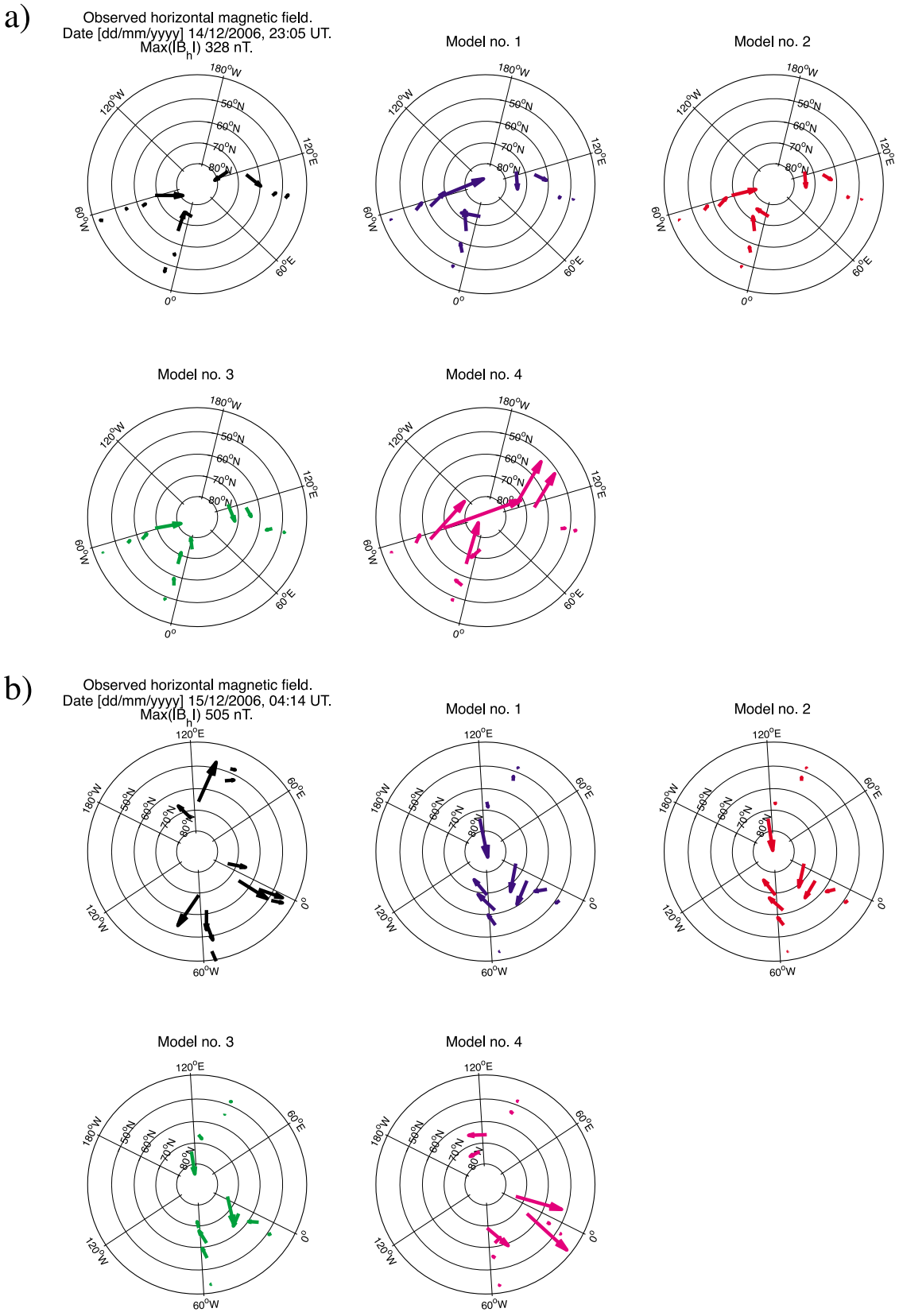

Figure 2. Snapshots of the observed and the modeled horizontal magnetic field vectors for two time instants at stations indicated in Table 2. The sources of the data in the polar plots are indicated. See Table 3 for model descriptions and color codes. Data in geomagnetic coordinates.

Table 3. Global MHD Models and the Model Setups Used in the Study

Model	Model Description[a]	Color Code[b]
1	BATS-R-US v7.73, grid: 0.25 Re, 700000 total	blue
2	BATS-R-US v7.73, grid: 0.25 Re, 2 million total	red
3	BATS-R-US v8.01 coupled to RCM, grid: 0.25 Re, 2 million total	green
4	OpenGGCM v3.1, grid: 0.3 Re, 3 million total	magenta

[a]"Grid" refers to the minimum magnetospheric cell width and the total number of cells.
[b]The color codes associated with each model. In the text, different model setups are referred as different "models."

where x_{obs} and x_{mod} are the observed and the modeled signals, respectively, $\langle \ldots \rangle_i$ indicates arithmetic mean (all means taken in this work are arithmetic means) take over i and σ^2_{obs} is the variance of the observed signal. Note that $PE = 1$ indicates a perfect prediction while $PE = 0$ means that the model predicts the signal equally well to a model that uses the mean value of the signal as a predictor.

[17] The models' ability to reproduce the observed ground magnetic field was assessed also by computing the power spectrum of the fluctuations. The power spectra for individual time series were computed as follows. First the time series was divided into 2 h long segments with 50% overlap between neighboring segments. The data within each segment was multiplied with the Hann window function [*Press et al.*, 1992, p. 554] prior to the fast fourier transform and the final spectrum was obtained by taking an average over the segments.

[18] The second applied metric was developed to quantify the models' capability to generate the GIC-related features of the magnetic field power spectra. For this, it was assumed that the horizontal geoelectric field ($\tilde{E}_x$, $\tilde{E}_y$) can be calculated from the observed or modeled ground magnetic field accurately by means of a plane wave approach [see, e.g., *Pirjola*, 2004]. In the plane wave approach the spectral domain electric field is obtained by multiplying the spectral domain magnetic field ($\tilde{B}_x$, $\tilde{B}_y$) by the surface impedance $\tilde{Z}(\omega)$ (tilde denotes quantities in the spectral domain and ω denotes the angular frequency). Further, it was assumed that GIC can be obtained from the relation

$$\tilde{GIC}(\omega) = a\tilde{E}_x + b\tilde{E}_y = \frac{a}{\mu_0}\tilde{Z}\tilde{B}_y - \frac{b}{\mu_0}\tilde{Z}\tilde{B}_x \tag{2}$$

where a and b are the system parameters that depend on the topology and the electrical properties of the investigated conductor system and μ_0 is the vacuum permeability. For power spectra the relation (2) takes the form of an inequality [*Pulkkinen et al.*, 2006]

$$|\tilde{GIC}(\omega)| \leq |\frac{a}{\mu_0}||\tilde{Z}||\tilde{B}_y| + |\frac{b}{\mu_0}||\tilde{Z}||\tilde{B}_x| \tag{3}$$

Motivated by the inequality (3), one then computes the logarithm of the ratios of the right-hand side of equation (3) and assumes $|a| = |b|$ to introduce m_s defined as

$$m_s(\omega) = \log\left[\frac{|\tilde{B}_x|_{obs} + |\tilde{B}_y|_{obs}}{|\tilde{B}_x|_{mod} + |\tilde{B}_y|_{mod}}\right] \tag{4}$$

which is a dimensionless quantity characterizing ability of a model to reproduce the GIC-related magnetic field fluctuations. To obtain a single number, one finally computes

$$M_s = \sqrt{\frac{1}{N}\sum_{\omega} m_s^2} \tag{5}$$

where N is the number of frequencies. M_s in equation (5) is a metric, called "log-spectral distance," measuring the positive definite distance from the perfect ($M_s = 0$) model performance.

4.1.3. Results

[19] Figure 2 shows snapshots of the observed and the modeled horizontal magnetic field vectors for two time instants. Figure 2a is an example of a "good" case, which indicates some obvious visual agreement between the observed and the modeled fields. Figure 2b in turn shows an example of a "bad" case where the observed and the modeled field patterns are very different. Although this is not a quantitative result, a visual inspection of the polar plots for all four storm events indicated that especially during the most intense phases of the storms the agreement between the observed and the modeled fields tends to be predominantly of the type in Figure 2b. Animations S1–S4 show the polar plots for each event.[1]

[20] Figure 3 shows the prediction efficiencies for stations in Table 2 for all four storm events. It is seen that there is a tendency indicating that the ground magnetic field behavior is more difficult to predict for stations in the midrange of the studied geomagnetic latitudes. Taking into account the storm time expansion of the auroral oval to lower latitudes, this may be an indication that, in agreement with the findings by *Raeder et al.* [2001b], the fluctuations associated with the auroral current systems are more difficult to predict than the fluctuations associated with the polar cap and subauroral currents. Also, it is seen that there are quite large differences between the predictability of different events: models 1–3 perform the best for event 3 while the model 4 performs the best for event 2. Generally, in agreement with the visual impression obtained from inspecting the polar plots in Figure 2, 1 min temporal resolution magnetic field variations are very challenging to predict as indicated by predominantly negative prediction efficiencies.

[21] The main observation of interest from Figure 3, however, is that there are systematic differences in the performances between different models. More specifically, prediction efficiencies tend to increase as the model complexity (in terms of grid resolution and breadth of the included physics) is increased from the model 1 to model 3 (see Table 3) and

[1]Auxillary materials are available in the HTML. doi:10.1029/2009JA014537.

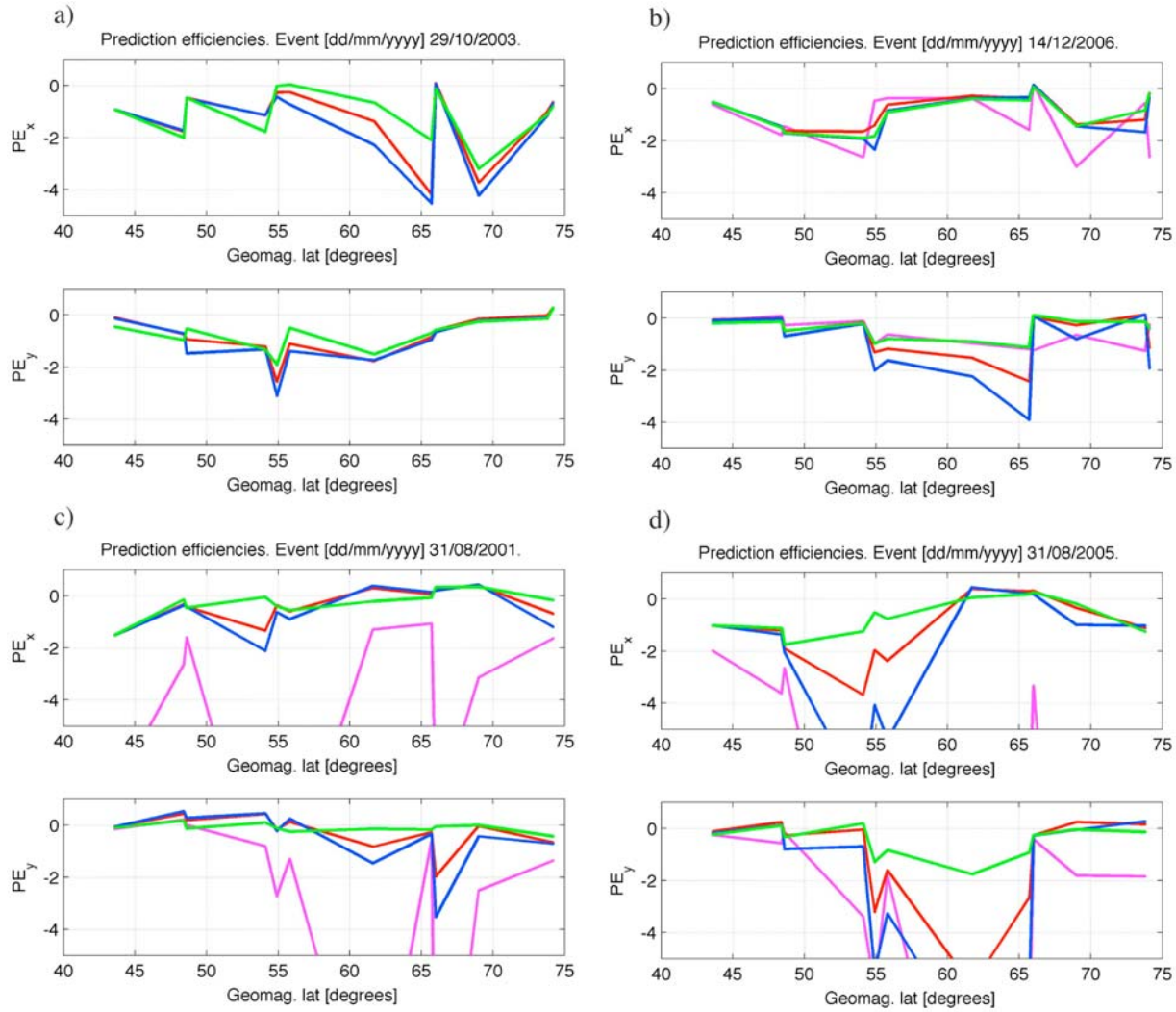

Figure 3. Prediction efficiencies for stations in Table 2 (stations ordered as a function of their geomagnetic latitude) for the storm events, (a) event 1, (b) event 2, (c) event 3, and (d) event 4, given in Table 1. The event dates are indicated. PE_x and PE_y are prediction efficiencies associated with horizontal magnetic field components, and different colors indicate efficiencies associated with different models. The color codes corresponding to different models are given in Table 3.

model 4 tends to provide poorer performance in comparison to models 1–3. To compress the information in Figure 3 into a single number, mean prediction efficiencies taken over different stations and field components were calculated. The mean prediction efficiencies are shown in Table 4, from which the systematic differences between the model performances observed from Figure 3 are easily confirmed.

[22] Figure 4 shows the mean (taken over stations) power spectra between the periods 2–120 min for the observed and the modeled horizontal magnetic field fluctuations for all storm events. Clearly, the general shape of the observed spectra is captured quite well by all models. However, especially models 1–3 tend to generate too little power at higher frequencies. Similar to prediction efficiencies in Figure 3, it is seen from Figure 4 that the higher-frequency portion of the power spectra is captured better as model complexity increases from the model 1 to model 3. However, in contrast to the prediction efficiency, model 4 seems to reproduce the observed power spectra best for all four events. These observations are confirmed from Table 5, which shows the log-spectral distance for each model for each event: except for a single value in the event 4 results, there is a systematic monotonic decrease of M_s for all events as one moves from model 1 to model 4.

4.2. Modeling Magnetic Field at the Geosynchronous Orbit

[23] The magnetic field at geostationary orbit is not only one of the key quantities in the studies of the inner mag-

Table 4. Mean Prediction Efficiencies[a]

Event	Model 1	Model 2	Model 3	Model 4
1	−1.25	−1.06	−0.91	NaN
2	−1.09	−0.79	−0.71	−0.96
3	−0.51	−0.32	−0.19	−4.74
4	−2.18	−1.19	−0.59	−6.24

[a]Mean is taken over horizontal magnetic field components and stations.

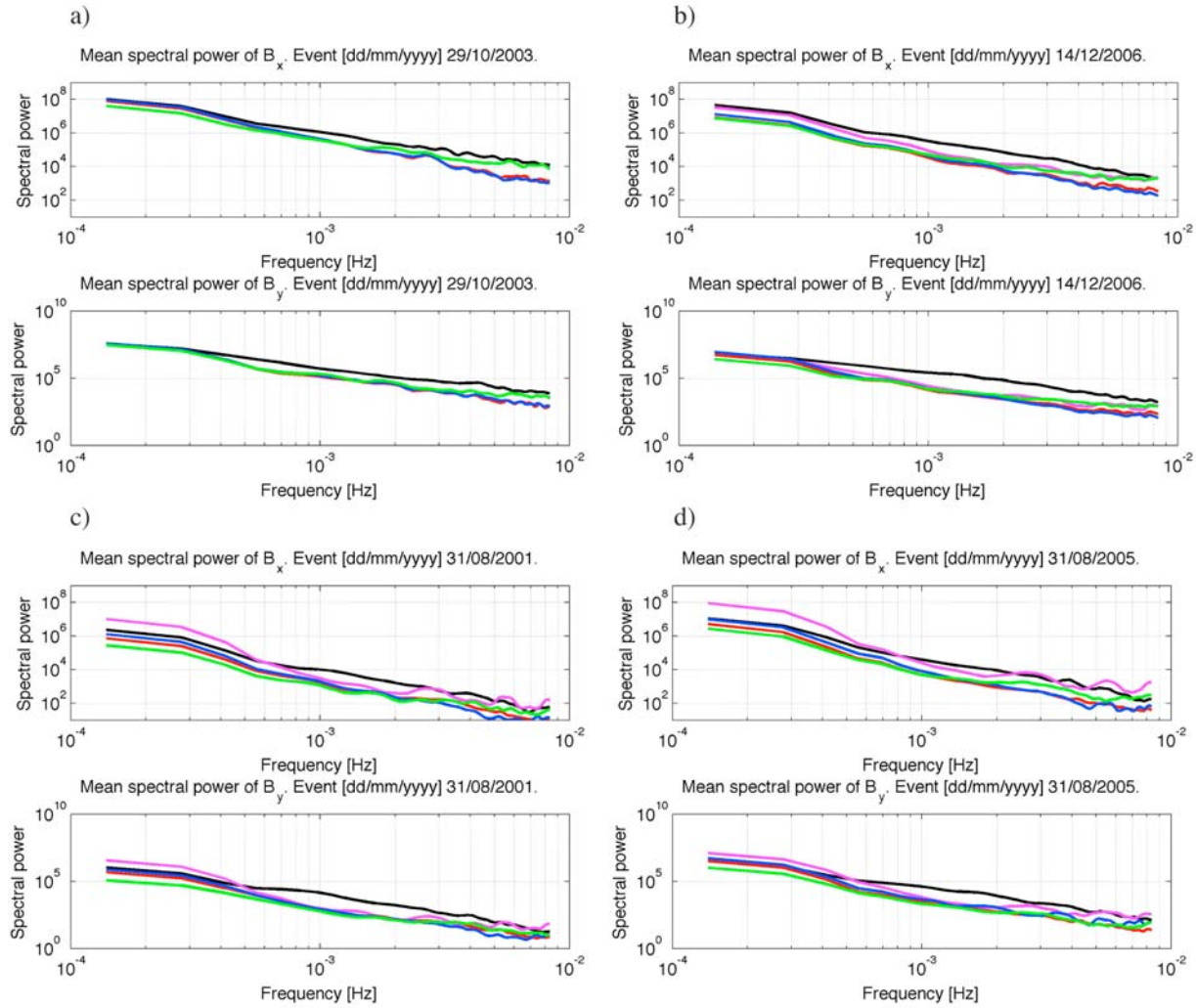

Figure 4. Mean (taken over stations in Table 2) power spectra for different horizontal magnetic field components for the storm events, (a) event 1, (b) event 2, (c) event 3, and (d) event 4, given in Table 1. The event dates are indicated. Different colors indicate the power associated with different models. Black curves indicate the observed spectra. The color codes corresponding to different models are given in Table 3.

netosphere but is also very important for characterizing the general state of the magnetosphere. For example, field dipolarizations due to substorm-related internal magnetospheric dynamics and dayside field compression due to dynamic pressure exerted by the solar wind on the magnetosphere are central features of geospace phenomena. Further, geostationary observations of the magnetospheric magnetic field are carried out continuously, most notably by the GOES, and the data covers extended time periods. Consequently, reproducing the observed storm time geostationary magnetic field fluctuations is one of the major goals in modern magnetospheric modeling.

[24] The global MHD models in Table 3 provide the magnetic field also in the inner magnetosphere. GOES orbit information was used to extract individual time series from the original global MHD data. Depending on the availability of the GOES data, different spacecraft were used for different storm events.

4.2.1. Preparation of the Observational Data

[25] One minute temporal resolution magnetic field data for applicable GOES satellites were downloaded from CDAWeb (GOES magnetometer; key parameters, the only type of data available). The used spacecraft are GOES 10 and GOES 12 for the event 1, GOES 11 and GOES 12 for the event 2, GOES 8 and GOES 10 for the event 3 and GOES 10 and GOES 12 for the event 4.

Table 5. Metric M_s for Ground Magnetic Field Perturbations Computed From Mean Power Spectra in Figure 4

Event	Model 1	Model 2	Model 3	Model 4
1	7.07	7.03	3.34	NaN
2	9.81	8.99	5.54	5.00
3	7.01	6.08	5.32	1.52
4	6.31	7.65	5.33	0.28

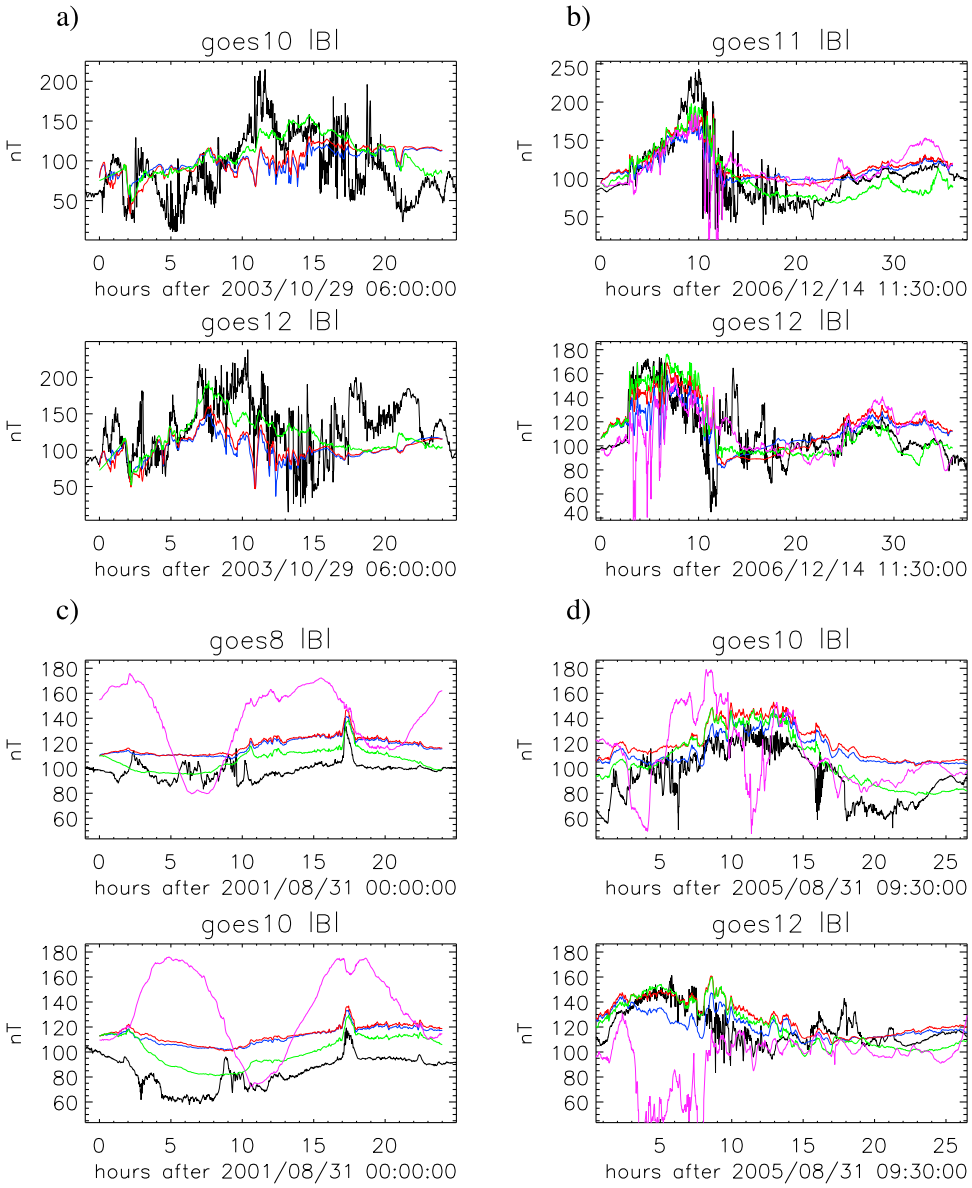

Figure 5. The modeled and the observed absolute magnitude of the magnetic field at GOES locations for the storm events, (a) event 1, (b) event 2, (c) event 3, and (d) event 4, given in Table 1. The spacecraft and the event dates are indicated. Different colors indicate the field associated with different models. Black curves indicate the observed field. The color codes corresponding to different models are given in Table 3.

[26] Small data gaps in observations and modeled data were interpolated to a regular 1 min cadence. In the power spectrum analysis any 2 h window that contained periods of negative B_z (indicating periods in the magnetosheath) in either the observations or the model results (B_z traces not shown) was excluded from the analysis. The presence and the timings of magnetopause crossings will be subject to a separate study. *Tsyganenko et al.* [2003] found that z component (GSM coordinates) of the GOES 8 magnetic field observations have a systematic offset of about 7 nT, which was subtracted here from the data prior to the analysis.

4.2.2. Model Performance Metrics

[27] As for the ground magnetic field, several different methods were used to evaluate the geostationary magnetic field predictions. First, the observed and the modeled 1 min absolute magnitude of the magnetic field were depicted on individual plots corresponding to each event and to each GOES used in the analyses. These plots were used as an initial visual assessment of the model performances. Again as for the ground magnetic field, the quantitative analysis was carried out by applying two different metrics. The first used metric is the prediction efficiency *PE* defined by equation (1). The geostationary absolute magnitude of the magnetic field was used in all metrics computations. Second, power spectra of the field fluctuations were calculated. The power spectra for individual time series were computed as was explained in section 4.1.2. Finally, the difference between the observed and the modeled geostationary mag-

Table 6. Mean Prediction Efficiencies[a]

Event	Model 1	Model 2	Model 3	Model 4
1	−0.32	−0.25	0.12	NaN
2	0.39	0.41	0.51	0.10
3	−6.06	−6.95	−1.49	−72.81
4	−0.17	−0.53	0.16	−4.29

[a]Mean is taken over GOES spacecraft.

netic field spectra was quantified in terms of log-spectral distance defined by equations (4) and (5).

4.2.3. Results

[28] Figure 5 shows the observed and the modeled geostationary magnetic field fluctuations for two GOES for each event in Table 1. Qualitatively, both storm time field behavior and the model performances are seen to vary from event to event. For example, whereas for event 2 all models seem to capture especially the low-frequency field fluctuations quite well, for event 3, despite the relatively smooth observed field behavior, most of the models have significant offset throughout the event.

[29] Table 6 shows the mean (mean over different GOES) prediction efficiencies for the data in Figure 5. The main observation of interest is, again similarly to the ground magnetic field part of the study, that there are systematic differences between the performances of different models. More specifically, prediction efficiencies tend to increase as the model complexity is increased from model 1 to model 3 (see Table 3) and model 4 tends to provide poorer performance in comparison to models 1–3. It is noted that increasing model complexity was observed to improve geosynchronous magnetic field predictions by *Huang et al.* [2006], who used

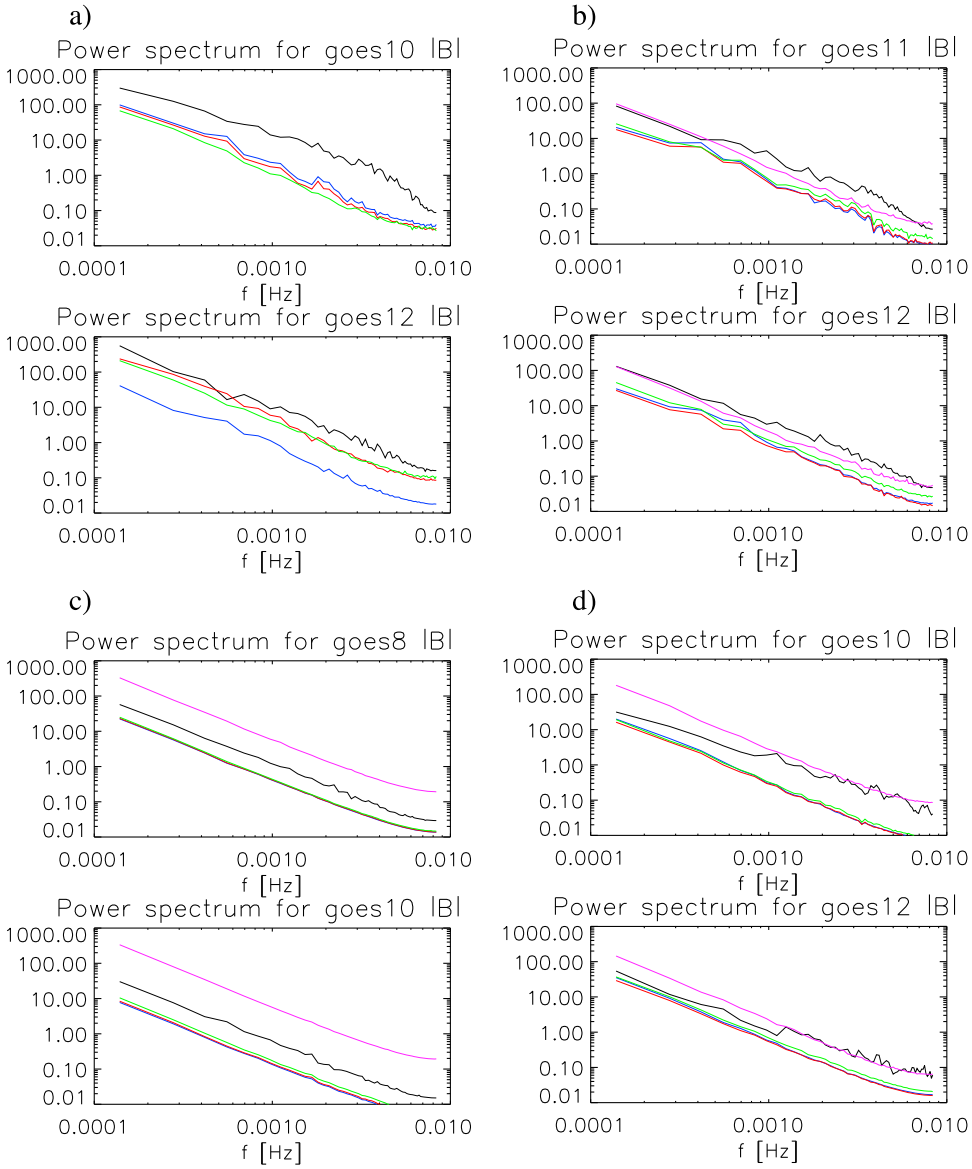

Figure 6. Spectra of the modeled and the observed geostationary magnetic field fluctuations in Figure 5 for the storm events, (a) event 1, (b) event 2, (c) event 3, and (d) event 4, given in Table 1. Different colors indicate the power associated with different models. Black curves indicate the observed spectra. The color codes corresponding to different models are given in Table 3.

Table 7. Metric M_s for Geostationary Magnetic Field Perturbations Computed From Power Spectra in Figure 6

Event	Model 1	Model 2	Model 3	Model 4
1	1.43	1.35	1.24	NaN
2	0.98	0.90	0.71	0.97
3	1.29	1.11	0.65	1.25
4	1.74	1.71	1.51	0.65

Lyon-Fedder-Mobarry global MHD code not utilized in this work. As seen from Table 6, because of the systematic model offsets (see Figure 5c), event 3 for which models 1–3 performed the best in terms of ground magnetic field perturbations, is clearly the most difficult event to predict efficiently. Further, prediction efficiencies are significantly better for event 2 in comparison to the other three events for all models.

[30] Figure 6 shows the power spectra between the periods 2–120 min for the observed and modeled geostationary magnetic field fluctuations for all storm events. The general shape of the observed spectra is captured quite well by all models. However, models 1–3 generate too little power throughout the spectrum whereas model 4 overestimates the power in the field fluctuations for the events 3 and 4. Table 7 shows the log-spectral distance for each model for each event and it is seen that there is a systematic monotonic decrease of M_s for all events as one moves from the model 1 to the model 3. For event 4 model 4 gives the smallest M_s while for the other events the performance of model 4 is in terms of the log-spectral distance between the models 1 and 2.

5. Discussion

[31] In this work a systematic evaluation of the ground and geostationary magnetic field predictions generated by a set of global magnetohydrodynamic models was carried out. It was seen that despite the different aspects of geospace represented by the two parameters, there were significant similarities between the metrics analyses of the two. More specifically, for both parameters and for the both metrics used, i.e., prediction efficiency and log-spectral distance, the performance improved as the model complexity was increased from model 1 to model 3 (see Table 3). This implies that increase in the spatial resolution and inclusion of more realistic inner magnetospheric physics makes the model better represent the physics of the system. Improving the spatial resolution in the near-Earth region improves the model's capability to generate realistic electric currents that map into the ionosphere. Consequently, from the ground magnetic field viewpoint having sufficient spatial resolution especially in the vicinity of the inner boundary of the MHD (the difference between models 1 and 2) is of particular importance. Similarly, introducing the kinetic inner magnetospheric physics is important particularly from the geostationary magnetic field viewpoint. As seen from Table 6, coupling RCM to MHD improves the model performance significantly for all events. The improvement is at least partly due to more realistic plasma pressure in the inner magnetosphere that, for example, causes the magnetosphere to inflate [see *De Zeeuw et al.*, 2004]. The effect of the inflation is seen clearly from Figure 5c where model 3 has a systematically smaller field magnitude and smaller offset with respect to the observations than the other models.

[32] Model 4 had a tendency to perform poorer in comparison to models 1–3 in terms of the prediction efficiency. However, for the ground magnetic field and for one of the storm events for geostationary magnetic field model 4 outperformed models 1–3 in terms of the log-spectral distance. The numerical scheme used to solve the MHD equations in the model 4 makes the model less diffusive than models 1–3 (for details on the used numerical schemes, see the references given above). This feature of the model 4 likely explains the models tendency to generate more field fluctuations, with comparable grid resolutions, in comparison to models 1–3. The presence of these fluctuations provides more accurate representation of the observed spectral characteristics of the magnetic field particularly at the high-frequency range of the studied frequencies. However, highly fluctuating modeled magnetic field easily generates large differences (in the sense of equation (1)) to the observed field if, for example, the spatial pattern of otherwise accurately modeled ionospheric currents is slightly rotated with respect to the true pattern.

[33] As seen from Tables 4–7, the performance of the models varied from event to event. For example, prediction efficiencies for ground magnetic field predictions were significantly better for event 3 for models 1–3. As event 3 was the smallest storm of the studied events, this is an indication that high-latitude ionospheric currents are more predictable for less active periods. This is understandable as during high-activity conditions ionospheric currents experience amplification of irregular, and possibly less predictable, fluctuations associated with, for example, substorms. In contrast, as seen from Table 6, geostationary magnetic field fluctuations were predicted the most efficiently by all models for the strong storm event 2. The positive prediction efficiencies associated with the event indicate that geostationary magnetic field fluctuations can in fact be quite predictable by means of global MHD modeling, especially if MHD output is modified by kinetic inner magnetospheric models.

[34] The work described in this paper was carried out in preparation for the Geospace Environment Modeling (GEM) 2008–2009 Challenge that will eventually provide more in-depth physics-based comparisons for a much larger number of models and a larger set of different geospace parameters. Further, alternative metrics such as event-based utility metric [e.g., *Weigel et al.*, 2006] that may be of more value for the users of space weather applications will be considered. The fundamental purpose of the GEM 2008–2009 Challenge is to quantify, for the given evaluation setup, the current state of the space physics modeling capability and to address the differences between various modeling approaches.

[35] **Acknowledgments.** The global MHD simulations used in this work were carried out at the Community Coordinated Modeling Center (CCMC) operated at NASA Goddard Space Flight Center. The authors wish to acknowledge the rest of the CCMC staff for their generous support throughout the work discussed in the paper. The results presented in this paper rely on data collected at magnetic observatories. We thank the national institutes that support them and INTERMAGNET for promoting high standards of magnetic observatory practice (http://www.intermagnet.org). Terry Onsager and Paul Loto'aniu of NOAA are acknowledged for their help with selection of the ground magnetometer stations used in the study and preparation of the GOES magnetic field data, respectively.

[36] Zuyin Pu thanks George Siscoe and another reviewer for their assistance in evaluating this paper.

References

Boteler, D. H., R. J. Pirjola, and H. Nevanlinna (1998), The effects of geomagnetic disturbances on electrical systems at the Earth's surface, *Adv. Space Res.*, *22*, 17–27.

De Zeeuw, D. L., S. Sazykin, R. A. Wolf, T. I. Gombosi, A. J. Ridley, and G. Tóth (2004), Coupling of a global MHD code and an inner magnetospheric model: Initial results, *J. Geophys. Res.*, *109*, A12219, doi:10.1029/2003JA010366.

Fok, M.-C., R. Wolf, R. Spiro, and T. Moore (2001), Comprehensive computational model of Earth's ring current, *J. Geophys. Res.*, *106*(A5), 8417–8424.

Fuller-Rowell, T. J., D. Rees, S. Quegan, R. J. Moffett, M. V. Codrescu, and G. H. Millward (1996), A coupled thermosphere ionosphere model (CTIM), in *Solar-Terrestrial Energy Program: Handbook of Ionospheric Models*, edited by R. W. Schunk, pp. 217–238, Utah State Univ., Logan.

Huang, C.-L., H. E. Spence, J. G. Lyon, F. R. Toffoletto, H. J. Singer, and S. Sazykin (2006), Storm-time configuration of the inner magnetosphere: Lyon-Fedder-Mobarry MHD code, Tsyganenko model, and GOES observations, *J. Geophys. Res.*, *111*, A11S16, doi:10.1029/2006JA011626.

Luhmann, J., S. Solomon, J. Linker, J. Lyon, Z. Mikic, D. Odstrcil, W. Wang, and M. Wiltberger (2004), Coupled model simulation of a Sun-to-Earth space weather event, *J. Atmos. Sol. Terr. Phys.*, *66*(15–16), 1243–1256, doi:10.1016/j.jastp.2004.04.005.

Molinski, T. (2002), Why utilities respect geomagnetically induced currents, *J. Atmos. Sol. Terr. Phy.*, *64*, 1765–1778.

Pirjola, R. (2004), Review on the calculation of surface electric and magnetic fields and of geomagnetically induced currents in ground-based technological systems, *Surv. Geophys.*, *23*, 71–90, doi:10.1023/A:101481600930.

Powell, K. G., P. L. Roe, T. J. Linde, T. I. Gombosi, and D. L. De Zeeuw (1999), A solution-adaptive upwind scheme for ideal magnetohydrodynamics, *J. Comput. Phys.*, *154*(2), 284–309, doi:10.1006/jcph.1999.6299.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery (1992), *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed., Cambridge Univ. Press, Cambridge, U. K.

Pulkkinen, A., and M. Engels (2005), The role of 3D geomagnetic induction in the determination of the ionospheric currents from the ground geomagnetic data, *Ann. Geophys.*, *23*, 909–917.

Pulkkinen, A., A. Viljanen, and R. Pirjola (2006), Estimation of geomagnetically induced current levels from different input data, *Space Weather*, *4*, S08005, doi:10.1029/2006SW000229.

Pulkkinen, A., M. Hesse, M. Kuznetsova, and L. Rastätter (2007), First-principles modeling of geomagnetically induced electromagnetic fields and currents from upstream solar wind to the surface of the Earth, *Ann. Geophys.*, *25*, 881–893.

Raeder, J., Y. Wang, and T. Fuller-Rowell (2001a), Geomagnetic storm simulation with a coupled magnetosphere-ionosphere-thermosphere model, in *Space Weather: Progress and Challenges in Research and Applications*, *Geophys. Monogr. Ser.*, vol. 125, edited by P. Song, H. J. Singer, and G. Siscoe, pp. 377–384, AGU, Washington, D. C.

Raeder, J., Y. Wang, T. Fuller-Rowell, and H. J. Singer (2001b), Global simulation of space weather effects of the Bastille Day storm, *Sol. Phys.*, *204*, 325–338.

Ridley, A. J., T. I. Gombosi, and D. L. De Zeeuw (2004), Ionospheric control of the magnetospheric configuration: Conductance, *Ann. Geophys.*, *22*, 567–584.

Siscoe, G., D. Baker, R. Weigel, J. Hughes, and H. Spence (2004), Roles of empirical modeling within CISM, *J. Atmos. Sol. Terr. Phy.*, *66*(15–16), 1481–1489, doi:10.1016/j.jastp.2004.03.028.

Skoug, R. M., J. T. Gosling, J. T. Steinberg, D. J. McComas, C. W. Smith, N. F. Ness, Q. Hu, and L. F. Burlaga (2004), Extremely high speed solar wind: 29–30 October 2003, *J. Geophys. Res.*, *109*, A09102, doi:10.1029/2004JA010494.

Spence, H., D. Baker, A. Burns, T. Guild, C.-L. Huang, G. Siscoe, and R. Weigel (2004), Center for integrated space weather modeling metrics plan and initial model validation results, *J. Atmos. Sol. Terr. Phys.*, *66*(15–16), 1499–1507, doi:10.1016/j.jastp.2004.03.029.

Taktakishvili, A., M. Kuznetsova, M. Hesse, L. Rastätter, A. Chulaki, and A. Pulkkinen (2007), Metrics analysis of the coupled Block Adaptive-Tree Solar Wind Roe-Type Upwind Scheme and Fok ring current model performance, *Space Weather*, *5*, S11004, doi:10.1029/2007SW000321.

Tanskanen, E., A. Viljanen, T. Pulkkinen, R. Pirjola, L. Häkkinen, A. Pulkkinen, and O. Amm (2001), At substorm onset, 40% of AL comes from underground, *J. Geophys. Res.*, *106*(A7), 13,119–13,134.

Tóth, G., et al. (2005), Space Weather Modeling Framework: A new tool for the space science community, *J. Geophys. Res.*, *110*, A12226, doi:10.1029/2005JA011126.

Tóth, G., D. L. De Zeeuw, T. I. Gombosi, W. B. Manchester, A. J. Ridley, I. V. Sokolov, and I. I. Roussev (2007), Sun-to-thermosphere simulation of the 28–30 October 2003 storm with the Space Weather Modeling Framework, *Space Weather*, *5*, S06003, doi:10.1029/2006SW000272.

Tsyganenko, N. A., H. J. Singer, and J. C. Kasper (2003), Storm-time distortion of the inner magnetosphere: How severe can it get?, *J. Geophys. Res.*, *108*(A5), 1209, doi:10.1029/2002JA009808.

Vasyliunas, V. M. (1970), Mathematical models of magnetospheric convection and its coupling to the ionosphere, in *Particles and Fields in the Magnetosphere*, edited by B. M. McCormac, pp. 60–71, D. Reidel, Dordrecht, Netherlands.

Viljanen, A., R. Pirjola, and O. Amm (1999), Magnetotelluric source effect due to 3D ionospheric current systems using the complex image method for 1D conductivity structures, *Earth Planets Space*, *51*, 933–945.

Wang, H., A. J. Ridley, and H. Lühr (2008), Validation of the Space Weather Modeling Framework using observations from CHAMP and DMSP, *Space Weather*, *6*, S03001, doi:10.1029/2007SW000355.

Weigel, R. S., T. Detman, E. J. Rigler, and D. N. Baker (2006), Decision theory and the analysis of rare event space weather forecasts, *Space Weather*, *4*, S05002, doi:10.1029/2005SW000157.

Yu, Y., and A. J. Ridley (2008), Validation of the space weather modeling framework using ground-based magnetometers, *Space Weather*, *6*, S05002, doi:10.1029/2007SW000345.

A. Chulaki, M. Hesse, M. Kuznetsova, A. Pulkkinen, and L. Rastätter, NASA Goddard Space Flight Center, Code 674, Greenbelt, MD 20771, USA. (antti.a.pulkkinen@nasa.gov)

J. Raeder, Space Science Center, University of New Hampshire, 245G Morse Hall, 39 College Rd., Durham, NH 03824, USA.

A. Ridley, Department of Atmospheric, Oceanic, and Space Sciences, University of Michigan, Ann Arbor, MI 48109, USA.

H. J. Singer, Space Weather Prediction Center, NOAA, 325 Broadway, Boulder, CO 80305, USA.