Geospace Model Evaluation
to Support Model Transition to Operations.
Phase I Report:
Ground Magnetic Field Perturbations dB/dt.
(Version 2013/08/31)

Community Coordinated Modeling Center

M. Kuznetsova, A. Pulkkinen, L. Rastätter, A. Glocer

CCMC, NASA Goddard Space Flight Center

————

Abstract.

This report evaluates the models' capability to reproduce observed ground magnetic field fluctuations, which are closely related to geomagnetically induced current phenomenon. One of the primary motivations of the work is to support NOAA SWPC in their selection of the next geospace numerical model that will be transitioned into operations. This report is based on two papers submitted to the Space Weather Journal [*Pulkkinen et al.*, 2013] and [*Rastätter et al.*, 2013b].

6 geomagnetic events and 12 geomagnetic observatories were selected for validation. While modeled and observed magnetic field time series are available for all 12 stations, the primary metrics analysis is based on 6 stations that were selected to represent the high-latitude and mid-latitude locations. Events-based analysis and the corresponding contingency tables were built for each event and each station. The elements in the contingency table were then used to calculate Probability of Detection (POD), Probability of False Detection (POFD) and Heidke Skill Score (HSS) for rigorous quantification of the models' performance.

In support of the project the CCMC developed a routine to calculate magnetic perturbations ΔB from snapshots of the current systems that are being produced by coupled models of the global magnetosphere-ionosphere system. Magnetic perturbations on the ground are calculated from currents in the magnetosphere, from field-aligned currents between the magnetosphere and the ionosphere, and the Hall- and Pedersen currents in the ionosphere.

The results of the post-processing tool were compared with runtime $\Delta\mathbf{B}$ calculations within the Space Weather Modeling Framework (SWMF) model. The same routine was applied for all global magnetosphere models.

The summary results of the metrics analyses are reported in terms of POD, POFD and HSS. Event by event contingency tables are provided in the appendix. An online interface built at CCMC is available for more detailed time series analyses.

1. Introduction

The geomagnetically induced current (GIC) problem [e.g., *Boteler et al.*, 1998; *Pirjola*, 2005] has received elevated international interest over the past 3-4 years, especially in terms of the potential impact on high-voltage power transmission systems. The current worst-case scenarios range from wide-scale voltage and system collapses [*North American Electric Reliability Corporation*, 2012] to catastrophic loss of a large number of high-voltage power transformers [*National Research Council*, 2008]. While better quantification of the hazard will require additional interdisciplinary science and power engineering investigations, it is commonly accepted that the problem is serious enough that actions need to be taken for mitigating the impact. Consequently, the space weather modeling and forecasting community is responding to this elevated need by supporting the operational utilization of the latest advancements in the science. More specifically, the community needs to work on new regional or even local predictions of the geomagnetic environment pertaining to GIC. Initial steps towards this goals have been taken both on the empirical and first-principles-based modeling fronts [e.g., *Weigel et al.*, 2003; *Wintoft*, 2005; *Weimer et al.*, 2010; *Pulkkinen et al.*, 2009; *Zhang et al.*, 2012]. The next logical step is to transition the new scientific capability into an operational setting, which is the primary motivation of the work at hand.

Understanding model capabilities to reproduce observed features in the signal of interest is a key element of space weather monitoring and forecasting. Quantification of the model performance becomes critical as one moves from the research to operational environment where inaccurate model estimates and misleading error bars can potentially lead to poor

and costly decisions by the end-user. Consequently, detailed model validation, tests for model robustness and general quality checks (e.g., model response to bad input data) are a central part of model transition to operations and of general interest to operational space weather forecasting entities.

Acknowledging the importance of rigorous model validation and building on the earlier reports on the topic [*Pulkkinen et al.*, 2010, 2011; *Rastätter et al.*, 2011], as well as the excellent work on geopace model validiaton conducted under the auspices of the Geospace Environment Modeling (GEM) Metrics and Validation Focus Group, NOAAs Space Weather Prediction Center (SWPC) requested the Community Coordinated Modeling Center (CCMC) evaluate geospace models available at the CCMC for possible transition to operations. This effort included the participation of model developers, as well as the CCMC, SWPC, and through GEM, the broader scientific community. Planning and discussions with modelers and the scientific community were held at GEM, the annual Space Weather Workshop in Boulder, and at meetings of the American Geophysical Union. One benefit of building on previous work done by the GEM Geospace Environment Modeling Challenge, is that, over time, we will be able to track model improvements as new and improved versions of existing models, or new models, are delivered to the CCMC.

In contrast to earlier GEM efforts on the topic the focus of the latest model validation effort was to study the models' capability to reproduce the observed "dB/dt events," i.e. rapid fluctuation of the ground magnetic field. The primary argumentation for studying dB/dt is that the time derivative of the ground magnetic field (referred to as "dB/dt") can be used as an indicator for the level of geomagnetically induced electric field, or geoelectric field, on the surface of the Earth [e.g., *Viljanen et al.*, 2001]. The geoelectric field, in

turn, is the primary physical quantity driving GIC. Consequently, although numerous additional complexities such as ground conductivity, conductor system configuration and other engineering details including high-voltage power transformer design are critical for more detailed assessment of the threat, dB/dt can be used as an indicator for a potential GIC hazard. Further, if data from an upstream monitor such as NASA's Advanced Composition Explorer (ACE) is used to produce dB/dt, one can generate short lead-time (15-30 min) forecast estimates of the potential hazard.

The validation work discussed in this report is a result of extensive collaboration between the geospace modeling community, NOAA SWPC and CCMC. The definition of the validation setting, selection of metrics and the general validation process were discussed comprehensively and agreed as the work progressed over the past approximate 2 years. All intermediate results of the analyses carried out by CCMC were communicated to the community and modelers and it was made certain that the model installations and tools at CCMC were acceptable to all participating groups. Generally, the validation process was made as transparent as possible including early communication of NOAA SWPC criteria for selecting models entering the transition process.

2. Selected events

Six geospace storm events listed in Table 1 were chosen for the study. Four of the events (events 1-4 in Table 1) were used in the earlier GEM Challenges [*Pulkkinen et al.*, 2010, 2011; *Rastätter et al.*, 2011] and two new "surprise events" not communicated to the modelers prior to the model and model setup delivery to CCMC were added to the list. The two new events were selected jointly by CCMC and NOAA SWPC scientists. Solar wind bulk plasma and the interplanetary magnetic field observations carried out

by Solar Wind Electron, Proton, and Alpha Monitor (SWEPAM) and MAG instruments onboard Advanced Composition Explorer (ACE) for the events are shown in Fig. 1. Due to limitations of the SWEPAM instrument during the October 2003 event (event 1), only low temporal resolution plasma velocity data could be constructed [*Skoug et al.*, 2004]. Further, the plasma density data for the event were obtained from the Geotail Plasma Wave Instrument. Earlier GEM Challenge events 1 and 2 are well-known coronal mass ejection-related major storm events and events 3 and 4 are less active periods associated with much more subtle changes in the solar wind driving. Events 1-4 are from the solar cycle 23. The new surprise event 5 is one of the first CME-related events of the cycle 24 and was of special interest due to the very large substorm event that was associated with the storm. Event 6 in turn was the first severe storm of the cycle 24.

Solar wind observations were propagated to model inflow boundaries by ballistic propagation and the x-component (Geocentric Solar Magnetospheric coordinate system) of the interplanetary magnetic field was set to zero. While solar wind propagation constitutes a source for modeling errors, the same solar wind input was used for all models, thus introducing identical uncertainty in the specification of the inflow boundaries.

3. Physical quantity used for the validation study

For each event in Table 1, the model performance was evaluated by comparing the observed vs predicted ground dB/dt. Throughout the paper dB/dt is defined as

$$dB/dt = \sqrt{(dB_x/dt)^2 + (dB_y/dt)^2} \tag{1}$$

where B_x and B_y indicate the two horizontal components of the magnetic field. One-minute geomagnetic observatory recordings were used to provide the observed signal. One-minute temporal resolution magnetic field recordings were downloaded via INTER-MAGNET (www.intermagnet.org). The data were transformed from geographic coordinates, as provided by INTERMAGNET, into geomagnetic dipole coordinates. IGRF 2000 coefficients were used to compute the coordinate transformation matrices as given by *Hapgood* [1992]. 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. Methods for computing dB_x and dB_y (delta-B) from model outputs are described below. The modeled delta-B were resampled by means of spline interpolation to match the time stamps of the observations.

4. Magnetometer stations

Following the GEM Challenges, 12 geomagnetic observatories (magnetometer stations) listed in Table 2 and shown in Fig. 2 were selected based on the global spatial and temporal coverage. Station PBQ was discontinued November 2007 and replaced by station SNK. Consequently, for events 5 and 6 station SNK was used in place of PBQ.

6 stations were selected out of the original 12 GEM Challenge stations to represent the high-latitude and mid-latitude locations. The selected high-latitude stations are PBQ/SNK, ABK and YKC and mid-latitude stations WNG, NEW, OTT (see Table 1 and Fig. 2). The selected 6 stations represent all three meridional chains used in the earlier Challenges and have equal weight on mid- and high-latitude locations. No observed data was available for station ABK for event 5. Although all 12 stations are available for

both observation and model prediction data sets and can be viewed via CCMC's online model validation interface only the above 6 stations are used in the results discussed in this report.

5. Models

Five models were evaluated in the model validation activity. These included empirical models by D. Weimer (Virginia Polytechnic Institute) and R. Weigel (George Mason University) and major US global magnetohydrodynamic (MHD) models from University of Michigan, Center for Integrated Space Weather Modeling (CISM) and University of New Hampshire. Also the Finnish Meteorological Institute's Grand Unified Ionosphere-Magnetosphere Coupling Simulation (GUMICS) global MHD group participated in the discussions associated with the validation work. However, while GUMICS is available for runs-on-request at CCMC, due to the current serial implementation of the model the group did not participate in the validation of the model itself. The full parallel version of GUMICS is expected to be available before the next round of operational geospace model validation.

Below each model and settings pertaining to the validation activity are described. Table 3 summarizes some of the key features of each individual model. A version of the Weimer model and all global MHD models discussed in this work are available at CCMC for runs-on-request.

5.1. Weimer empirical ground magnetic field prediction model

The empirical model supplied by Daniel Weimer provides values of the magnetic perturbations at the surface of the Earth, for vector components in the North, East, and

Vertical (down) directions. Vectors are returned as a function of location in either geographic latitude/longitude coordinates, or corrected geomagnetic latitude/magnetic local time coordinates. More specifically, the model internally uses "Modified Apex" coordinates [*Richmod*, 1995; *Emmert et al.*, 2010; *VanZandt et al.*, 1972]. If geographic locations are specified on input, then the output vectors are also in geographic coordinates such that positive North is toward the geographic pole; otherwise North is toward the corrected geomagnetic apex pole. For the purpose of the challenge the output vectors are rotated into geomagnetic coordinates. The data needed to drive the model are the Geocentric Solar Magnetospheric (GSM) y- and z-components of the interplanetary magnetic field (IMF), the solar wind velocity, the dipole tilt angle of the Earth's magnetic field, and the solar F10.7 index. Ideally, 25-minute mean values of the IMF and solar wind velocity should be used, with a 20-minute delay after propagation to the magnetosphere's bow shock [*Weimer and King*, 2008].

The model uses Spherical Cap Harmonic Analysis (SCHA) [*Haines*, 1985] within a cap that extends down to 33.4 degrees apex latitude. SCHA coefficients up to order $m = 3$ and degree $l = 16$ are used. Measurements of the IMF and solar wind on the ACE spacecraft, from February 1998 through December 2005, were used to generate the model, as well as measurements from over 120 magnetometer stations. Details of data preparation and initial tests, are provided by *Weimer et al.* [2010]. A least-error-fit was used to find how each of the SCHA coefficients varies as a linear function of the input values, with 17 terms for each coefficient. In order to handle the non-linear "saturation" response of the ionosphere, the fits are derived separately within 23 bins, divided according to the magnitude of the IMF.

More recently an improved version of this model has been developed, but it was not available for the validation work discussed in this paper [*Weimer*, 2013]. The latest version extends down to the geomagnetic equator, using data from 143 magnetometer stations. It uses spherical harmonics up to degree $l = 31$, and divides the IMF measurements into 29 bins.

5.2. Weigel empirical ground magnetic field prediction model

Three models were used for this study. All of the models were developed using available 1-minute ground magnetometer measurements from World Data Centre for Geomagnetism (Edinburgh) [(http://www.wdc.bgs.ac.uk/) in the time interval January 1, 2000 through December 31, 2006 excluding event intervals that occurred in this time range.

The first model is the climatological average of ΔB (disturbance field) for each component computed by taking the average of ΔB in 48 local times.

The second and third models are linear impulse response filters using the method of *Weigel* [2007]. Both models predict geomagnetic disturbance G (ΔB or dB/dt) in vector direction i using

$$G_i(t, LT) = h_{\Delta, LT} + \sum_{t'=0}^{Nc} vB_s(t - t')h(t', LT) \tag{2}$$

where the solar wind velocity v, and the rectified $-z$ component of the interplanetary magnetic field, $B_s = 1/2(|B_z| - B_z)$, are from the OMNI high-resolution data set (http://omniweb.gsfc.nasa.gov/ow_min.html), and the h coefficients depend on local time, LT. The model coefficients are determined using a least-squares minimization of the prediction error.

The model for $G_i = \Delta B_i$ has $Nc = 12$ and predicts G_i at 48 local times. Physically, this model predicts the ground magnetic field given the past six hours of solar wind measurements. Data were resampled to place the predictions on a 1-minute time grid.

The model that predicts $G = dB/dt$ has $Nc = 4$ and predicts G at 1440 local times. The model $G = dB/dt$ was used in the metrics analyses carried out in this paper.

5.3. Space Weather Modeling Framework

The Space Weather Modeling Framework (SWMF) [*Tóth et al.*, 2005, 2012] (http://csem.engin.umich.edu/swmf) is a flexible software framework designed to model a variety of space physics phenomena. The SWMF divides the complex space physics systems into physics domains. The domains used in the SWPC modeling challenge are the Global Magnetosphere (GM), the inner magnetosphere (IM) and the Ionosphere Electrodynamics (IE).

The GM model is the Block-Adaptive Tree Solarwind Roe-type Upwind Scheme (BATS-R-US) [*Powell et al.*, 1999; *Gombosi et al.*, 2004]. In the work described in this paper the semi-relativistic MHD equations [*Gombosi et al.*, 2002] are solved. We use an explicit/implicit time stepping scheme [*Tóth et al.*, 2006] with a 5 second time step (potentially reduced by the adaptive time step control scheme if necessary). The computational domain extends from $32\,R_E$ upstream to $-224\,R_E$ downstream in the x direction and $\pm 128\,R_E$ in the y and z coordinates (GSM). The inner boundary is at $2.5\,R_E$ distance from the center of the Earth. The domain is discretized with a block-adaptive Cartesian grid. The roughly 1 million grid cells vary in size from $1/4\,R_E$ near the inner boundary to $8\,R_E$ in the distant tail. The boundary conditions are the usual (see referred publications for details), except for the density at the inner boundary, which is set as

$\rho_{\text{inner}} = 28 + 0.1\,\text{CPCP}$ where CPCP is the average of the northern and southern cross polar cap potentials measured in keV, and the density is measured in amu/cm^3. In these runs, we used the artificial wind scheme [*Sokolov et al.*, 2002] with Koren's limiter [*Koren*, 1993] ($\beta = 1.2$) and the 8-wave scheme [*Powell et al.*, 1999].

The IM domain is represented by the Rice Convection Model (RCM) [*Wolf et al.*, 1982; *Toffoletto et al.*, 2003]. RCM solves for the bounce averaged and isotropic but energy resolved particle distribution of electrons and various ions. We used the standard RCM settings except for one modification: we added an exponential decay term to the RCM equations, so that the phase space density decays towards zero with 10 hour e-folding rate. With this modification the Dst index of the coupled model recovers better after large storms.

The IE domain is represented by the Ridley Ionosphere Model (RIM) [*Ridley et al.*, 2004]. RIM uses the field-aligned currents obtained from GM and the F10.7 flux (set as an input parameter for each event) to calculate particle precipitation and conductances based on empirical relationships. RIM solves a Poisson-type equation for the electric potential on a 2D spherical grid. We set the lower latitude boundary to 10°.

In the work described in this paper the BATS-R-US and RIM models are coupled every 5 seconds, while the BATS-R-US with RCM as well as the RIM to RCM couplings are done every 10 seconds. In the BATS-R-US – RIM coupling the MHD model calculates the field-aligned currents (FAC) at $3\,R_E$ and maps it down to the ionospheric grid. The electric field obtained by RIM, is mapped back to the inner boundary of GM, where the $\mathbf{E} \times \mathbf{B}/B^2$ velocity is calculated. The cross polar cap potentials are also sent to GM, and are used to set the density at the inner boundary. In the RIM to RCM coupling the

electric potential is passed and interpolated onto the RCM grid. In the BATS-R-US to RCM coupling BATS-R-US finds the closed field line region and calculates field volume integrals with an efficient parallel field line tracing algorithm [*Glocer et al.*, 2009a]. The integrated GM density and pressure are applied as outer boundary conditions for the IM model assuming a 90% H^+ to 10% O^+ number density ratio. In the RCM to BATS-R-US coupling the GM grid cell centers are traced to the inner boundary along the magnetic field lines with an efficient parallel algorithm [*De Zeeuw et al.*, 2004]. The BATS-R-US pressure and density are nudged towards the RCM values with a 20 second relaxation time.

In addition to the basic variables used in the various models, the SWMF can also calculate various plasma parameters along satellite trajectories, ionospheric foot-points of satellites, integrated line-of-sight images, various geomagnetic indexes (Dst, Kp), as well as local magnetic perturbations [*Yu and Ridley*, 2008]. The SWMF can model space weather events starting from the Sun all the way to the Earth [*Tóth et al.*, 2007]. The magnetospheric components of the SWMF have been validated in several studies [*Ridley et al.*, 2002; *Zhang et al.*, 2007; *Glocer et al.*, 2009c; *Welling and Ridley*, 2010]. In addition to the magnetospheric models used in this study, the SWMF also contains the radiation belt and the polar wind components [*Glocer et al.*, 2009a, b], and the CRCM and RAM-SCB inner magnetosphere models [*Buzulukova et al.*, 2010; *Zaharia et al.*, 2010]. The upgraded model version with CRCM and radiation belt components is available at CCMC for runs-on-request since 2012.

5.4. Lyon-Fedder-Mobarry model with Magnetosphere-Ionosphere Coupler and Electrodynamics Solver

The Coupled Magnetosphere Ionosphere Thermosphere Model (CMIT) [*Wang et al.*, 2004; *Wiltberger et al.*, 2004] couples the Lyon-Fedder-Mobarry global magnetosphere model (LFM) [*Lyon et al.*, 2004] with the Thermosphere-Ionosphere-Electrodynamic Global Circulation Model (TIEGCM) [*Roble and Ridley*, 1994] via the Magnetosphere Ionosphere Coupler Solver (MIX) [*Merkin et al.*, 2010] to provide a comprehensive global simulation of geospace response to solar and solar wind drivers. The LFM portion of the model solves the ideal magnetohydrodynamic equations to describe the interaction of the solar wind plasma with the plasma in geospace. This portion of the model requires that the solar wind and IMF conditions be specified typically from ACE or WIND spacecraft observations. These conditions are assumed to be constant along planar fronts propagating through the computational domain. The LFM is electrodynamically coupled to the ionosphere through the MIX model. MIX solves for the cross polar cap potential taking currents from the magnetospheric domain and conductance from the ionosphere. In order to obtain the conductance information MIX uses a series of empirical relationships described in *Wiltberger et al.* [2009] to transform the MHD parameters at the inner boundary into a characteristic energy and flux of precipitating electrons. The ionospheric component uses the electron flux information along with a parameterization of the solar extreme ultraviolet (EUV) flux driven by the F10.7 index to compute the conductance. It is noted that in this validation work TIEGCM was not used. Instead of the full ionosphere-thermosphere system only the ionospheric electrodynamics was treated via MIX. The

model version with TIEGCM has been installed at CCMC in July 2012 and is available for runs-on-request since August 2012.

For the validation work discussed in this paper we used a modest resolution version of the model that provides reliable performance on small amount of computational resources. In the LFM the simulation grid was 53(radial)x48(azimuthal)x64(polar) points allowing for a typical resolution in the inner magnetosphere of roughly 0.4 R_E. The electrodynamic grid used in MIX was 2x2 degrees covering the high latitude region down to a magnetic colatitudes of 45 degrees. The model runs faster than real-time on 24 processors. In the magnetosphere the typical time step is approximately 0.1 seconds. The electrodynamic coupling between the ionosphere and magnetosphere is updated every 5 seconds. The computational model can provide a vast array of information relevant to space weather ranging from the magnetic fields at geosynchronous orbit to the ground magnetic field perturbations.

CMIT and its component models have been used to study a variety of process in geospace ranging from magnetic storms [*Goodrich et al.*, 1998] to substorms [*Lopez et al.*, 1998; *Wiltberger et al.*, 2000] including driving by CMEs [*Baker et al.*, 2004] and CIRs [*Wiltberger et al.*, 2012]. The model was validated by developers against numerous measurements of magnetopause crossings [*Lopez et al.*, 2006; *Garcia and Hughes*, 2007], ground magnetometer observations [*Wiltberger et al.*, 2003] and climatology data from Geotail [*Guild et al.*, 2008a, b]. The version of the model used in this repott does not include an inner magnetosphere model, but coupling with the Rice Convection Model has recently been completed [*Pembroke et al.*, 2012] and will be part of a future release to the CCMC and the next round of operational geospace model validation.

5.5. Open General Geospace Circulation Model

The Open General Geospace Circulation Model (OpenGGCM) global MHD model simulates the interaction of the solar wind with the magnetosphere-ionosphere-thermosphere system. Besides numerically solving the MHD equations with high spatial resolution in a large volume containing the magnetosphere, the model also includes ionospheric processes and their electrodynamic coupling with the magnetosphere. The coupling between the magnetosphere and the ionosphere is an essential part of the model because the ionosphere controls, to a large extent, magnetospheric convection, by providing the resistive closure of the field aligned currents that are generated from the interaction of the solar wind with the magnetosphere [*Raeder et al.*, 1996, 1998]. Processes that occur in the near-Earth region on polar cap and auroral field lines and that are inherently kinetic have been parametrized in the model using empirical relationships. These processes include the field aligned potential drops that are associated with upward field aligned currents, electron precipitation caused by the field aligned potential drops [*Knight*, 1972], and the diffuse electron precipitation that is caused by pitch angle scattering of plasma sheet electrons [*Lyons et al.*, 1979; *Robinson et al.*, 1987; *Weimer et al.*, 1987; *Kennel and Petschek*, 1966]. The electron precipitation parameters and the ionosphere potential are then passed to the Coupled Thermosphere Ionosphere Model (CTIM), which is coupled to the MHD part of the code. CTIM [*Fuller-Rowell et al.*, 1996] is a dynamic model of the ionosphere and thermosphere with a long heritage, covering the globe from 80 km to several 1000 km altitude, and following several neutral and ionic species and their photochemical interactions. CTIM computes self-consistently the ionospheric Pedersen and Hall conductances,

which are then used to solve the ionospheric potential equation [see *Raeder*, 2003, for details].

The OpenGGCM requires as input solar wind and IMF data, and the F10.7 solar radio flux as a proxy for solar UV/EUV radiation. Solar wind and IMF data are ballistically propagated from the monitor location to the upstream boundary of the simulation. Furthermore, we calculate the normal direction of the solar wind fronts using the minimum variance method [*Sonnerup and Cahill*, 1967, 1968] in order to extrapolate the single point measurements to the entire inflow surface. Without such treatment the IMF B_x component cannot change in time without violating $\nabla \cdot \mathbf{B} = 0$ [*Raeder et al.*, 2001c].

The OpenGGCM computes all magnetospheric and ionospheric quantities that are necessary to determine GICs [*Raeder et al.*, 2001a]. Other quantities of interest to space weather can also be derived, such as the total and equivalent ionospheric current, ionosphere electron content, and neutral density affecting low-Earth orbit satellites [*Li et al.*, 2011]. The model has been used for a variety of studies, for example, for the study of substorms [*Raeder et al.*, 2001c, 2008, 2010; *Ge et al.*, 2011; *Gilson et al.*, 2012], storms [*Raeder et al.*, 2001a, b], flux transfer events and dayside reconnection [*Raeder*, 2006; *Dorelli et al.*, 2012; *Muhlbacher et al.*, 2005; *Connor et al.*, 2012], ionospheric convection [*Vennerstrom et al.*, 2005, 2006; *Siscoe et al.*, 2004; *Lu et al.*, 2012], and plasma entry under northward IMF [*Li et al.*, 2005, 2008, 2009, 2011]. A more detailed description of the code and the methods used can be found in *Raeder* [2003] and *Raeder et al.* [2008].

6. Implementation of delta-B computations for physics-based models

In support of the project the CCMC developed a routine to calculate magnetic perturbations delta-B from snapshots of the current systems that are being produced by

coupled models of the global magnetosphere-ionosphere system [*Rastätter et al.*, 2013b]. Magnetic perturbations on the ground are calculated from currents in the magnetosphere, from field-aligned currents between the magnetosphere and the ionosphere, and the Hall- and Pedersen currents in the ionosphere. The same routine was applied for all global magnetosphere models. The calculations are the same regardless of the model that writes the outputs and thus can be easily applied to any additional coupled magnetosphere-ionosphere model supported by the Community Coordinated Modeling Center in the future. The two empirical models (see Table 3) provided direct predictions of the magnetic field at the used station locations.

The delta-B values are calculated from three contributions in the magnetosphere-ionosphere system:

6.1. Magnetosphere currents:

Electric currents from the magnetosphere outside of the "current pickup radius" of the magnetosphere model are used to calculate the magnetic perturbation $\Delta\mathbf{B}$ by using the Biot-Savart formula, similar to *Rastätter et al.* [2013a]:

$$\Delta\mathbf{B} = \frac{\mu_0}{4\pi} \sum \frac{\mathbf{J} \times \mathbf{R}}{\mathbf{R}^3} dV \tag{3}$$

Here, $\mathbf{R} = \mathbf{X}_J - \mathbf{X}_{station}$ is the vector between the position of the current element $\mathbf{X}_J$ and the magnetometer position $(X_{station})$, and dV is the volume element of the grid cell that contains current element $\mathbf{J}$.

Every magnetosphere MHD model has a near-Earth boundary that is separated from the ionosphere electrodynamics module by a certain distance to avoid the very strong magnetic field and large Alfvén velocities that restrict the maximum allowed time step

in the numerical schemes *Powell et al.* [1999]; **?**]; **?**. The inner boundary for the SWMF

runs is located at $2.5R_E$ distance from the Earth's center. Electric currents mapped into

the ionosphere are picked up from a distance slightly farther away from Earth to ensure

that the currents are being taken from within the magnetosphere grid and not from the

boundary. In SWMF this additional distance is $0.5R_E$, which means that field-aligned

currents are picked up at $R_0 = 3R_E$ The margin of $0.5R_E$ is slightly larger than a cell

size (dC) measured diagonally (with $dx = 0.25R_E$: $dC = \sqrt{3}dx \sim 0.433R_E$). Only

magnetosphere grid cells with volumes dV that are centered at positions $\mathbf{R} = [x, y, z]$

beyond R_0 from the Earth's center are considered in the summation of Equation 3.

6.2. Field-aligned currents:

The field-aligned currents (FAC) fill the gap region between the inner boundary of the

magnetospheric model grid ($3R_E$ or $3.5R_E$ from the Earth's center) and the ionosphere

(at 110 km altitude). During post-processing the magnetic field-aligned currents (FAC)

are picked up from the magnetosphere and assumed to follow a dipole magnetic field to

the ionosphere. We used the original model grids in the ionosphere and used the radial

current density (J_r) reported by the ionosphere electrodynamic modules of each model.

The region between 110 km ($r_{\mathrm{iono}} = (1. + 110/6371.2)R_E = 1.01727R_E$) and the pickup

radius of the magnetospheric currents at $R_0 = 3R_E$ is represented by a spherical grid with

an adjustable radial grid spacing and the model's grid in the ionosphere for longitude

(azimuth) angles ϕ between 0 and 2π and latitude angles λ between $-\pi/2$ at the south

pole and $\pi/2$ at the north pole. FAC are arranged on filaments emanating from the

ionosphere grid that run along lines of the dipolar magnetic field $\mathbf{B}_{\mathrm{dip}}$ that determines the

orientation and strength of the FAC in the gap region. In spherical coordinates $(\hat{r}, \hat{\theta}, \hat{\phi})$

the dipole field is

$$\mathbf{B}_{\text{dip}} = \frac{3(\mathbf{m}\cdot\mathbf{r})\mathbf{r} - r^2\mathbf{m}}{r^5} = \frac{1}{r^3}\begin{pmatrix} 2\cos(\theta) \\ -\sin(\theta) \\ 0 \end{pmatrix} = \frac{1}{r^3}\begin{pmatrix} 2\sin(\lambda) \\ -\cos(\lambda) \\ 0 \end{pmatrix} \tag{4}$$

with $\theta = \pi/2 - \lambda$. The magnetic field strength is described by:

$$B_{\text{dip}}(r, \lambda) = \frac{\sqrt{3\sin^2(\lambda) + 1}}{r^3} \tag{5}$$

Positions (r, θ) on dipole field lines satisfy the relation $(\sin^2(\theta)/r = \cos^2(\lambda)/r = \text{constant})$. Using r, r_{iono}, λ_2 and λ_{iono} we can calculate λ_2 using

$$r\cos^2(\lambda_2) = r_{\text{iono}}\cos^2(\lambda_{\text{iono}}). \tag{6}$$

Equation 6 is used to calculate the latitude $\lambda_2(r, \lambda_{iono})$ along a dipole field line on each radial level of the grid (r) for each footprint located at λ_{iono} and r_{iono} in the ionosphere. Each volume element in the FAC region is then computed for each filament (index i) using the half-distance between adjacent filaments (located at index $[j-1]$ and $[j+1]$):

$$dV(i, j) = dr\, (r[i])^2 \frac{\lambda_2[j+1] - \lambda_2[j-1]}{2} \cos(\lambda_2[j]) \tag{7}$$

For the polar axes ($j = 1$: south, $j = N$: north), the expression of Equation 7 is replaced by:

$$dV[i, 1] = dr\, (r[i])^2 \frac{\lambda_2[2] - (-\frac{\pi}{2})}{2} \cos\left(\frac{\lambda_2[1] + (-\frac{\pi}{2})}{2}\right) \tag{8}$$

$$dV[i, N] = dr\, (r[i])^2 \frac{\frac{\pi}{2} - \lambda_2[N-1]}{2} \cos\left(\frac{\lambda_2[N-1] + \frac{\pi}{2}}{2}\right) \tag{9}$$

The finite difference in latitude angle (λ_2) in Equation 7, that involves positions across the pole, is replaced by the latitude difference between the respective pole and the first element of λ_2 adjacent to the pole. The argument of the co-sine in Equation 7 is being replaced by the half distance between the polar axis and the first grid position away

from the polar axis as shown in Equations 8 and 9. This is needed to obtain the volume

elements that fill the region around the axes.

To construct the FAC filaments, we use the radial component of the FAC (J_r) that is

reported by the output provided by the ionosphere electrodynamics module of SWMF. J_r

results from current densities that are encountered at the current pickup radius R_0 in the

magnetosphere and mapped into the ionosphere. To obtain the strength and sign of the

actual field-aligned currents, J_r is divided by the sine of the dipole inclination angle I

$$\sin(I) = -\sin\left[\arctan\left(\frac{2\sin(\lambda)}{\cos(\lambda)}\right)\right] \tag{10}$$

with λ here and in following equations denoting the latitude position on the ionosphere

grid. The ratio of the field strength (Equation 5) at the grid position on the FAC filament

divided by the magnetic field strength at the ionosphere foot point is used to scale the

current at the ionosphere altitude ($J_r(\theta, \phi)/\sin(I)$) to yield the current strength along the

FAC filament:

$$J_{\text{FAC}}(r, \lambda_2(\lambda), \phi) = \frac{J_R(\lambda, \phi)}{\sin(I)} \frac{[3\cos^2(\lambda_2) + 1]^{1/2} r_{\text{iono}}^3}{[3\cos^2(\lambda) + 1]^{1/2} r^3} \tag{11}$$

As with magnetosphere currents, magnetic perturbations from FAC are calculated using

the Biot-Savart formula (Equation 3). In contrast to the D_{st} study [*Rastätter et al.*,

2013a], the contribution of the field-aligned currents to the magnetic perturbations at

the magnetometer stations is non-negligible. The distance vector ($\mathbf{R}$) between the FAC

element and the station position is not always in the poloidal plane unlike the vector from

the Earth's center to the current element in the case of the D_{st} calculation. The sum of all

FAC elements may have nonzero components in each of the three directions. In addition,

we have to consider local magnetic north, east and downward components of $\Delta\mathbf{B}$ at each

station location instead of a single component aligned with the magnetic dipole axis for D_{st}.

6.3. Ionospheric currents:

We used the original model grids in the ionosphere and used the Cartesian components of the ionospheric height-integrated current density (J_x, J_y, J_z) reported by the ionosphere electrodynamic modules of each model. The size of the surface elements dS at each ionosphere grid position $(\theta, \phi) = (\pi/2 - \lambda, \phi)$ are given by:

$$dS = r_{\text{iono}} \sin(\theta) d\,\theta d\,\phi = r_{\text{iono}}^2 \cos(\lambda) d\,\lambda\, d\,\phi \tag{12}$$

We use the Biot-Savart formula (Equation 3) for the ionosphere currents in the same manner as for the magnetosphere but we replace the magnetospheric grid cell volume dV by the surface element dS and the magnetosphere current densities (in units of A/m^2) with the height-integrated ionospheric current densities (in A/m).

6.4. Local magnetic coordinate system:

After combining the three contributions in Solar Magnetospheric (SM) coordinates, the three cartesian components of $\Delta\mathbf{B}$ are then converted to (North, East, Down) or $(-\theta, \phi, -r)$ components in local spherical coordinates at each station. In the "dB/dt" study, the horizontal components (North, East) are used to compute the time derivatives that are then compared to the respective components reported by each station.

7. Milestones of CCMC delta-B tool development

A baseline tool was developed by November 2011. After a review of the outputs and comparisons with SWMF runtime generated time series in summer of 2012 the tool was rewritten to calculate contributions from the magnetosphere, FAC and ionosphere sepa-

rately. This allowed for more efficient debugging. In the course of the rewrite, a parallelized re-execution of calculations was implemented. Several bugs in the calculations were found and corrected. The algorithms for the FAC and ionosphere current contributions were rewritten to use the currents on the original ionospheric model grids. The current version of the tool has been finalized in October 2012. This version was used to produced delta-B time series for all global magnetosphere models evaluated in this report.

8. Delta-B tool testing and sensitivity analysis

The results of the CCMC Delta-B post-processing tool were compared with runtime $\Delta\mathbf{B}$ calculations within the Space Weather Modeling Framework (SWMF) model. Running SWMF at the CCMC enables us to use both the model's own output of $\Delta\mathbf{B}$ (ID: 9a_SWMF) and the global magnetospheric and ionospheric current system post-processing to compute $\Delta\mathbf{B}$ (ID: 9_SWMF).

8.1. Grid to cover field-aligned current region

The region that holds the field-aligned currents (FAC) between the magnetosphere's inner boundary and the altitude of the ionosphere (assumed at 110 km) is filled with a spherical grid. The grid starts with the latitude and longitude positions as defined in the ionosphere electrodynamics module of the SWMF model. In these runs RIM uses a 1-degree spacing in latitude and a 2-degree spacing in longitude.

The radial resolution of the grid that is used to calculate the effects of FAC is a free parameter of the post-processing algorithm as well as in the SWMF model itself. The implementation inside the SWMF model version that was delivered to CCMC uses 800 layers (a number of 400 was mentioned in *Yu et al.* [2010]). To be able to process this

many layers, the model distributes the calculation of $\Delta\mathbf{B}$ among all 64 processors that were used to run the model and eliminates FAC filaments with a current strength that falls below a certain (small) threshold. Thus, the time taken by the $\Delta\mathbf{B}$ calculation is negligible compared to the other tasks performed by the model.

During our separate post processing, we only rely on a few processors to perform the calculation for each time step and therefore we must limit the size of the radial grid. The post-processing tool needs to complete a calculation before the magnetosphere model writes another snapshot of the current system (which is a minute later in the "dB/dt" study). We performed a grid convergence test using grids with a radial spacing of $1/15 R_E$, $1/30 R_E$, $1/60 R_E$ and $1/120 R_E$ corresponding to 29, 59, 118 and 237 layers between $1.017 R_E$ (110 km altitude) and $3 R_E$, respectively. Results are shown in Figure 3 for two stations for the "AGU Storm" (Event 2). Figure 3a and Figure 3b show the portion of the $\Delta\mathbf{B}$ signal from the FAC for Yellowknife (YKC) and Ottawa (OTT), respectively. The black trace is the result obtained by SWMF, and the colored traces are from the different radial resolutions (red: 1/15, yellow: 1/30 and blue: 1/60). On the global scale, the three resolutions are very similar and for most times lie on top of each other (masked by the blue line). Figure 3c and Figure 3d show the difference of the three signals compared to the SWMF model signal.

We see that resolutions of 1/30 and finer yield nearly identical results and only 1/15 differs slightly more (about 5%) from the SWMF results than the other resolutions. A resolution of $1/30 R_E$ is sufficient to provide reasonable results with this algorithm. Finer resolutions yield no better results and consume unnecessary computing cycles. We chose $dr = 1/30 R_E$ for the model challenge described in *Pulkkinen et al.* [2013]. This resolu-

tion allows for real-time calculation of the contribution while providing the best results available.

8.2. Elimination of FAC filaments below a significance threshold

A possible way to save computational time is to eliminate FAC filaments with FAC strengths that fall below a small threshold value (SWMF uses 0.0001 in normalized units). We tested our implementation to see whether implementing a similar elimination made any difference in terms of results and computation time. We found no measurable difference in $\Delta\mathbf{B}$ values that were obtained using the reduced set of FAC filaments compared to the calculation using the full set. With our implementation of the calculation in the Interactive Data Language (IDL$^{\circledR}$ by Exelis Vis), we also did not notice a significant reduction in terms of the execution time. IDL (version 8.1) uses multi-threading and employed up to four processors during the calculation. The lack of speed-up using the elimination procedure may be explained by the fragmentation of large data arrays which are used in the summation. Larger, contiguous data arrays may be processed as fast (or even faster) than several arrays that have a shorter combined length. We chose to keep the full set of FAC filaments to determine the magnetic perturbations and were able to perform the calculation for the twelve stations faster than real time using the radial resolution that is sufficient to provide high-quality results.

8.3. Accuracy of station location in geomagnetic coordinates

From the FAC signal we see that there is a systematic difference between our calculated signals and the results written by the SWMF model. To assess one possible source of the difference we used the magnetometer station locations in geographic coordinates

(GEO) obtained from the INTERMAGNET web site (http://www/intermagnet.org) and the magnetic coordinates (MAG) as specified in a list in Table 2. The list was compiled for the 2008 GEM modeling challenge (http://ccmc.gsfc.nasa.gov/GEM_metrics_08) and the positions in magnetic coordinates were specified using International Geomagnetic Reference Field (IGRF) with parameters valid during 2000-2005. These parameters fit the magnetic field conditions for the four original events defined for the GEM challenge and these positions were used by the SWMF model. Magnetic latitudes and longitudes, however, are time dependent in our post-processing algorithm. We start with geographic coordinates and convert them to magnetic coordinates for each event using the applicable parameters specified by the IGRF for the year of each event. To perform the coordinate transformation in our post processing, we use GEOPACK-2008 (http://geo.phys.spbu.ru/∼tsyganenko/modeling.html) with IGRF-11 [*Finlay et al.*, 2010] coefficients that are definite through year 2010 and constitute extrapolations through year 2015.

Figure 4 shows the effect of the use of geographic (GEO) and magnetic coordinates (MAG) on our calculation relative to the SWMF results for the ionosphere, usually the strongest overall signal (Figure 4a and Figure 4b) and also for the FAC (Figure 4c and Figure 4d). In a global scale (Figure 4a) the ionosphere signal agrees well for the station shown. The difference plots for each $\Delta\mathbf{B}$-component (Figure 4b), however, show that there are considerable differences at times within the event at a level of about 10%-15% of the total signal.

We note that the SWMF approach to transform coordinates with a fixed dipole orientation typically results in an SM station location that is of the order of one de-

gree away from the position calculated by GEOPACK and applicable IGRF coeffi-
cients. This is consistent with GEOPACK and IGRF placing the magnetic pole at lati-
tude=79.6 and longitude=288.4 for 2001/01/01 and at latitude=80.0 and longitude=287.8
for 2010/01/01, which is at least one degree away from the SWMF specification (SWMF's
share/Library/src/ModPlanetConst.f90: geographic latitude = 79.0 degrees, longitude =
289.1 degrees). This fact is demonstrated by plotting the $\Delta\mathbf{B}$ traces for four virtual sta-
tion positions located one degree away north, south, east and west of Yellowknife (YKC)
as shown in Figure 4b). A one-degree displacement of a station does give rise to a change
of the $\Delta\mathbf{B}$ signal that is comparable to the observed discrepancy. For Event 2, the SWMF
signal at YKC resembles best the signal obtained for a slight (less than one degree) north-
ern deviation (compare red trace with yellow) early in the event (on 12/14) and nearly a
one-degree westerly deviation (red trace comparable with the dark blue trace) later in the
event (between 0:00 UT and 12 UT on 12/15 for north and set components). However,
later in the event no correlation can be seen with either of the 4 colored trace for the
adjacent locations.

For the magnetic perturbations computed from FAC, the one-degree displacement of
the station location yields effects that are smaller than the difference with the SWMF
results. Clearly, another source has to be identified to account for the difference.

The contributions of ionosphere and field-aligned currents on $\Delta\mathbf{B}$ can be strongly af-
fected by a mis-representation of a station location. Current filaments may be located
nearly overhead and a shifted station location may change the magnitude and even di-
rection of the magnetic contribution from that filament. Evidence of this can be seen in

Figure 4 where the FAC signal can be vastly different for short times during an event whereas in general the signals are similar.

The magnetosphere contribution is less affected by a possible error in the station location since the magnetic field perturbation exists on a larger spatial scale and the station location only slightly changes the spherical components of $\Delta\mathbf{B}$

8.4. Magnetospheric inner boundary

It was found that the magnetosphere current system near the inner boundary of the LFM model gave rise to an unphysical baseline in the D_{st} index. The additional (time-independent) signal can reach 1000 nT in magnitude when all currents were considered. The amplitude diminishes to about 100 nT for currents beyond $3R_E$ and became negligible ($30 nT$) when currents beyond $4R_E$ only were considered. Therefore we have set the inner boundary for purposes of calculating the magnetosphere contribution and to obtain the field-aligned currents to $4R_E$ for the LFM model. This is $1R_E$ farther than the inner boundary of $3R_E$ used in the CISM-DX [**?**] module that calculates D_{st} and reflects the fact that we are using locations on the Earth's surface and not the center of the Earth.

9. Selected metrics

Based on the earlier GEM Challenge experiences and operational needs in terms of dB/dt prediction capability, it was agreed that the model validation should be built on event-based analyses. An event is defined here as follows: within a forecast window $0 \leq t \leq t_f$, the absolute value of the parameter of interest exceeds an event threshold $|x_{thres}|$ (here dB/dt). The windows are moved over the time series in non-overlapping segments and events for given t_f and $|x_{thres}|$ are recorded for both the measured and the

modeled x. By comparing threshold crossing for both observed and modeled time series one can then build a four-element matrix known as contingency table. The table reports the number of correct hits, false alarms, missed events and correct no events [e.g., *Lopez et al.*, 2007].

In this work the length of the analysis window t_f was selected to be 20 minutes and the thresholds dB/dt .3, .7, 1.1 and 1.5 nT/s were used. The selected thresholds represent values that both span lower and higher ranges of rates of change and are also in the "midrange" in a sense that enough threshold crossing could be detected for good statistics. We carried out systematic sensitivity analyses to study the impact of the selected forecast window length. While predictability of events gets somewhat poorer with shorter window lengths, the ranking of the models did not change significantly as a function of the analysis window length (not shown). Consequently, it was concluded that varying the analysis window length between 10-45 minutes did not change the central results notably.

The elements of the contingency table contain the number of correctly predicted threshold crossings H **(hits)**, the number of false alarms F, the number of missed crossings M and the number of correctly predicted no crossings N. The set $\{H, F, M, N\}$ can be used to compute a number of different metrics quantifying the performance of individual models. In this study three metrics proposed by NOAA SWPC were selected for use in the final analyses. The selected metrics are Probability of Detection (POD), Probability of False Detection (POFD) and Heidke Skill Score (HSS). We describe each metric more in detail in the following subsections.

9.1. Probability of Detection

POD is defined for the set $\{H, F, M, N\}$ as

$$POD = \frac{H}{H + M} \tag{13}$$

The metric measures the fraction of observed threshold crossings which where correctly forecast. It ranges from 0 to 1 with 1 being a perfect score. Since a model providing artificially large signal amplitudes will tend to generate large H and large POD the metric should be used in conjunction with POFD defined below.

9.2. Probability of False Detection

POFD is defined for the set $\{H, F, M, N\}$ as

$$POFD = \frac{F}{F + N} \tag{14}$$

The metric measures the fraction of correctly predicted no crossings that were incorrectly forecast as crossings. POFD ranges from 0 to 1, with 0 being a perfect score. Similar to POD, a model predicting artificially low signal amplitudes will provide low F and small PODF and thus the metric should be used in conjunction with POD.

9.3. Heidke Skill Score

HSS is defined for the set $\{H, F, M, N\}$ as

$$HSS = \frac{2(HN - MF)}{(H + M)(M + N) + (H + F)(F + N)} \tag{15}$$

The metric measures the fraction of correctly predicted threshold crossings after eliminating those predictions that would be correct purely by random chance. It ranges from

negative infinity to 1. Negative values indicate that random forecast is better than the model prediction, 0 indicates no skill (as good as random) and 1 indicates a perfect score.

It is noted that for HSS to be meaningful measure of the model performance a variety of states of the system should be studied. For example, perfect prediction of no 1.5 nT/s crossings for a weak event ($H = M = F = 0$) is reported as $HSS = 0/0$, which is not defined. Consequently, to guarantee well-defined HSS one should be careful to pick thresholds that are crossed for the selected sets of events.

10. Results of Metrics Analysis

To demonstrate a typical storm-time situation, Figs. 5 and 6 show example time series of the observed vs modeled dB/dt for the event 2 (Table 1). Data for all events and stations are viewable via CCMC's online visualization interface accessible at http://ccmc.gsfc.nasa.gov/challenges/dBdt/metrics_results.php. As is seen from the figures, models do not generally capture the dB/dt fluctuations point-by-point, which is not surprising considering the complex waveform of the signal. However, the amplitudes of high-latitude dB/dt fluctuation especially in the beginning of the event are reproduced to a degree by the models. Although all models miss some of the observed activity, for example, at station ABK (Fig. 5) around 26-32 MLT, the capability, at times, to reproduce comparable dB/dt amplitudes, indicates that the models may provide utility in capturing events within given forecast windows. We will quantify this capability to capture the events using metrics discussed in Section 9.

The final metrics-based analyses were carried out for each individual model using events and stations described in Section 2 and the corresponding contingency tables with elements $\{H, F, M, N\}$ were generated for each model for each event and station for dB/dt

thresholds of .3, .7, 1.1 and 1.5 nT/s. Here we will report only the results integrated, i.e. summed contingency table elements, over all events. The summary results are integrated also separately over high-latitude (PBQ/SNK, ABK, YKC) and mid-latitude stations (WNG, NEW, OTT). Figs. 7 and 8 show the corresponding POD, POFD and HSS for all participating models.

The focus of this paper is to report on the process, metrics, and initial results from the evaluation of physics-based and empirical models that predict regional ground-based dB/dt variations during strong geomagnetic activity. Future work is needed to understand where model improvements are needed to better represent observations. At this stage of the work, it is important to quantify model capabilities, and to provide information that will be used to assess whether or not these models provide useful guidance for improved forecasts of regional ground-based magnetic field perturbations. For model settings analyzed in this study it is quite clear that for a given set of stations, events, and metrics, the model 9_SWMF provides the highest POD and HSS for most of the thresholds. As an indication that large dB/dt events are still a challenge to capture accurately, for threshold 1.5 nT/s none of the models is capable of providing POD or HSS greater than 0.5.

We emphasize that for optimal statistics the summary results reported here are obtained by integrating over selected stations and all events. The results and the ranking of the models vary from station to station and event to event. Further, event 1 dominates the statistics for larger dB/dt thresholds due to the strength of the Halloween storm event. The breakdown of the results for individual events (skill score plots and contingency tables) for further analysis are presented at the web page dedicated to the operational geospace model validation at the CCMC website:

http://ccmc.gsfc.nasa.gov/challenges/dBdt/index.php. Skill score plots and contingency table for individual events are also presented in Appendixes A and B in the Supplementary Material attached to this report. Note that statistics for individual events studies may be insufficient for some thresholds and further interpretation of event-by-event results should be done with caution. Figures 9 and 10 show POD, POFD and HSS integrated over the "surprise events" 5 and 6 only.

11. Time line visualization and metrics analysis system

Model verification studies such as the one described in this work have been accompanied by an online visualization and analysis tool that allows a user to plot any subset of available model simulations for a predefined set of time periods ("Events") and calculate skill scores used in the various studies themselves. This tool is being used to quickly compare results for accurate range, data gaps and to add additional runs as they become available to new (and existing) studies.

All data analyzed in this study are viewable via CCMC's on-line visualization interface accessible at **http://ccmc.gsfc.nasa.gov/challenges/dBdt/metrics_results.php** at the special page at the CCMC website dedicated to the operational geospace model validation: http://ccmc.gsfc.nasa.gov/challenges/dBdt/index.php. Fig. 11 shows the screen shot of the entry page to the interface. To access the time line visualization interface the user can click on the name of the selected ground station for the selected event. TIme series also can be downloaded by clicking on the link at the bottom of the entry page. Regional-K time series have been added August 2013 to support the Phase II validation study.

The online visualization consists of three elements: 1) The Data Tree with lists of model runs; 2) The front-end Visualization Interface; 3) The back-end Calculation Software.

The **Data Tree** contains the name of the study (e.g., "GEM2008" for the GEM metics challenges, or "CETI2010" for ionospheric model challenge initiated in 2010). A list of model runs contains the short run ID (e.g. 9_SWMF), followed by the sub-directory that holds the data and a short description that is displayed when data files are available (see Figure 12).

The **Visualization Interface** is a Perl script that uses four basic inputs to descend into the data directory tree to offer the user a selection of available model runs and the variables involved in that study.

- The metrics campaign (e.g., "GEM2008", "CETI2010").

- The event used in the study, such as the time period from the "Halloween Storm" (Event 1) or the "AGU Storm" (Event 2).

- The observatory or satellite used to obtain the data, such as the magnetometer station "YKC", "MEA" (this study) or the "GOES-10" spacecraft in the magnetospheric magnetic field study [*Rastätter et al.*, 2011].

- The type of metrics study selected, currently a number from 1 through 12, where 4 denotes the $\Delta\mathbf{B}$ values used in *Rastätter et al.* [2013b] and 12 denites the dB/dt values used in *Pulkkinen et al.* [2013] and this report. The type of study determines which physical variables are offered for comparison and analysis. A drop-down list displays variables such as $\Delta B_{\mathrm{North}}$, ΔB_H or $|\mathbf{B}|$).

The interface also displays the start and end time of the selected event and a list of model runs that can be selected for display and analysis (Figure 12a). The user can

modify the time range, the variable from the drop-down list and can select the color and line style for each run in the plot. A checkbox can be activated to perform skill score analysis and the vertical plot range can be set explicitly.

The back-end **Calculation Software** that renders the plot images and lists the skill score values is a program written in the Interactive Data Language (IDL by Exelis). The interface creates a driver program in IDL that transmit the selected options to the back-end program. When executed, the program fetches the data and model results, selects or calculates the requested quantity, and performs the analysis (i.e., computes skill scores or threshold-based contingency tables) for each model run for the selected time interval. Results are displayed as an image (or set of images), and a listing of scores (Figure 12b).

In the case of the study of $\Delta\mathbf{B}$ and most other studies, skill scores such as Prediction Efficiency, Correlation Coefficient, and Yield as described in *Rastätter et al.* [2013a] are calculated and listed. In the case of dB_h/dt, the threshold-based contingency tables *Pulkkinen et al.* [2013] are generated for each selected model run. Figure 12b shows a result for a plot of the north component of $\Delta\mathbf{B}$ with skill scores and a plot of the power spectrum for the selected analysis window size and time period.

12. Milestones of model deliveries and run executions

The original deadline for model delivery to CCMC was set for January 31st 2011. Actual delivery or the last update vary depending on particular situation at each modeling team. CCMC had extensive communications with the model developers to guarantee correct installation and to ensure the usage of appropriate settings for each model. Based on a variety of tests such as code robustness carried out at CCMC, model developers provided revisions to the model settings. The final selection of all model settings for most of the

models was accomplished by mid-August 2011. To allow for simulations in a realistic real-time computational environment, it was required that settings for all models were such that the simulations would run not slower then twice the modeled physical time on 64 Beowulf cluster processors. This approach also ensured that model settings are selected to produce comparable speed for the same computational platform and computational resources.

12.1. SWMF

Model release: January 31, 2011. There were no significant modifications in the code after the original delivery date. Modification were required to adjust the performance on CCMC computers using CCMC compilers. Simulation runs for Events 1- 4 were completed in May and June 2011. This model setting is currently implemented for CCMC Runs-on-request. This model settings are also being successfully tested for continuous and event-based real-time simulations (to replace the 2008 version that is continuously running in real-time at the CCMC for years). SInce January 31, 2011 several SWMF upgrades were installed, tested and implemented for CCMC runs-on-request. Upgrades include versions coupled with CRCM and Fok Radiational Belt model. These versions are not included in this study.

12.2. OpenGGCM

Initial model release: Feb. 8, 2011. Model updates were required after the initial robustness tests. The extension for the final delivery was agreed with all code developer. The updated version was delivered in Aug. 2011. The source code was locked since then. Due to various reasons, it was necessary to rerun simulations several time and to make

modifications to simulation setting suggested by model developers. For the simulations
of the event 2 the numerical instability result in generation of earthward flow from the
far-tail boundary (X=-355 R_E) of the simulation box that impacted the whole solution
starting mid-day on Dec. 15, 2006. Similar problems were detected for the event 3.
Subsequently, the model settings for events 2 and 3 were modified by the CCMC to avoid
numerical instabilities development at the simulation box boundaries. The simulations
box was widened and elongated and the model run was completed with $310 \times 144 \times 144$
(6M) cells in a box from $-535R_E < X < 30.01R_E$ and $-62.058R_E < Y, Z < 62.058R_E$.
The box dimensions and the number of cells in each direction were chosen to have a grid
that is as close as possible to the original grid in the original box ($270 \times 120 \times 120$ cells
in $-300.01R_E < X < 30.01R_E$ and $-38.01R_E < Y, Z < 38.01R_E$). Further, during
the verification of runs results in December 2012 it was found that simulated events 1-4
propagated solar wind erroneously twice from ACE spacecraft position ($X = 232R_E$) to
the model upstream boundary at 30 R_E. The error was introduced during the second
model upgrade that unexpectedly included solar wind propagation option in the input file
provided by developers. These solar wind propagation errors were corrected in the revised
simulations.

12.3. LFM

Model release: January 14, 2011. Modification were required to adjust the performance
on CCMC computers using CCMC compilers. Simulation runs for Events 4 was completed
in July of 2011, and for Events 1 and 2 in October and November of 2011, respectively.
Simulations for Event 3 were rerun in December 2012 due to the erroneous solar wind
data propagation delay detected during rigorous model output testing. The cause of the

erroneous solar wind delay is unknown and is being investigated. Note: The process of LFM installation is more cumbersome then for other models.

In August 2013 Michael Wiltberger inspected model results at high latitude stations and reported on August 22, 2013 that among three runs of different versions of the LFM model the version used in the dB/dt study showed significantly weaker responses (Delta-B during the storm) at high latitudes than the other versions also used a CCMC but not submitted by model developers for this study.

Three versions of the model were run at CCMC using the same inputs for the GEM20-08 storms: a) 1_LFM-MIX with version LFM-MIX_1_0_4_pre_valid (delivered in 2010); b) 2_LFM-MIX with version LTR-2_1_1 submitted for this study, and c) 2_CMIT with version LTR-2_1_5, including TIE-GCM ionosphere (delivered to CCMC in June 2012, well-after the dead-line for model submission and not included in this study).

The post processing step that was developed at the CCMC to calculate the Delta-B values from the different current systems was reevaluated. The calculation was found to be correct and no contributions were missing. A rerun of the calculation on the three model runs (for Event 2) did not show any discrepancy compared to the time lines already available for on-line visualization, download and used in the study.

The contribution from the ionosphere current was isolated as being the cause of the difference. An inspection of 2-dimensional distributions of ionospheric conductances at the peak time of the event (2:00 UT on Dec. 15, 2006) revealed that the three runs show considerable difference in the auroral contribution to the Hall conductance. The 2_LFM-MIX run has a very weak Hall conductance in the auroral oval, unlike the two other runs.

Since the Hall currents provide the dominant contribution to Delta-B at high-latitude stations, the large difference is explained.

Inspection of the model's inputs and outputs show that a parameter driving the intensity of the Hall conductance driven by auroral field-aligned currents and precipitation (SIGMA_FACTOR) was indeed different (1 instead of 3 in the 2_CMIT run) in the model version used for the dB/dt and regional K study. The internal parameter was never touched by CCMC run preparation procedures. Therefore the time series used for the report V.20130419 have not been modified. The sensitivity of the results to the ionosphere conductance model is an important fining that should be taken into account in further model improvements.

12.4. Weimer

Model release: April 19, 2011. There were no significant modifications of the source code since the original delivery. Executions to generate Delta-B and dB/dt tImeseries for papers [*Pulkkinen et al.*, 2013] and [*Rastätter et al.*, 2013b] and for the V. 20130419 of the report were performed in April of 2012 (referred in the papers and in the V. 20130419 of the Phase I report as 5_WEIMER).

In the beginning of August 2013 Dan Weimer discovered discrepancies between the results that had been obtained at CCMC with results that he himself generated in February of 2012. In the ensuing investigation three problems were identified, two of which affect already published results of the dB/dt study [*Pulkkinen et al.*, 2013].

• An inspection of the driver program used at CCMC revealed a misprint of a variable name that left the solar wind B_y input to the Delta-B model undefined. The model, in turn, did not halt but quietly used $B_y = 0$.

• The model's outputs as used in the dB/dt study was in geographic coordinates. This did not affect the results of the dB/dt study since time derivatives of the full horizontal magnetic field perturbation were used. For the Phase II validation study of regional K values the East and North components are used and thus need to be provided in the correct coordinates. Dan Weimer provided code to perform the rotation of the horizontal (North, East) components.

• When running both statistical models (Weimer and Weigel Delta-B), it had been assumed that solar wind data should be provided at the Earth's position. The Weimer Delta-B model requires the data at the bow shock location and an additional 20 minutes of delay.

To account for this in the model rerun performed as 6_WEIMER, the driver program was modified to provide a 17-minute delay to the solar wind data before running the model to make up for the difference (20 minute delay minus an average solar wind travel time of 3 minutes from the bow shock to Earth).

6_WEIMER outputs generated for V.20130831 reports (dB/dt and Regional-K) reflect the recent corrections that were found necessary in running the Weimer model: 1) the driver routine now properly feeds in By from the solar wind; 2) applies a 17-minute delay to the solar wind input; 3) rotates model output into magnetic dipole coordinates.

12.5. Weigel

Initial model release: February 13, 2011. We had multiple problems and questions when executing the model and a few more re-releases were necessary to address issues with different versions of Octave (used as substitute for Matlab). It was also realized by model developer that the originally released version was designed for the delta-B studies

and is not appropriate for dB/dt calculations. A final model version designed for dB/dt predictions was released on March 15, 2012. Therefore different model versions are used for dB/dt study included in Phase I report and for delta-B and local K predictions (Phase II report). Final reruns were last done in October 2012.

The approach to solar wind propagation was revisited in August 2013. it had been assumed that solar wind data should be provided at the Earth's position, while the Weigel model would have required the solar wind at the bow shock location (about a 3-minute difference from the data propagated to Earth that were used) but no additional delay. In a conference call with Robert Weigel on August 20, it was concluded that the difference was insignificant in the dB/dt study with its 20-minute analysis windows and will be negligible in the regional-K study that uses 3-hour windows.

13. Discussion

In this work, coordination among the CCMC, NOAA SWPC, modelers and science community has resulted in the evaluation of several geospace models capable of predicting the fluctuation of the ground magnetic field. The work was a continuation of earlier GEM modeling challenges and was designed to support model transition into operations at NOAA SWPC. The primary NOAA interest in this specific effort was to study models' capability to reproduce the observed dB/dt, which can be used as an indicator for GIC activity.

The report describes the method developed at the CCMC by *Rastätter et al.* [2013b] to calculate the magnetic perturbation delta-B at selected magnetometer station locations from physics-based coupled models of the magnetosphere-ionosphere system. The calcu-

lations are performed on snapshots of the full current systems in the magnetosphere and ionosphere that are written at regular time intervals during a simulation run.

To validate the delta-B tool we compared components of the new post-processing algorithm to the existing implementation of a $\Delta\mathbf{B}$ calculation within the SWMF magnetosphere-ionosphere model. Both the SWMF model and our post-processing tool essentially generate the same $\Delta\mathbf{B}$ results that include contributions from all the currents in the coupled magnetosphere-ionosphere system. We identified minor differences between the SWMF model results and results from our algorithm that can be attributed to the different methods used to convert between the coordinate systems that are being used. SWMF uses a fixed geographic position of the magnetic dipole to convert between Magnetic (MAG), Solar-Magnetic (SM) and Geocentric-Solar-Magnetospheric (GSM) coordinates. We use Geopack-2008 that accounts for secular variations in the orientation of the geo-magnetic dipole. The difference in position is of the order of one degree in latitude or longitude and explains the differences seen in the magnetic perturbation signals.

We were able to perform the post-processing of one-minute-resolution model outputs in real-time on a few processors. This means that it takes less than a minute per set of outputs to read the currents, calculate and write the magnetic perturbations for every station in the list that was used in the study. The model grid in the ionosphere was used to set up a spherical grid between the ionospheric altitude and the current pickup radius in the magnetosphere. We performed a grid convergence study and found that good results can be achieved even when using only about a tenth of the altitude layers compared to the SWMF model. Differences between the SWMF runtime calculations and

the CCMC delta-B post-processing tool are negligible compared to the differences seen between observations and modeled results.

We described the functionality of the online model-observation comparison tool that has been implemented to compare time lines in support of this evaluation as well as other model validation studies performed at the CCMC.

We reported here the metrics results integrated over selected stations and all events. Model 9_SWMF provided the highest POD and HSS for all dB/dt thresholds used to build the event detection-based contingency tables. However, we emphasize that the metrics results vary from station to station and event to event. One should thus be cautious in making general interpretations without studying the more detailed breakdown of the analysis. For this purpose, along with CCMC's online analysis interface, contingency tables for each individual event are available at the "Operational Geospace Model Validation" pages at the CCMC website and at the appendixes in the Supplementary Material attached to this report.

Finally, the key question is "are the models good enough to provide tangible value for the end-user in need to mitigate GIC?" This is a multifaceted complex question and the answer most likely varies from user to user. Based on the summary results for POD, POFD and HSS with the dB/dt threshold of 1.5 nT/s, it is clear that predicting large dB/dt is still a challenge. POD and HSS were below 0.5 for all models for the dB/dt threshold of 1.5 nT/s. Users requiring localized predictions for large dB/dt with high likelihood of event detection may not be satisfied with the current state-of-the-art. However, we saw that models have the capability to capture the general level of enhanced activity.

Consequently, users satisfied with more rough characterization of dB/dt activation over the storm periods may be able to use the models for generating actionable information.

The models validated in this paper can provide short lead-time dB/dt predictions. The meaning of "short" will vary as a function of the speed of transient structures in the solar wind and the computational capacity available for model execution. Lead-times of 15-30 minutes at best can be expected for fast coronal mass ejection events. Obviously, continuous high-quality upstream solar wind plasma and magnetic field monitoring used to drive the models is also required. It is important to acknowledge that while providing 24/7 data stream, ACE SWEPAM plasma experiment has limitations during strong solar energetic particle events [*Skoug et al.*, 2004] often associated with major Earth-directed coronal mass ejections. The Deep Space Climate Observatory (DSCOVR) mission that will replace ACE as the primary upstream monitor is expected to launch 2014.

Finally, one of the results of this effort to evaluate geospace models for transition from the research environment to operations is that it has accelerated the delivery of new versions of models to the CCMC for use by the science community. It has also resulted in the rigorous validation of models and initiated feedback from the operations to research that will ultimately result in a better understanding of where model improvements are most needed.

14. Supplementary Comments

The analysis presented in this report demonstrated that for the selected physical parameter (dB/dt) and threshold-based metrics physics-based models are ranked higher than empirical models for most of events and thresholds. The top ranked model contains additional physics (ring current effects) that are not included in other models. This is an

indication that incorporation of inner magnetosphere into global magnetosphere models is important for further improvements of model performance. Comparative analysis of LFM simulations with different settings for the ionosphere conductance model reveled the sensitivity of the results to the model for the Hall conductance. Therefore development of realistic ionosphere conductance model is an important element of further model improvement.

One should be cautious in making general interpretations of the results of metrics studies presented in this report. Our experience with different types of metrics studies demonstrated that model ranking depends on physical parameters selected for evaluation as well as on metrics type. For different physical parameters and different approaches to model-data comparison the top-ranking model can be different. By no means model ranking presented in this report can be interpreted as ranking of overall model quality.

It is clear that predicting large dB/dt is still a challenge. POD and HSS were below 0.5 for all models for the dB/dt threshold of 1.5 nT/s. This is an indication that model performance may benefit from grid resolution increase at the inner boundary of the simulation domain and in areas of currents systems that are mapped to the ionosphere.

Simulations for this evaluation study used level 2 solar wind input data for most of the events. Real-time input data may degrade the performance and change model ranking. For real-time calculations input data streams have to be automatically analyzed for data gaps, bad data points, etc.

Time series analyzed in this report were obtained using CCMC post-processing tool that includes contribution from different sources in ionosphere and magnetosphere. Thus,

we are dealing with value added products. Using other tools may affect metrics analysis results.

Runs performed for this study demonstrated model robustness for a set of events that lasted no more than 48 hours. Continues simulations for longer period of time impose additional requirements on model robustness. All of the models evaluated in this study are implemented at the CCMC Runs-on-request system. Continues real-time simulations are also being tested. In some cases we have to modify model settings (e.g., by changing the simulation grid and/or by switching to more diffusive but more robust numerical scheme) to allow the run to come through the requested time interval. Modification of model settings can lead to changes in skill scores and model ranking. Continues use of models at CCMC for Runs-on-request and real-time runs will increase the statistics for further model performance evaluation. Timeline analysis tools developed at the CCMC in support of this study allow real-time post-processing and skill score calculations in automatic fashion.

During the 2-years model evaluation process there were significant upgraded to almost all models participated in this study. Periodic reevaluations of model performance and progress over time using metrics, post-processing tools and CCMC on-line automated validation systems developed during this study is recommended.

References

Baker, D. N., R. S. Weigel, E. J. Rigler, R. L. McPherron, D. Vassiliadis, C. N. Arge, G. L. Siscoe, and H. E. Spence (2004), Sun-to-magnetosphere modeling: CISM forecast model development using linked empirical methods, *Journal of Atmospheric and Solar-*

Terrestrial Physics, 66, 1491, doi:10.1016/j.jastp.2004.04.011.

Birn, J., et al. (2001), Geospace Environmental Modeling (GEM) Magnetic Reconnection Challenge, *Journal of Geophysical Research*, *106*(A3), 3715–3719, doi: 10.1029/1999JA900449.

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.

Buzulukova, N., M.-C. Fok, A. Pulkkinen, M. Kuznetsova, T. E. Moore, A. Glocer, P. C. Brandt, G. Toth, and L. Rastatter (2010), Dynamics of ring current and electric fields in the inner magnetosphere during disturbed periods: CRCM-BATS-R-US coupled model, *J. Geophys. Res.*, 115:A05210.

Connor, H. J., J. Raeder, and K. H. Trattner (2012), Dynamic modeling of cusp ion structures, *J. Geophys. Res.*, 117, A04203, doi:10.1029/2011JA017203.

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 magnetosphere model: Initial results, *J. Geophys. Res.*, 109(A12):A12219.

Dorelli, J. C., M. Hesse, M. M. Kuznetsova, L. Rastaetter, and J. Raeder (2004), A new look at driven magnetic reconnection at the terrestrial subsolar magnetopause, *J. Geophys. Res.*, 109 , A12,216, doi: 10.1029/2004JA010,458.

Emmert, J. T., A. D. Richmond, and D. P. Drob (2010), A computationally compact representation of MagneticApex and QuasiDipole coordinates with smooth base vectors, *J. Geophys. Res.*, 115, A08322, doi:10.1029/2010JA015326.

Finlay, C. C., et al. (2010), International Geomagnetic Reference Field: the eleventh generation, *Geophysical Journal International*, *183*, 1216–1230, doi:10.1111/j.1365-

246X.2010.04804.x.

Fukushima, N. (1976), Generalized theorem for no ground magnetic effect of vertical currents connected with Pedersen currents in the uniform-conductivity ionosphere, *Report of Ionososphere and Space Research in Japan,*, *30*, 35–40.

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 *STEP Report*, edited by R. W. Schunk, p. 217, Scientific Committee on Solar Terrestrial Physics (SCOSTEP), NOAA/NGDC, Boulder, Colorado.

Garcia, K. S., and W. J. Hughes (2007), Finding the Lyon-Fedder-Mobarry magnetopause: A statistical perspective, *J. Geophys. Res*, 112, 06229, doi:10.1029/2006JA012039.

Ge, Y. S., J. Raeder, V. Angelopoulos, M. Gilson, and A. Runov (2011), Interaction of dipolarization fronts within multiple bursty bulk flows in global MHD simulations of a substorm on 27 February 2009, *J. Geophys. Res.*, 116, A00I23, doi:10.1029/2010JA015758.

Gilson, M. L., J. Raeder, E. Donovan, Y. S. Ge, and E. L. Kepko (2012), Global simulation of proton precipitation due to field line curvature during substorms, *J. Geophys. Res.*, 117, A05,216.

Glocer, A., G. Tóth, M. Fok, and T. Gombosi (2009a), Integration of the radiation belt environment model into the space weather modeling framework, *J. Atmos. Solar-Terr. Phys.*, 71:1653–1663.

Glocer, A., G. Tóth, T. I. Gombosi, and D. Welling (2009b), Polar Wind Outflow Model (PWOM): Modeling ionospheric outflows and their effect on the magnetopshere, inital results, *J. Geophys. Res.*, 114:A05216.

Glocer, A., G. Tóth, Y. J. Ma, T. Gombosi, J.-C. Zhang, and L. M. Kistler (2009c), Multi-fluid BATS-R-US: Magnetospheric composition and dynamics during geomagnetic storms, initial results, *J. Geophys. Res.*, 114, A12203, doi:10.1029/2009JA014418.

Gombosi, T.I., G. Tth, D.L. De Zeeuw, K.C. Hansen, K. Kabin, and K. G. Powell (2002), Semi-relativistic magnetohydrodynamics and physics-based convergence acceleration, *J. Computational Phys.*, 177, 176-205.

Gombosi, T.I., K. G. Powell, D. L. De Zeeuw, C. R. Clauer, K. C. Hansen, W. B. Manchester, A. J. Ridley, I. I. Roussev, I. V. Sokolov, Q. F. Stout, and G. Tóth (2004), Solution-adaptive magnetohydrodynamics for space plasmas: Sun-to-Earth simulations, *Computing in Science and Engineering*, 06(2):14–35, Mar/Apr.

Goodrich, C. C., J. G. Lyon, M. Wiltberger, R. E. Lopez, and K. Papadopoulos (1998), An overview of the impact of the January 10-11, 1997 magnetic cloud on the magnetosphere via global MHD simulation, *Geophys. Res. Lett.*, 25(1), 25372540, doi:10.1029/98GL01159.

Guild, T. B., H. E. Spence, E. L. Kepko, V. Merkin, J. G. Lyon, M. Wiltberger, and C. C. Goodrich (2008a), Geotail and LFM comparisons of plasma sheet climatology: 1. Average values, *J. Geophys. Res*, 113, 04216, doi:10.1029/2007JA012611.

Guild, T. B., H. E. Spence, E. L. Kepko, V. Merkin, J. G. Lyon, M. Wiltberger, and C. C. Goodrich (2008b), Geotail and LFM comparisons of plasma sheet climatology: 2. Flow variability, *J. Geophys. Res*, 113, 04217, doi:10.1029/2007JA012613.

Haines, G. V. (1985), Spherical cap harmonic analysis, *J. Geophys. Res.*, 90 (B3), 2583.

Hapgood, M.A. (1992), Space Physics Coordinate Transformations: A User Guide, *Planet. Space Sci.*, Vol. 40, No. 5, 711-717.

Kennel, C. F., and H. E. Petschek (1966), Limit on stably trapped particle fluxes, *J. Geophys. Res.*, 71, 1.

Knight, S. (1972), Parallel electric fields, *Planet. Space Sci.*, 21, 741.

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

Li, W., J. Raeder, J. Dorelli, M. ieroset, and T. D. Phan (2005), Plasma sheet formation during long period of northward IMF, *Geophys. Res. Lett.*, 32, L12S08, doi: 10.1029/2004GL021,524.

Li, W., J. Raeder, J. C. Dorelli, M. Thomsen, and B. Lavraud (2008), Solar wind entry into the magnetosphere under northward IMF conditions, *J. Geophys. Res.*, 113, A04, 204.

Li, W., J. Raeder, M. Oieroset, and T. D. Phan (2009), Cold dense magnetopause boundary layer under northward IMF: Results from THEMIS and MHD simulations, *J. Geophys. Res.*, 114, A00C15, doi:10.1029/2008JA013, 497.

Li, W., J. Raeder, and D. Knipp (2011), The relationship between dayside local poynting flux enhancement and cusp reconnection, J. Geophys. Res., 116, A08, 301, doi:10.1029/2011JA016,566.

Lopez, R. E., C. C. Goodrich, M. Wiltberger, K. Papadopoulos, and J. G. Lyon (1998), Simulation of the March 9, 1995 substorm and initial comparison to data, *Geospace Mass and Energy Flow: Results From the International Solar-Terrestrial Physics Program*, 104, 237-245, doi:10.1029/GM104p0237.

Lopez, R. E., S. Hernandez, M. Wiltberger, J. Lyon, and C. Goodrich (2006), Initial Results from the Simulation of the Halloween 2003 Storms, *Advances in Geosciences*, 2, 191.

Lopez , R. E., S. Hernandez, M. Wiltberger, C.-L. Huang, E. L. Kepko, H. Spence, C. C. Goodrich, and J. G. Lyon (2007), Predicting magnetopause crossings at geosynchronous orbit during the Halloween storms, *Space Weather*, 5, S01005, doi:10.1029/2006SW000222.

Lu, G., et al. (2011), Reversed two-cell convection in the northern and southern hemi-spheres during northward interplanetary magnetic field, *J. Geophys. Res.*, 116, A12,237.

Lyon, J. G., J. A. Fedder, and C. M. Mobarry (2004), The Lyon-Fedder-Mobarry (LFM) global MHD magnetospheric simulation code, *Journal of Atmospheric and Solar-Terrestrial Physics*, 66, 1333, doi:10.1016/j.jastp.2004.03.020.

Lyons, L. R., D. Evans, and R. Lundin (1979), An observed relation between magnetic field aligned electric fields and downward electron energy fluxes in the vicinity of auroral forms, *J. Geophys. Res.*, 84, 457.

Matsumoto, H., et al. (1994), Plasma wave observations with geotail spacecraft, *J. Geomag. Geoelectr.*, *46*, 59–95.

Merkin, V. G., and J. G. Lyon (2010), Effects of the low-latitude ionospheric boundary condition on the global magnetosphere, *J. Geophys. Res*, 115(A), 10202, doi:10.1029/2010JA015461.

Muhlbacher, S., C. J. Farrugia, J. Raeder, H. K. Biernat, and R. B. Torbert (2005), A statistical investigation of dayside erosion showing saturation response, *J. Geophys. Res.*, 110, A11, 207, doi:10.1029/2005JA011,177.

National Research Council (2008), Severe Space Weather Events-Understanding Societal and Economic Impacts: A Workshop Report, *The National Academies Press*, Washington, DC.

North American Electric Reliability Corporation GMD Task Force (2012), 2012 Special Reliability Assessment Interim Report: Effects of Geomagnetic Disturbances on the Bulk Power System, *NERC*, February 2012.

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, *J. Geophys. Res.*, 117(A2), doi:10.1029/2011JA016979.

Pirjola, R., Effects of space weather on high-latitude ground systems, *Advances in Space Research*, 36, 2231-2240, 2005.

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. Comp. Phys.*, 154:284–309.

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*, Vol. 4, S08005, doi:10.1029/2006SW000229.

Pulkkinen, A., M. Hesse, S. Habib, L. Van der Zel, B. Damsky, F. Policelli, D. Fugate (2009), and W. Jacobs, Solar Shield: forecasting and mitigating space weather effects on high-voltage power transmission systems, *Natural Hazards*, doi:10.1007/s11069-009-

9432-x.

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*, 115, A03206, doi:10.1029/2009JA014537.

Pulkkinen, A., M. Kuznetsova, A. Ridley, J. Raeder, A. Vapirev, D. Weimer, R. S. Weigel, M. Wiltberger, G. Millward, L. Rastätter, M. Hesse, H. J. Singer and A. Chulaki (2011), Geospace Environment Modeling 2008-2009 Challenge: ground magnetic field perturbations, *Space Weather*, Vol. 9, S02004, doi:10.1029/2010SW000600.

Pulkkinen, A., M. M. Kuznetsova, L. Rastätter, and H. Singer (2013), Community-wide validation of ground magnetic field perturbation predictions of geospace models to support model transition to operations, *Space Weather*, Vol. 11, doi:10.1002/swe.20056.

Raeder, J. (2003), Global Magnetohydrodynamics A Tutorial, in *Space Plasma Simulation*, edited by J. Buchner, C. T. Dum, and M. Scholer, Springer Verlag, Berlin Heidelberg New York.

Raeder, J. (2006), Flux transfer events: 1. Generation mechanism for strong southward IMF, *Ann. Geophys.*, 24, 381.

Raeder, J., J. Berchem, and M. Ashour-Abdalla (1996), The importance of small scale processes in global MHD simulations: Some numerical experiments, in *The Physics of Space Plasmas*, edited by T. Chang and J. R. Jasperse, vol. 14, p. 403, MIT Cent. for Theoret. Geo/Cosmo Plasma Phys., Cambridge, Mass.

Raeder, J., J. Berchem, and M. Ashour-Abdalla (1998), The Geospace Environment Modeling grand challenge: Results from a Global Geospace Circulation Model, *J. Geophys.*

Res., 103, 14,787.

Raeder, J., Y. L. Wang, and T. Fuller-Rowell (2001a), Geomagnetic storm simulation with a coupled magnetosphere - ionosphere - thermosphere model, in *Space Weather, AGU Geophys. Monogr. Ser.*, edited by P. Song, G. Siscoe, and H. J. Singer, vol. 125, p. 377, American Geophysical Union.

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

Raeder, J., D. Larson, W. Li, E. L. Kepko, and T. Fuller-Rowell (2008), OpenGGCM simulations for the THEMIS mission, *Space Sci. Rev.*, 141, 535, doi:10.1007/s11,21400894215.

Raeder, J., P. Zhu, Y. Ge, and G. L. Siscoe (2010), OpenGGCM simulation of a substorm: Axial tail instability and ballooning mode preceding substorm onset, *J. Geophys. Res.*, 115 , A00l16.

Raeder, J., et al. (2001c), Global simulation of the geospace environment modeling substorm challenge event, *J. Geophys. Res.*, 106, 381.

Rastätter, L., 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, Vol. 9, S04005, doi:10.1029/2010SW000617.

Rastätter, L., et al. (2013a), Geospace Environment Modeling 2008-2009 Challenge: D_{st} index, *Space Weather*, *11*, 20036, doi:10.1002/swe.20036.

Rastätter, L., M. Kuznetsova, A. Pulkkinen, G. Toth (2013b), Community-wide geospace model validation: computation of ΔB at the CCMC, submitted to *Space Weather*, 2013.

Richmond, A. D. (1995), Ionospheric electrodynamics using magnetic apex coordinates, *J. Geomag. Geoelectr.*, 47, 191.

Ridley, A. J. , K. C. Hansen, G. Tóth, D. L. De Zueew, T. I. Gombosi, and K. G. Powell (2002), University of Michigan MHD results of the GGCM metrics challenge, *J. Geophys. Res.*, 107(A10):1290.

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

Robinson, R. M., R. R. Vondrak, K. Miller, T. Dabbs, and D. Hardy (1987), On calculating ionospheric conductances from the flux and energy of precipitating electrons, *J. Geophys. Res.*, 92, 2565.

Roble, R. G., and E. C. Ridley (1994), A Thermosphere - Ionosphere - mesosphere - electrodynamics General Circulation Model (TIME-GCM): Equinox solar minimum simulations (30-500km), *Geophys. Res. Lett.*, 21, 417420.

Skoug, R. M., Gosling, J. T., Steinberg, J. T., McComas, D. J., Smith, C. W., Ness, N. F., Hu, Q., Burlaga, L. F., (2004), Extremely high speed solar wind: 2930 October 2003, *J. Geophys. Res.*, 109, A09102, doi:10.1029/2004JA010494.

Siscoe, G., J. Raeder, and A. Ridley (2004), Transpolar potential saturation models compared, *J. Geophys. Res.*, 109, A09,203, doi:10.1029/2003JA010,318.

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

Sonnerup, B. U. O., and L. J. Cahill (1967), Magnetopause structure and attitude from Explorer 12 observations, *J. Geophys. Res.*, 72, 171.

Sonnerup, B. U. O., and L. J. Cahill (1968), Explorer 12 observations of the magnetopause current layer, *J. Geophys. Res.*, 73, 1757.

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), 13119-13134.

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.R. Clauer, D. L. De Zeeuw, K. C. Hansen, K. J. Kane, W. B. Manchester, K. G. Powell, A. J. Ridley, I. I. Roussev, Q. F. Stout, O. Volberg, Richard A. Wolf, S. Sazykin, A. Chan, Bin Yu, and József Kóta (2005), Space Weather Modeling Framework: A new tool for the space science community, *J. Geophys. Res.*, 110:A12226.

Tóth, G., D. L. De Zeeuw, T. I. Gombosi, and K. G. Powell (2006), A parallel explicit/implicit time stepping scheme on block-adaptive grids, *J. Comput. Phys.*, 217:722–758.

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 October 28-30, 2003 storm with the Space Weather Modeling Framework, *Space Weather Journal*, 5:S06003.

Tóth, G. et al (2012), Adaptive numerical algorithms in space weather modeling, *J. Comput. Phys.*, 231.

VanZandt, T. E., W. L. Clark, and J. M. Warnock (1972), Magnetic apex coordinates: A magnetic coordinate system for the ionospheric f2 layer, *J. Geophys. Res.*, 77, 2406.

Vennerstrom, S., T. Moretto, L. Rastaetter, and J. Raeder (2005), Field-aligned currents during northward interplanetary field: Morphology and causes, *J. Geophys. Res.*, 110, A06,205, doi:10.1029/2004JA010,802.

Vennerstrom, S., T. Moretto, L. Rastaetter, and J. Raeder (2006), Modeling and analysis of solar wind generated contributions to the near-Earth magnetic field, *Earth, Planets, Space*, 58, 451-461.

Viljanen, A., H. Nevanlinna, K. Pajunpää, and A. Pulkkinen (2001), Time derivative of the horizontal geomagnetic field as an activity indicator, *Ann. Geophys.*, 19, 1107-1118.

Wang, W., M. Wiltberger, A. G. Burns, S. C. Solomon, T. L. Killeen, N. Maruyama, and J. G. Lyon (2004), Initial results from the coupled magnetosphere-ionosphere-thermosphere model: thermosphere-ionosphere responses, *Journal of Atmospheric and Solar-Terrestrial Physics*, 66(1), 1425-1441, doi:10.1016/j.jastp.2004.04.008.

Weigel, R. S., A. J. Klimas, D. Vassiliadis (2003), Solar wind coupling to and predictability of ground magnetic fields and their time derivatives, *J. Geophys. Res.*, 108 (A7), 1298, doi:10.1029/2002JA009627.

Weigel, R. S. (2007), Solar wind time history contribution to the day-of-year variation in geomagnetic activity, *J. Geophys. Res.*, 112, A10207, doi:10.1029/2007JA012324.

Weimer, D. R., D. A. Gurnett, C. K. Goertz, J. D. Menietti, J. L. Burch, and M. Sugiura (1987), The current voltage relationship in auroral current sheets, *J. Geophys. Res.*, 92, 187.

Weimer, D. R., and J. H. King (2008), Improved calculations of interplanetary magnetic field phase front angles and propagation time delays, *J. Geophys. Res.*, 113, A01105, doi:10.1029/2007JA012452.

Weimer, D. R., C. R. Clauer, M. J. Engebretson, T. L. Hansen, H. Gleisner, I. Mann, and K. Yumoto (2010), Statistical maps of geomagnetic variations as a function of the interplanetary magnetic field, *J. Geophys. Res.*, 115, A10320, doi:10.1029/2010JA015540.

Weimer, D. R. (2013), An empirical model of ground-level geomagnetic perturbations, *Space Weather*, in press, doi: 2012SW000880.

Welling, D. T., and A. J. Ridley (2010), Validation of swmf magnetic field and plasma, *Space Weather*, 8.

Wintoft, P. (2005), Study of the solar wind coupling to the time difference horizontal geomagnetic field, *Annales Geophysicae*, 23, 1949-1957.

Wiltberger, M., T. I. Pulkkinen, J. G. Lyon, and C. C. Goodrich (2000), MHD simulation of the magnetotail during the December 10, 1996, substorm, *Journal of Geophysical Research*, 105(A), 2764927664, doi:10.1029/1999JA000251.

Wiltberger, M., J. G. Lyon, and C. C. Goodrich (2003), Results from the Lyon-Fedder-Mobarry global magnetospheric model for the electrojet challenge, *Journal of Atmospheric and Solar-Terrestrial Physics*, 65(1), 1213-1222, doi:10.1016/j.jastp.2003.08.003.

Wiltberger, M., W. Wang, A. G. Burns, S. C. Solomon, J. G. Lyon, and C. C. Goodrich (2004), Initial results from the coupled magnetosphere ionosphere thermosphere model: magnetospheric and ionospheric responses, *Journal of Atmospheric and Solar-Terrestrial Physics*, 66(1), 1411-1423, doi:10.1016/j.jastp.2004.03.026.

Wiltberger, M., R. S. Weigel, W. Lotko, and J. A. Fedder (2009), Modeling seasonal variations of auroral particle precipitation in a global-scale magnetosphere-ionosphere simulation, *Journal of Geophysical Research*, 114(A), 01204, doi:10.1029/2008JA013108.

Wiltberger, M., L. Qian, C.-L. Huang, W. Wang, R. E. Lopez, A. G. Burns, S. C. Solomon, Y. Deng, and Y. Huang (2012), CMIT study of CR2060 and 2068 comparing L1 and MAS solar wind drivers, *Journal of Atmospheric and Solar-Terrestrial Physics*, doi:10.1016/j.jastp.2012.01.005.

Wolf, R.A., M. Harel, R. W. Spiro, G.-H. 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.

Zaharia, S., V. K. Jordanova, D. Welling, and G. Toth (2010), Self-consistent inner magnetosphere simulation driven by a global mhd model, *J. Geophys. Res.*, 115:A12228.

Zhang, J., M.W. Liemohn, D.L. De Zeeuw, J.E. Borovsky, A.J. Ridley, G. Tóth, 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, *J. Geophys. Res.*, 112:A04208.

Zhang, J. J., C. Wang, and B. B. Tang (2012), Modeling geomagnetically induced electric field and currents by combining a global MHD model with a local one-dimensional method, *Space Weather*, 10, S05005, doi:10.1029/2012SW000772.

Yu, Y., and A.J. Ridley (2008), Validation of the Space Weather Modeling Framework using ground-based magnetometers, *Space Weather*, 6.

Yu, Y., A. 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*, *115*, A08207, doi: 10.1029/2009JA014869.
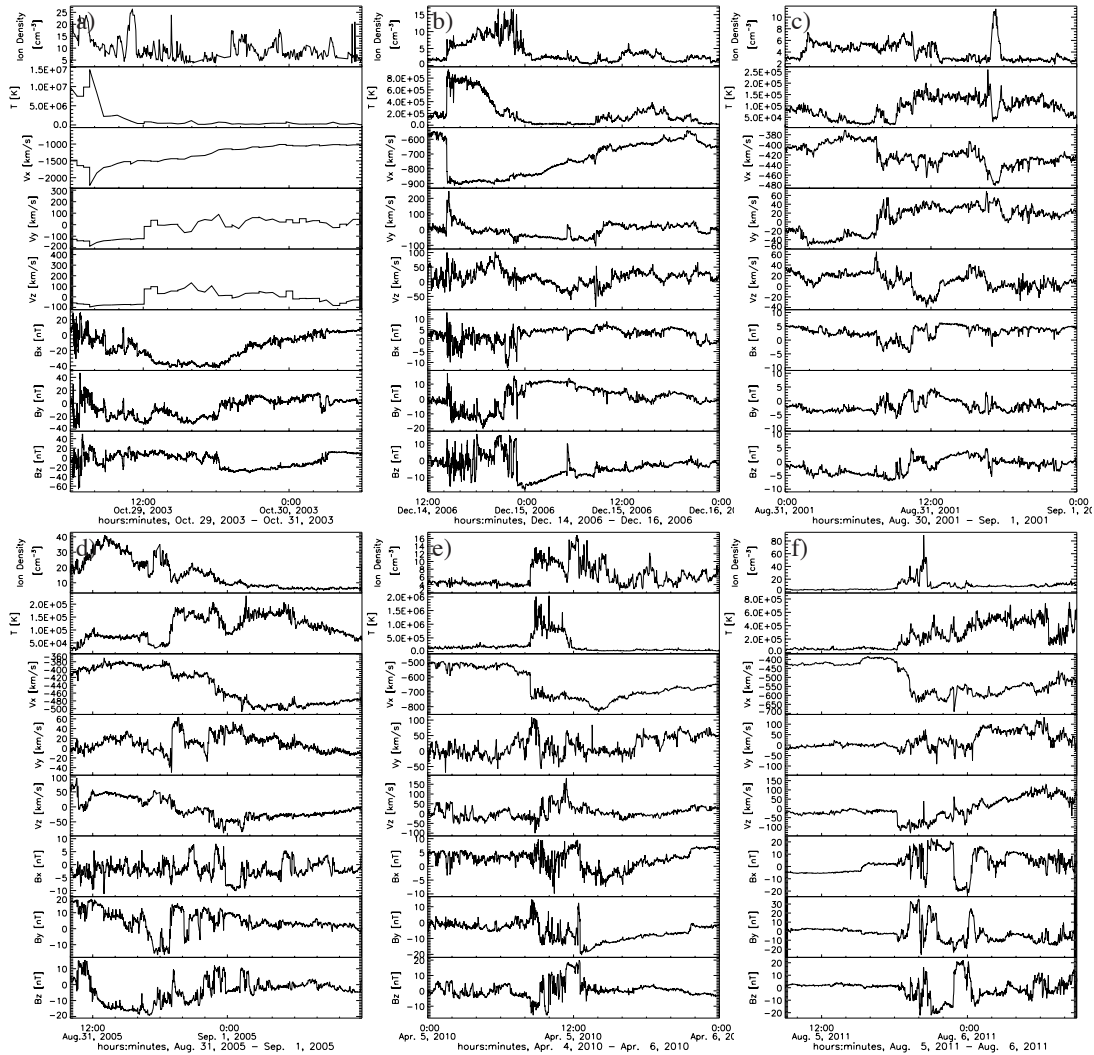
Figure 1.   Solar wind bulk plasma and the interplanetary magnetic field observations (in each panel from top to bottom: plasma density, plasma temperature, x-component of the plasma flow velocity, y-component of the plasma flow velocity, z-component of the plasma flow velocity, x-component of the magnetic field, y-component of the magnetic field, z-component of the magnetic field) for the studied storm events (panels a-f corresponding to events 1-6) given in Table 1.   Data in Geocentric Solar Magnetospheric coordinates.. See the text for details.

Table 1. Geospace events studied in the validation activity. The last two columns give the minimum Dst index and the maximum Kp index of the event, respectively.

Event #	Date and time	min(Dst)	max(Kp)
1	October 29, 2003 06:00 UT - October 30, 06:00 UT	-353 nT	9
2	December 14, 2006 12:00 UT - December 16, 00:00 UT	-139 nT	8
3	August 31, 2001 00:00 UT - September 1, 00:00 UT	-40 nT	4
4	August 31, 2005 10:00 UT - September 1, 12:00 UT	-131 nT	7
5	April 5, 2010 00:00 UT - April 6, 00:00 UT	-73 nT	8-
6	August 5, 2011 09:00 UT - Aug 6, 09:00 UT	-113 nT	8-

Table 2. The locations of the geomagnetic observatories used in the study. Bold typeface stations indicate the six stations (stations PBQ and SNK are alternates, see the text for details) used in the final analyses. The table lists the stations by chain from low to high magnetic latitude.

Station		Geographic		Geo-Magnetic	
Name	IAGA Code	lat.	lon.	lat.	lon.
Fresno	FRN	37.09	240.28	43.52	305.25
Newport	NEW	48.27	242.88	54.85	304.68
Meanook	**MEA**	54.62	246.65	61.57	306.20
Yellowknife	**YKC**	62.48	245.52	68.93	299.36
Fredericksburg	FRD	38.20	282.63	48.4	353.38
Ottawa	**OTT**	45.40	284.45	55.63	355.31
Poste de la Baleine	**PBQ**	55.28	282.26	65.46	351.81
Iqaluit	IQA	63.75	291.48	73.98	5.24
Fürstenfeldbrück	FUR	48.17	11.28	48.38	94.61
Wingst	**WNG**	53.74	9.07	54.12	95.0
Abisko	**ABK**	68.36	18.82	66.06	114.66
Hornsund	HRN	77.00	15.37	73.88	125.99
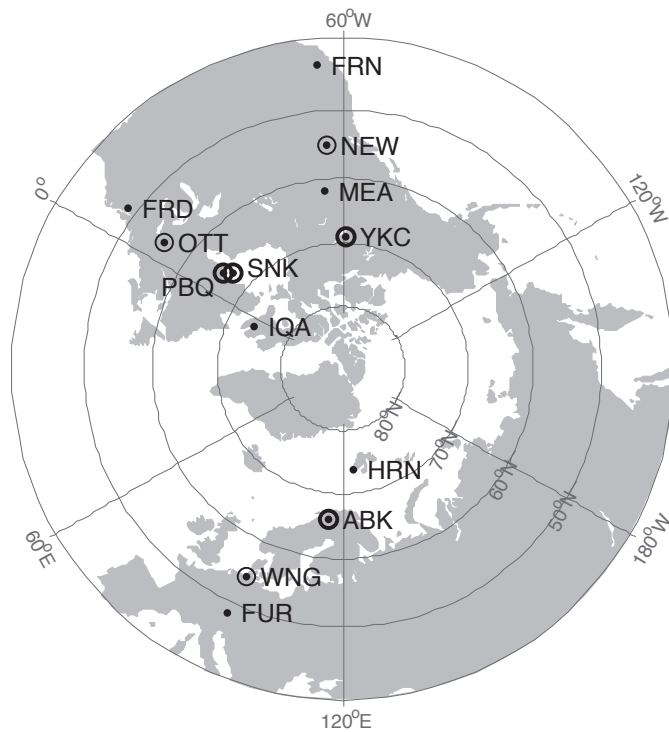
Figure 2.   The locations and the station codes of the geomagnetic observatories used in the study. Geomagnetic dipole coordinates are used. Thick and thin circles indicate high-latitude and mid-latitude stations, respectively, used in the final analyses in Section 10.

Table 3. Models analyzed in the validation effort. Each model is assigned a unique model identifier given by the leftmost column of the table. The table indicates the model setting, and if applicable, the number of cells and the minimum spatial resolution used in the global MHD part of the model. See text in Section 5 for details.

Identifier	(model version) Model	Grid (# of cells, min. re
2_LFM-MIX	(LTR-2.1.1) LFM coupled with ionospheric electrodynamics	163,000, 0.4 R_E
3_WEIGEL	empirical model	N/A
4_OPENGGCM	(OpenGGCM 4.0) global MHD coupled with CTIM	3.9 million, 0.25 R_E
6_WEIMER	empirical model	N/A
9_SWMF	(SWMF 2011-01-31) BATS-R-US coupled with RIM and RCM	1 million, 0.25 R_E
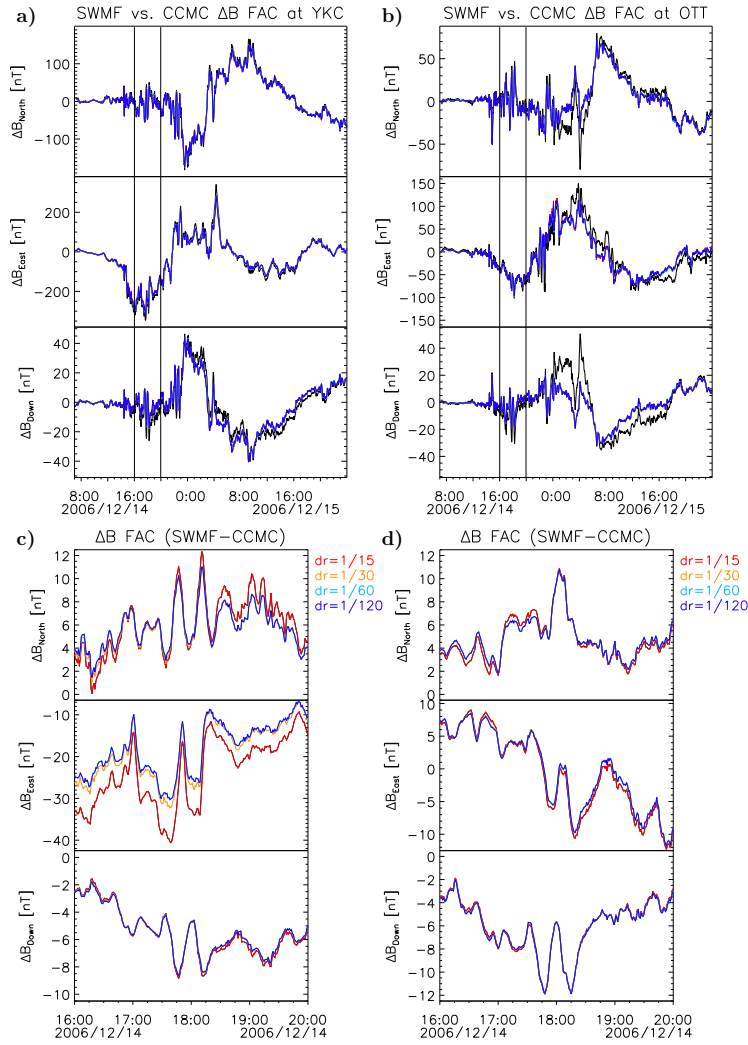
Figure 3.   Grid convergence study for field-aligned current contribution to $\Delta\mathbf{B}$. Figure 3a shows results for the high-latitude station YKC, Figure 3b shows results for the mid-latitude station OTT. The black trace shows the SWMF result and the colored traces show the different grid resolutions used by the CCMC calculation: red: $dr = 1/15R_E$, orange: $dr = 1/30R_E$, blue: $dr = 1/60R_E$. dark blue: $dr = 1/120R_E$ At the global scale the traces are virtually identical. Figure 3c and Figure 3d show in greater detail the colored traces between 16:00 and 20:00 on 2006/12/14, indicated by vertical lines in Figure 3a) and Figure 3b. The middle traces in Figure 3c for YKC show the largest differences between the resolution settings. SWMF signal is not shown in the shorter time intervals in Figure 3c and Figure 3d.
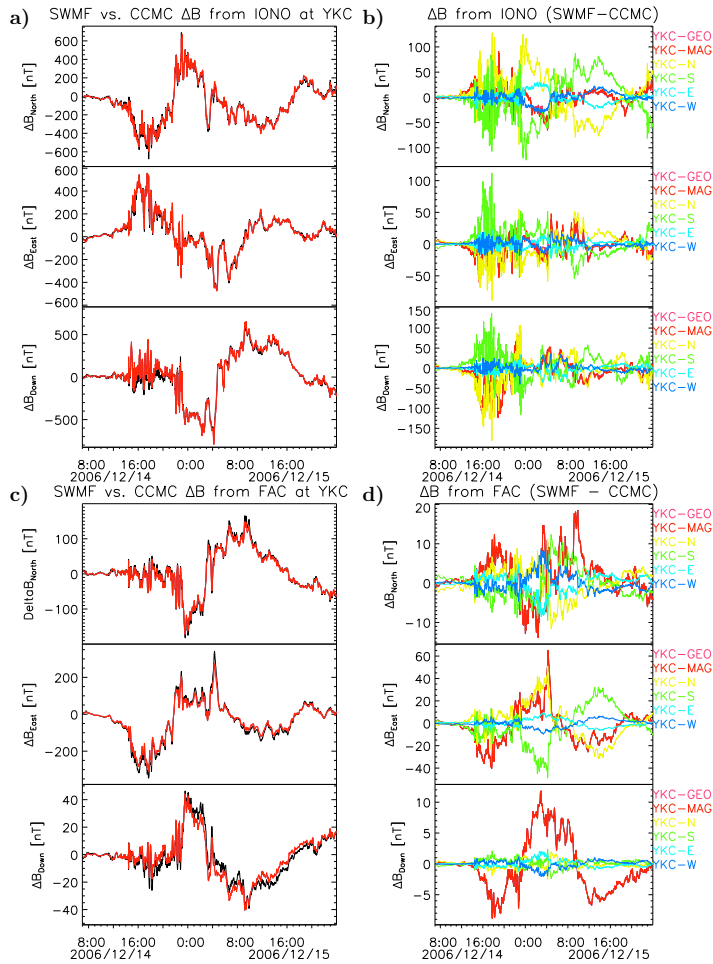
Figure 4. Effect of station location on ionosphere current and FAC contributions to $\Delta \mathbf{B}$ at YKC. Figure 4a: Ionosphere current contribution: SWMF (black trace), CCMC post-processing using either station locations in geographic (magenta) or magnetic coordinates (red). The two CCMC traces are indistinguishable on this scale and differ slightly from the SWMF trace. Figure 4b: Differences between SWMF and CCMC (magenta and red, nearly on top of each other) and the difference between the CCMC signal at the station location and 4 locations one degree away in geographic coordinates in each direction (north: "YKC-N" [yellow], south: "YKC-S" [green], east: "YKC-E" [light blue], west: "YKC-W" [dark blue]). Figure 4c: $\Delta \mathbf{B}$ from FAC in same format as Figure 4a. Figure 4d: Differences between the SWMF and CCMC for FAC contribution to $\Delta \mathbf{B}$. Colored traces are derived in the same manner as in Figure 4b.
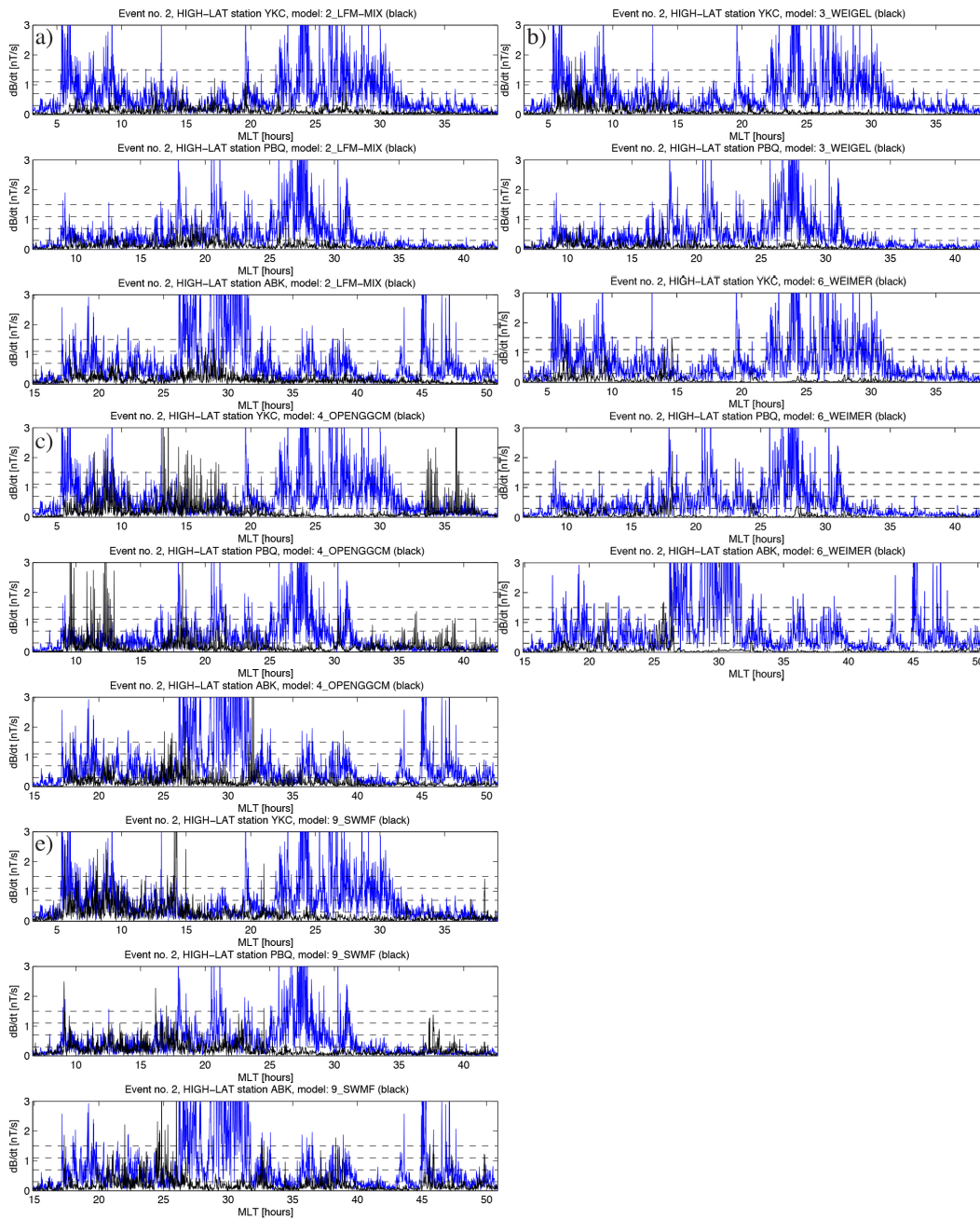
Figure 5. Time series of the observed (blue curves) vs modeled (black curves) dB/dt at the three high-latitude stations indicated in Table 2 for event 2 indicated in Table 1. The time is magnetic local time (MLT) and the dashed lines indicate the dB/dt thresholds of .3, .7, 1.1 and 1.5 nT/s. Panels a)-e) show results for models 2_LFM-MIX, 3_WEIGEL, 4_OPENGGCM, 6_WEIMER and 9_SWMF, respectively (see Table 3).
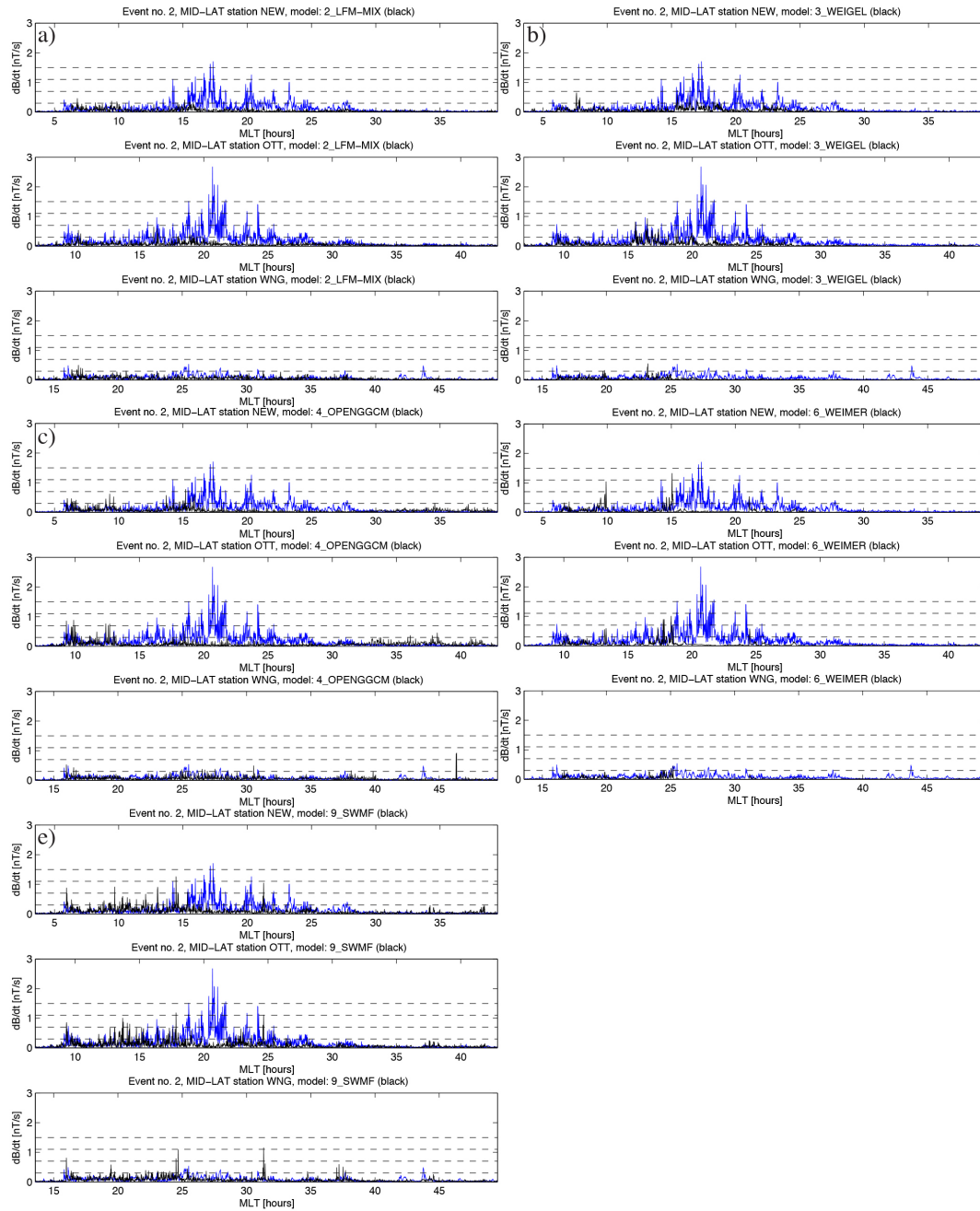
Figure 6.   Time series of the observed (blue curves) vs modeled (black curves) dB/dt at the three mid-latitude stations indicated in Table 2 for event 2 indicated in Table 1. The time is magnetic local time (MLT) and the dashed lines indicate the dB/dt thresholds of .3, .7, 1.1 and 1.5 nT/s. Panels a)-e) show results for models 2_LFM-MIX, 3_WEIGEL, 4_OPENGGCM, 6_WEIMER and 9_SWMF, respectively (see Table 3).
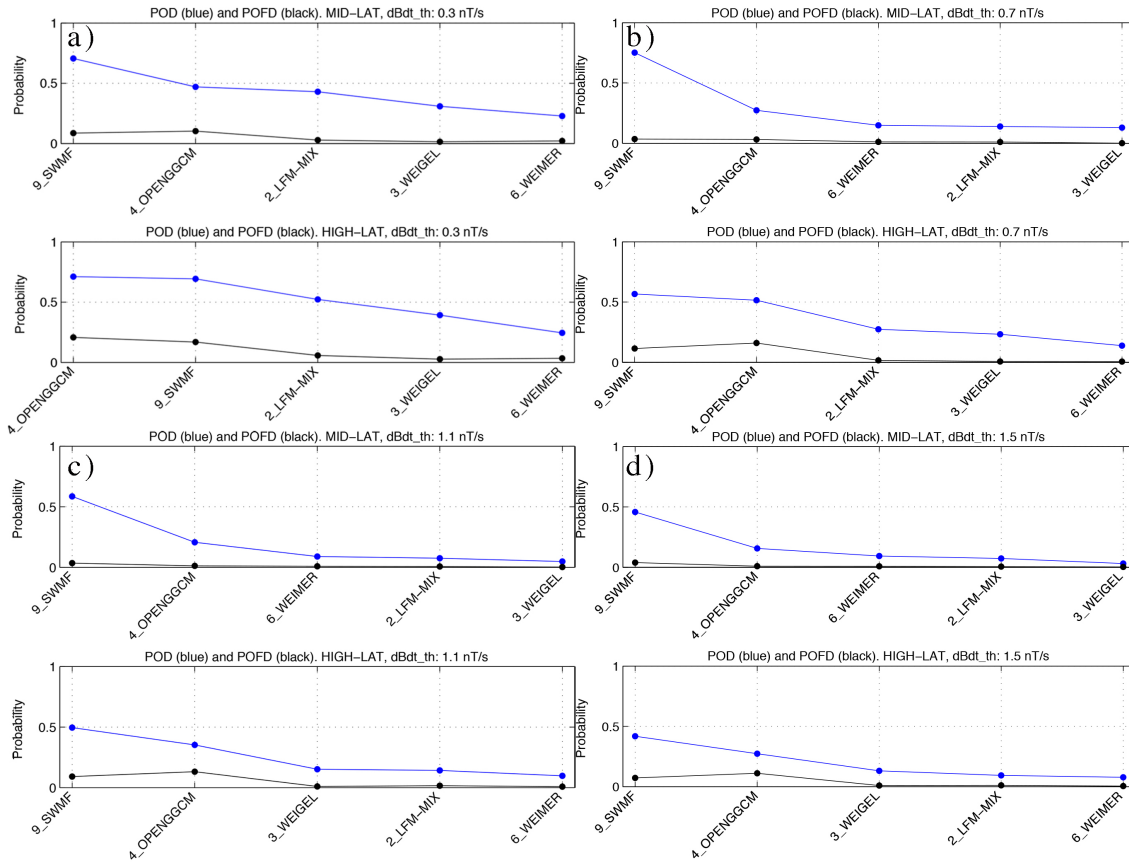
Figure 7.   Probability of Detection (POD) (blue curve) and Probability of False Detection (POFD) (black curve) defined in Section 9 for the dB/dt thresholds a) .3 nT/s, b) .7 nT/s, c) 1.1 nT/s and d) 1.5 nT/s. In all panels the top panel shows POD and POFD obtained by integrating over the three mid-latitude stations and the bottom panel shows POD and POFD obtained by integrating over the three high-latitude stations. The models (see Table 3) are ordered according to their POD. The model with the largest POD is the leftmost in all panels.
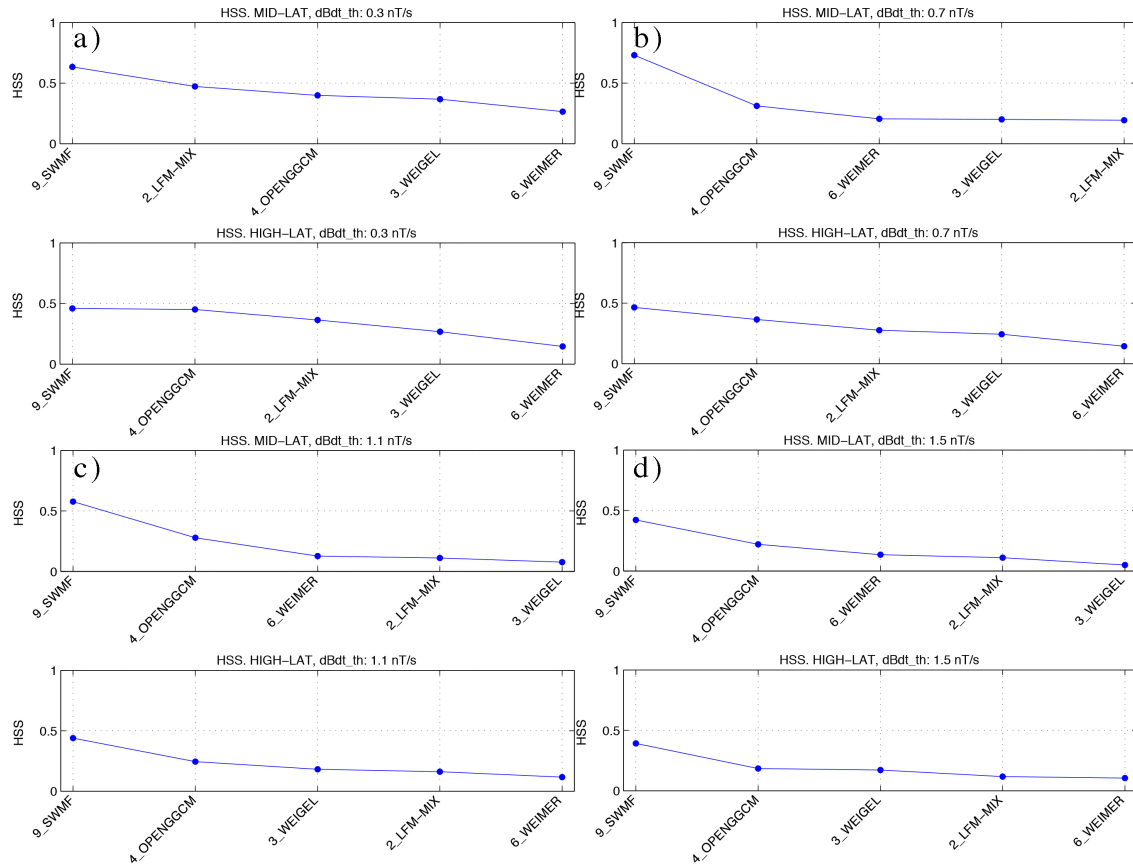
Figure 8. Heidke Skill Score (HSS) defined in Section 9 for the dB/dt thresholds a) .3 nT/s, b) .7 nT/s, c) 1.1 nT/s and d) 1.5 nT/s. In all panels the top panel shows HSS obtained by integrating over the three mid-latitude stations and the bottom panel shows HSS obtained by integrating over the three high-latitude stations. The models (see Table 3) are ordered according to their HSS. The model with the largest HSS is the leftmost in all panels.
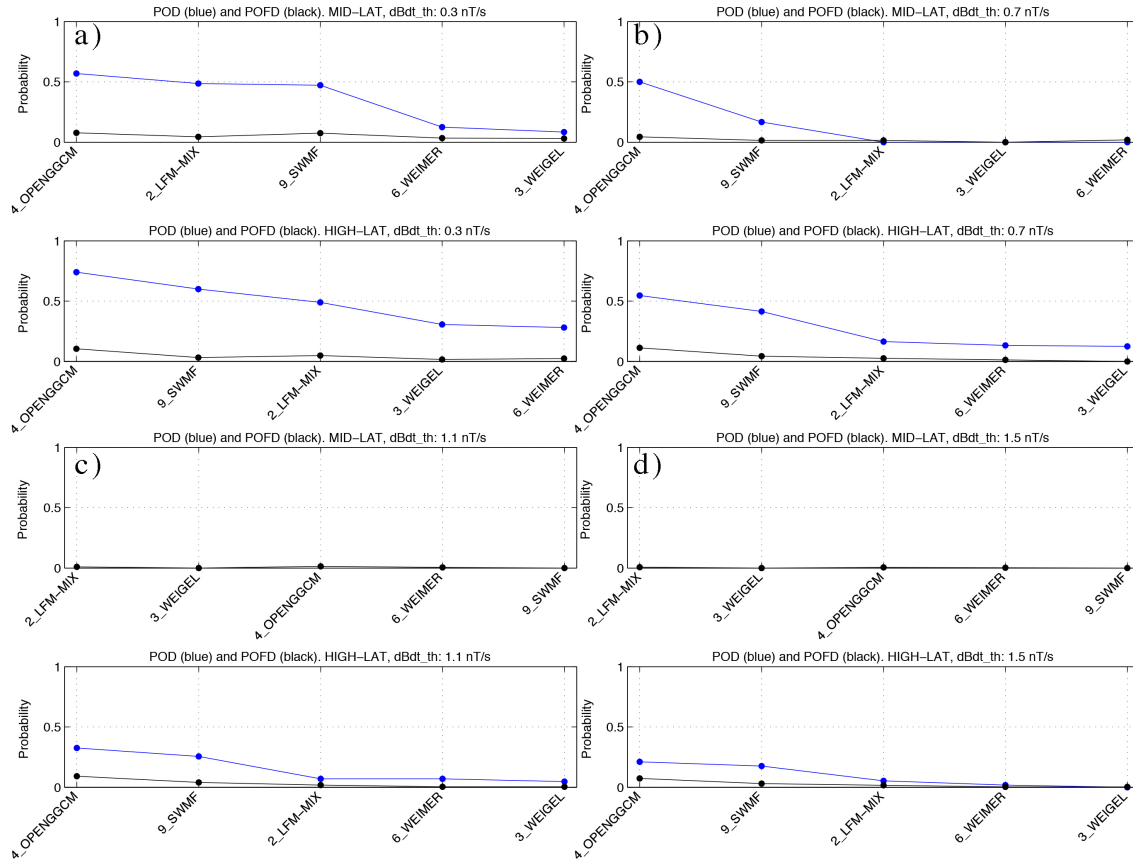
Figure 9. Probability of Detection (POD) (blue curve) and Probability of False Detection (POFD) (black curve) for "surprise events" 5 and 6 only, for the dB/dt thresholds a) .3 nT/s, b) .7 nT/s, c) 1.1 nT/s and d) 1.5 nT/s. In all panels the top panel shows POD and POFD obtained by integrating over the three mid-latitude stations and the bottom panel shows POD and POFD obtained by integrating over the three high-latitude stations. The models (see Table 3) are ordered according to their POD. The model with the largest POD is the leftmost in all panels.
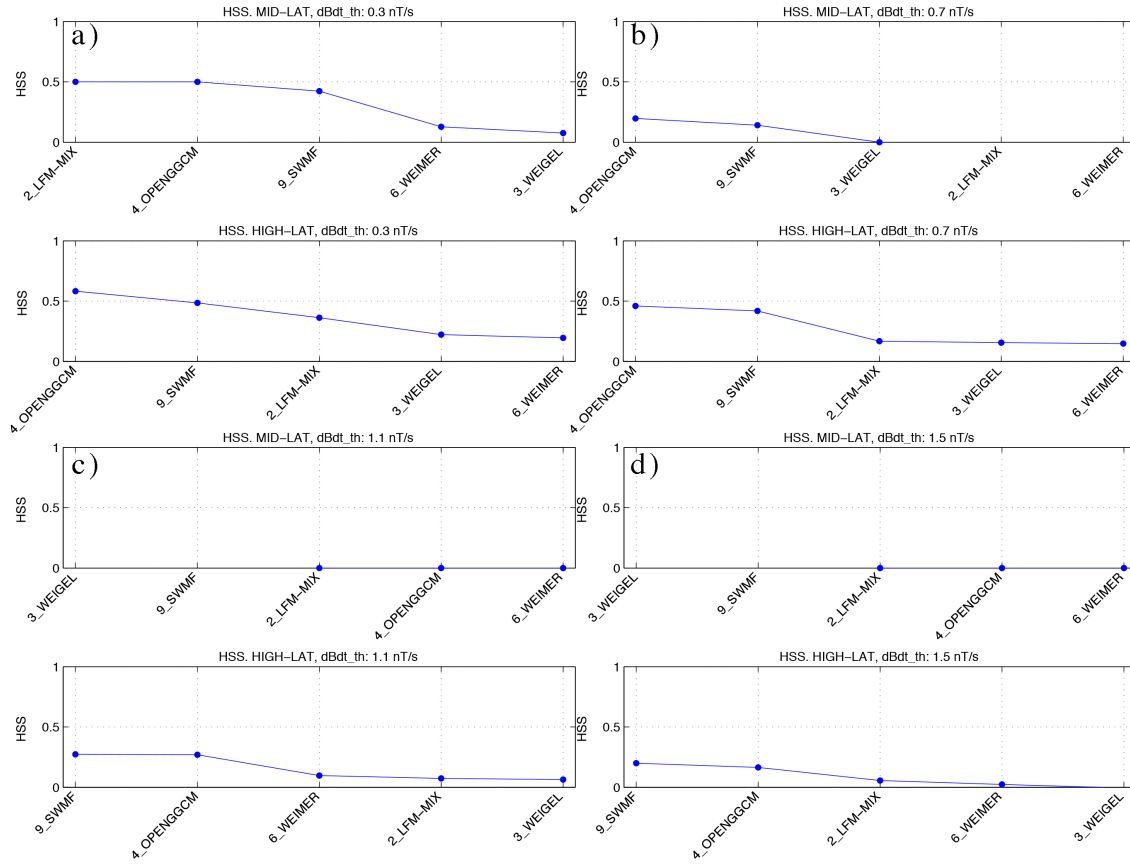
Figure 10.   Heidke Skill Score (HSS) for "surprise events" 5 and 6 only, for the dB/dt thresholds a) .3 nT/s, b) .7 nT/s, c) 1.1 nT/s and d) 1.5 nT/s. In all panels the top panel shows HSS obtained by integrating over the three mid-latitude stations and the bottom panel shows HSS obtained by integrating over the three high-latitude stations. The models (see Table 3) are ordered according to their HSS. The model with the largest HSS is the leftmost in all panels.

Groundbased magnetic perturbations dB/dt and Regional-K study results

Event	Magnetic perturbations on the ground dB/dt	Magnetic perturbations on the ground Delta B	Regional-K
2006/12/14 (doy 348) 12:00 UT - 12/16 00:00 UT	ABK FRD FRN FUR HRN IQA MEA NEW OTT PBQ WNG YKC	ABK FRD FRN FUR HRN IQA MEA NEW OTT PBQ WNG YKC	ABK NEW OTT PBQ WNG YKC
2001/08/31 (doy 243) 00:00 UT - 09/01 00:00 UT	ABK FRD FRN FUR IQA MEA NEW OTT PBQ WNG YKC	ABK FRD FRN FUR IQA MEA NEW OTT PBQ WNG YKC	ABK NEW OTT PBQ WNG YKC
2005/08/31 (doy 243) 10:00 UT - 09/01 12:00 UT	ABK FRD FRN FUR HRN MEA NEW OTT PBQ WNG YKC	ABK FRD FRN FUR HRN MEA NEW OTT PBQ WNG YKC	ABK NEW OTT PBQ WNG YKC
2003/10/29 (doy 302) 06:00 UT - 10/30 06:00 UT	ABK FRD FRN FUR HRN IQA MEA NEW OTT PBQ WNG YKC	ABK FRD FRN FUR HRN IQA MEA NEW OTT PBQ WNG YKC	ABK NEW OTT PBQ WNG YKC
2010/04/05 (doy 095) 00:00 UT - 2010/04/06 00:00	ABK FRD FRN FUR HRN IQA NEW OTT SNK WNG YKC	ABK FRD FRN FUR HRN IQA NEW OTT SNK WNG YKC	ABK NEW OTT SNK WNG YKC
2011/08/05 (doy 217) 09:00 UT - 2011/08/05 09:00	ABK FRD FRN FUR HRN IQA NEW OTT SNK WNG YKC	ABK FRD FRN FUR HRN IQA NEW OTT SNK WNG YKC	ABK NEW OTT SNK WNG YKC

You can also download Delta B, dB/dt and K-index timeseries files.

Figure 11.   The screen shot of the entry page to the interface **http://ccmc.gsfc.nasa.gov/challenges/dBdt/metrics_results.php** at the special page at the CCMC website dedicated to the operational geospace model validation. To access the time line visualization interface the user can click on the name of the selected ground station for the selected event. Time series also can be downloaded by clicking on the link at the bottom of the entry page. Regional-K time series have been added to support the Phase II validation study.
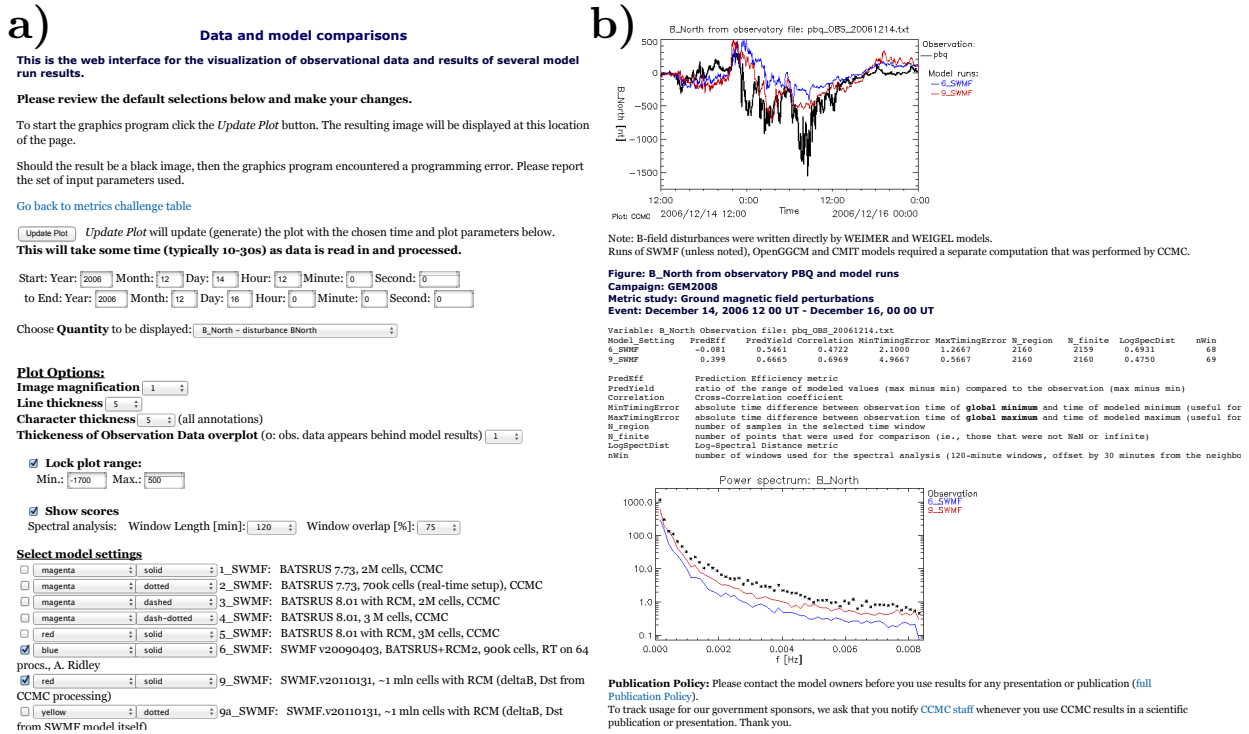
Figure 12.   Screen shot of plot interface. Figure 12a shows the interface displaying plot options and a (partial) list of available model runs for the selected study, event, and observatory (here: magnetometer station). The user can select the time period, a plot quantity from the drop-down list (here displaying "B_North" as the selected quantity), and under '"Plot Options" may modify the size of the displayed image, set the vertical plot range, and elect whether to calculate skill scores. Spectral analysis or event-based metrics have additional parameters such as window size, window overlap or threshold value (not shown). The list of runs contains options to select the color and line style of the model output traces. Figure 12b shows results returned. The skill scores include the Prediction Efficiency, Yield, Correlation Coefficient and Timing Errors (used in *Rastätter et al.* [2013a]) with explanations. The top image is the time plot and the bottom the spectral power plot with model runs data rendered with the selected colors and line styles.