

J. Space Weather Space Clim. 2020, **10**, 33  
 © N.K. Kwagala et al., Published by EDP Sciences 2020  
<https://doi.org/10.1051/swsc/2020034>



Available online at:  
[www.swsc-journal.org](http://www.swsc-journal.org)

Topical Issue - Geomagnetic Storms and Substorms:  
 a Geomagnetically Induced Current perspective

RESEARCH ARTICLE

OPEN ACCESS

# Validating the Space Weather Modeling Framework (SWMF) for applications in northern Europe

## Ground magnetic perturbation validation

Norah Kaggwa Kwagala<sup>1,\*</sup>, Michael Hesse<sup>1,3</sup>, Therese Moretto<sup>1</sup>, Paul Tenfjord<sup>1</sup>, Cecilia Norgren<sup>1</sup>, Gabor Tóth<sup>2</sup>, Tamas Gombosi<sup>2</sup>, Håkon M. Kolstø<sup>1</sup>, and Susanne F. Spinnangr<sup>1</sup>

<sup>1</sup> Space Plasma Physics Group, Department of Physics and Technology, University of Bergen, 5020 Bergen, Norway

<sup>2</sup> Department of Climate and Space, Center for Space Environment Modeling, University of Michigan, Ann Arbor, 48109 MI, USA

<sup>3</sup> Southwest Research Institute, San Antonio, 78238 TX, USA

Received 28 February 2020 / Accepted 29 June 2020

**Abstract**—In this study we investigate the performance of the University of Michigan’s Space Weather Modeling Framework (SWMF) in prediction of ground magnetic perturbations ( $\Delta B$ ) and their rate of change with time ( $dB/dt$ ), which is directly connected to geomagnetically induced currents (GICs). We use the SWMF set-up where the global magnetosphere provided by the Block Adaptive Tree Solar-wind Roe-type Upwind Scheme (BATS-R-US) MHD code, is coupled to the inner magnetosphere and the ionospheric electrodynamics. The validation is done for  $\Delta B$  and  $dB/dt$  separately. The performance is evaluated via data-model comparison through a metrics-based approach. For  $\Delta B$ , the normalized root mean square error (nRMS) and the correlation coefficient are used. For  $dB/dt$ , the probability of detection, the probability of false detection, the Heidke skill score, and the frequency bias are used for different  $dB/dt$  thresholds. The performance is evaluated for eleven ground magnetometer stations located between  $59^\circ$  and  $85^\circ$  magnetic latitude and spanning about five magnetic local times. Eight geomagnetic storms are studied. Our results show that the SWMF predicts the northward component of the perturbations better at lower latitudes ( $59^\circ$ – $67^\circ$ ) than at higher latitudes ( $>67^\circ$ ), whereas for the eastward component, the model performs better at high latitudes. Generally, the SWMF performs well in the prediction of  $dB/dt$  for a 0.3 nT/s threshold, with a high probability of detection  $\approx 0.8$ , low probability of false detection ( $< 0.4$ ), and Heidke skill score above zero. To a large extent the model tends to predict events as often as they are actually occurring in nature (frequency bias 1). With respect to the metrics measures, the  $dB/dt$  prediction performance generally decreases as the threshold is raised, except for the probability of false detection, which improves.

**Keywords:** SWMF model validation / ground magnetic perturbation / M–I coupling currents / global MHD / GICs

## 1 Introduction

The awareness of the importance of space weather with respect to safety and life on Earth has increased and gained international interest. Among the space weather effects of concern are ground magnetic perturbations, which arise as a result of currents from the solar wind–magnetosphere–ionosphere coupling/interaction. The rate of change of the ground magnetic perturbations with time ( $dB/dt$ ) is directly linked to geomagnetically induced currents (GICs) which have potential impact on high-power voltage transmission systems (e.g., Boteler et al., 1998;

Pirjola, 2005). In addition to this, unexpected geomagnetic perturbations ( $\Delta B$ ) can disrupt navigation activities that are dependent on the ground geomagnetic field direction, such as directional drilling used by the oil and gas industry (e.g., Reay et al., 2005). Therefore, the ability to forecast or predict these effects of the solar wind–magnetosphere–ionosphere coupling in space and on the ground is crucial. A number of challenges and campaigns have been carried out in the past related to evaluating how well the existing models perform in reproducing observations (e.g., Pulkkinen et al., 2011, 2013; Rastätter et al., 2013). The most recent challenge for the ground magnetic perturbations is the community-wide validation of geospace model magnetic perturbations predictions to support model

\*Corresponding author: [norah.kwagala@uib.no](mailto:norah.kwagala@uib.no)

transition to operation in Pulkkinen et al. (2013). The University of Michigan's space weather modelling framework (SWMF) (Tóth et al., 2005, 2012) was then selected and transitioned to operation at the NOAA Space Weather Prediction Center (SWPC). In our study, we further assess the performance of the SWMF in predicting regional and localised ground magnetic perturbations in the northern Europe sector. With quantifiable metrics, we aim at further understanding the strengths and weaknesses of SWMF in prediction of ground magnetic perturbations.

The SWMF is a software framework that enables coupling of different space physics systems in a single simulation. In this study, for the global magnetosphere domain the Block-Adaptive Roe-type Upwind Scheme (BATS-R-US) MHD code (Powell et al., 1999) is used and coupled to the Rice Convection Model (RCM) of the inner magnetosphere (IM) (Wolf et al., 1982; Sazykin et al., 2002; Toffoletto et al., 2003), and the Ridley ionosphere model (RIM) (Ridley et al., 2004) which is an ionospheric electrodynamics (IE) solver. BATS-R-US is an adaptive mesh magnetohydrodynamics (MHD) code that solves the ideal MHD equations throughout the magnetosphere (Powell et al., 1999; De Zeeuw et al., 2000).

Understanding model capabilities to reproduce observed features in the signal of interest is a key element of space weather monitoring and forecasting. One of the key elements investigated in this work is the rate of change of the horizontal component of ground magnetic perturbations with time,  $dB/dt$ . The primary argument for studying  $dB/dt$  is that the time derivative of ground magnetic field can be used as an indicator for level of geomagnetically induced electric field or geoelectric field, on the surface of the Earth (e.g., Viljanen et al., 2001). Hence regional and local predictions of the ground geomagnetic perturbations are a necessity for modelling and forecasting space weather impacts in areas such as power grids and resource pipelines. The other element studied in this work is the magnetic perturbation on the ground  $\Delta B$  which is important in navigation aspects such as directional drilling, where the geomagnetic field is followed for navigation.

The paper is organised as follows. The setup settings of the SWMF simulations, as well as the storm events and ground magnetometer used, are described in Section 2. Section 3 details the metrics used to quantify the model performance. The results are presented in Section 4 followed by a discussion of our results in Section 5 and finally a summary is given in Section 6.

## 2 Validation setting

### 2.1 SWMF configuration

The SWMF is driven by solar wind plasma (velocity and density) and interplanetary magnetic field input data at the upstream L1 location from the advanced composition explorer (ACE) and WIND satellites. The other input is the F10.7 cm flux. Using a tool provided on the Community Coordinated Modeling Center (CCMC) website ([https://ccmc.gsfc.nasa.gov/requests/GetInput/get\\_ace\\_K.php](https://ccmc.gsfc.nasa.gov/requests/GetInput/get_ace_K.php)), data were propagated from the satellite position to  $x = 33 R_E$  which is just outside the upstream boundary of the simulation box, with averaged  $V_x$  velocity. The results presented in this paper are for runs with highest resolution of  $0.125 R_E$  close to the inner boundary and lowest

resolution of  $8 R_E$  in the distant tail. The computational domain extends from  $32 R_E$  upstream to  $-224 R_E$  downstream in the  $x$  direction and  $\pm 128$  in the  $y$  and  $z$  GSM coordinates. The inner boundary is  $2.5 R_E$  from Earth's center. The Sokolov scheme (Sokolov et al., 2002) with Koren's limiter (Koren, 1993) and  $\beta = 1.2$  is used. The integrated density and pressure are applied as outer boundary conditions for inner magnetosphere model assuming a 80%  $H^+$  to 20%  $O^+$  number density ratio.

The global magnetosphere (GM) component of the SWMF transfers energy and mass from the solar wind to the magnetosphere and it is responsible for the convection in the magnetosphere (Welling & Ridley, 2010). The GM component provides the IE component with field aligned current density and the IM component with magnetic field structure and plasma density (De Zeeuw et al., 2004). In return, the GM component receives density and pressure information from the IM component and electric field from the IE component. The IM component also receives electric potential from IE component.

### 2.2 Storm events

Eight storm events are examined in this study. Six of these (events 1–5, 7) were used in the GEM Challenges (e.g., Pulkkinen et al., 2010, 2011, 2013; Rastätter et al., 2011, 2013), while event 8 is one of those recommended by Welling et al. (2018) for inclusion in the validation suite and event 6 is a new event included in our validation. Events 3, 7, and 8 are known major storms, the AGU, the Halloween and the St. Patrick's Day storms, respectively. A list of all the 8 storm events as well as their F10.7 flux, maximum AE index, and the minimum SYM-H index is shown in Table 1.

### 2.3 Ground magnetometer stations

Table 2 shows the names and geomagnetic coordinates of the eleven ground magnetometer stations used in this study. It should be noted that not all the stations had data for all of the eight events. The last column in Table 2 shows the events for which we do not have data at the corresponding stations. The geomagnetic coordinates of the stations are entered as inputs of the SWMF and act as locations of virtual magnetometer stations where the model predicts/calculates the ground geomagnetic perturbations. These are the predicted perturbations that are compared to the actual perturbations measured at the ground-based stations. The magnetometers used here are located in the northern Europe region between magnetic latitude  $59^\circ$  and  $85^\circ$ , spanning 5 h of MLT as shown in Figure 1. The data from the stations were transformed from geographic coordinates into geomagnetic dipole coordinates using the Apex python tool which uses the IGRF-12 with coefficients from 1900 to 2020 (Thébault et al., 2015).

## 3 Selected metrics

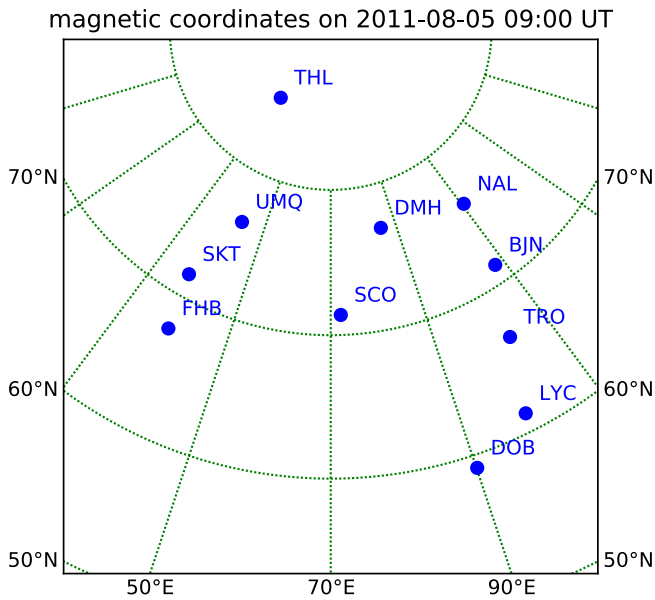
The performance of the model for each event is evaluated via model-data comparison. The first parameter evaluated is the ground magnetic field perturbation  $\Delta B$ . Only the horizontal components of the ground geomagnetic field are investigated in the analysis presented in this paper. The  $\Delta B$  comparison is done using the normalized root mean square error (nRMS) defined in

**Table 1.** List of events.

Event	Start time	End time	f10.7	AE	SYM-H
1	31 August 2001, 0000 UT	01 September 0000 UT	192.2	959	-46
2	31 August 2005, 0900 UT	01 September 1200 UT	85.6	2063	-119
3	14 December 2006, 1100 UT	16 December 0000 UT	90.5	2284	-211
4	05 April 2010, 0000 UT	06 April 0000 UT	79.3	2565	-67
5	05 August 2011, 0900 UT	06 August 0900 UT	112.5	2611	-126
6	22 January 2012, 0000 UT	22 January 2345 UT	136.6	1028	-79
7	29 October 2003, 0600 UT	30 October 0600 UT	275.4	4056	-391
8	16 March 2015, 0000 UT	19 March 0000 UT	113.2	2298	-234

**Table 2.** The locations of the geomagnetic observatories used in the study.

Station name	Station code	Geomag. latitude	Geomag. longitude	No data
Thule	THL	85.0	30.8	0
Danmarkshavn	DMH	77.0	85.4	Events 4, 5
Uummannaq	UMQ	76.5	43.1	Event 1
Ny Ålesund	NAL	76.0	110.6	0
Sukkertoppen	SKT	71.6	37.3	Events 1, 3
Scoresbysund	SCO	71.4	72.2	Event 1, 4, 8
Bjørnøya	BJN	71.3	108.0	0
Fredrikshåb	FHB	67.6	39.0	Events 2, 3
Tromsø	TRO	66.5	102.9	0
Lycksele	LYC	61.3	99.3	Events 2, 3, 7
Dombås	DOB	59.1	90.1	Event 5



**Fig. 1.** Ground magnetometer locations.

equation (1) and cross correlation coefficient. This is done independently for the northward ( $\Delta B_n$ ) and eastward ( $\Delta B_e$ ) components. The subscripts “p” and “o” in equation (1) refer to prediction and observation, respectively:

$$\text{nRMS} = \frac{\sqrt{\langle (\Delta B_p - \Delta B_o)^2 \rangle}}{\sqrt{\langle \Delta B_o^2 \rangle}}. \quad (1)$$

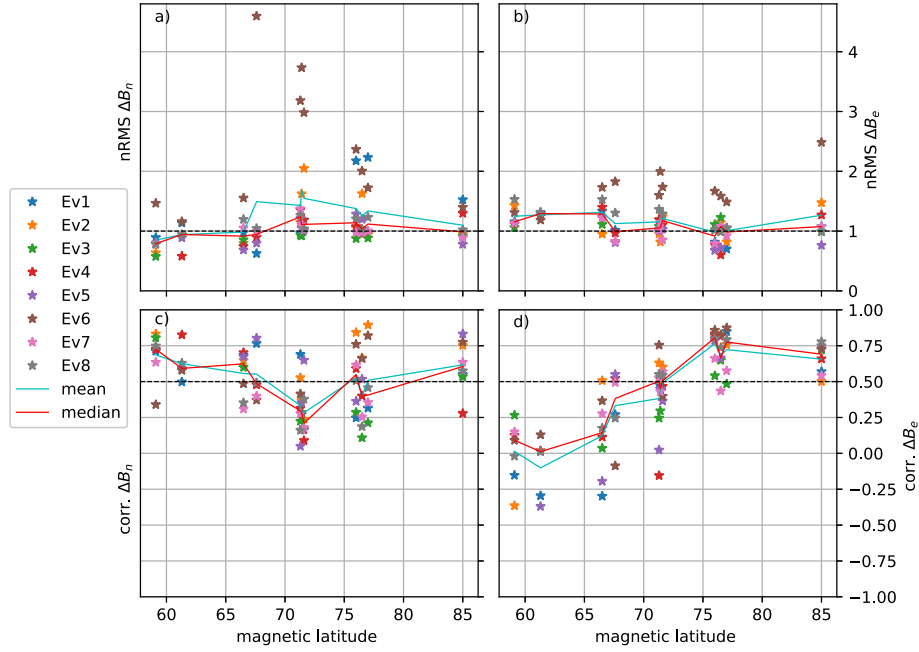
The average is taken over the entire simulation period. A nRMS error of zero indicates that the model exactly reproduces the observations, while nRMS less than one indicates that the model prediction is in good agreement with the observations. A nRMS error above one indicates that the model prediction significantly diverges from the observations, i.e., may possibly have opposite trends or large offsets.

The second parameter evaluated in this work is the  $dB/dt$ , similar to what was used in prior work (e.g., Pulkkinen et al., 2013). The magnitude of the horizontal  $dB/dt$  is directly related to the geomagnetically induced currents (GICs) since it generates the geoelectric field (e.g., Viljanen et al., 2004; Pulkkinen et al., 2006). Following Pulkkinen et al. (2013), for each event and station  $dB/dt$  is defined as

$$dB/dt = \sqrt{(dB_n/dt)^2 + (dB_e/dt)^2} \quad (2)$$

where  $B_n$  and  $B_e$  are the two horizontal components (northward and eastward) of the magnetic field. The magnetometer stations record data at different intervals of 1 s, 10 s and 20 s. For the validation, we take the data point every minute to match the time interval of data extracted from the model. The central difference is used to calculate  $dB/dt$  for both observation and model data.

An event-based analysis is carried out, where an event is defined as the absolute value exceeding an event threshold within a forecast window  $0 \leq t \leq t_f$ . For this study we use thresholds 0.3, 0.7, 1.1, and 1.5 nT/s and a forecast window of 20 min. These are the same thresholds that have been used in the previous validation studies (e.g., Pulkkinen et al., 2013). For each storm event a contingency table is obtained containing the number of correctly predicted threshold crossings



**Fig. 2.** Normalised root mean square error (nRMS) in panels a and b and cross correlation coefficient (corr.) between the ground magnetometer observations and the model  $\Delta B_n$  and  $\Delta B_e$  in panels c and d for all events and magnetometer stations. The cyan and red lines show the mean and median metrics value at each magnetometer station.

$H$  (hits), the number of false alarms  $F$ , the number of missed crossings  $M$ , and the number of correctly predicted no crossings  $N$ . The set  $(H, F, M, N)$  is used to calculate different metrics values to quantify the model performance. The selected metrics include probability of detection (POD), probability of false detection (POFD), Heidke skill score (HSS), and model frequency bias (FB). The first three metrics have been used in the previous validation challenges, while the FB is an additional metric recommended for inclusion by [Welling et al. \(2018\)](#).

### 3.1 Probability of detection (POD)

POD is defined as

$$\text{POD} = \frac{H}{H + M} \quad (3)$$

which is a measure of the fraction of observed threshold crossings that were correctly forecast. POD ranges from 0 to 1 with 1 being the perfect score.  $\text{POD} > 0.5$  implies that the model tends to predict more correct crossings than it misses.

### 3.2 Probability of false detection (POFD)

POFD is defined as

$$\text{POFD} = \frac{F}{F + N} \quad (4)$$

which measures the fraction of observed no crossing events that are incorrectly forecast as crossings. POFD ranges from 0 to 1 with 0 as the perfect score.  $\text{POFD} < 0.5$  implies that the model tends to predict more correct no crossings than false crossings.

### 3.3 Heidke skill score (HSS)

HSS is defined as

$$\text{HSS} = \frac{2(HN - MF)}{(H + M)(M + N) + (H + F)(F + N)} \quad (5)$$

which 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 one. Negative values indicate that random forecast is better than model prediction while zero is no skill, just as good as random. 1 is the perfect score.

### 3.4 Frequency bias

Frequency bias (FB) is defined as

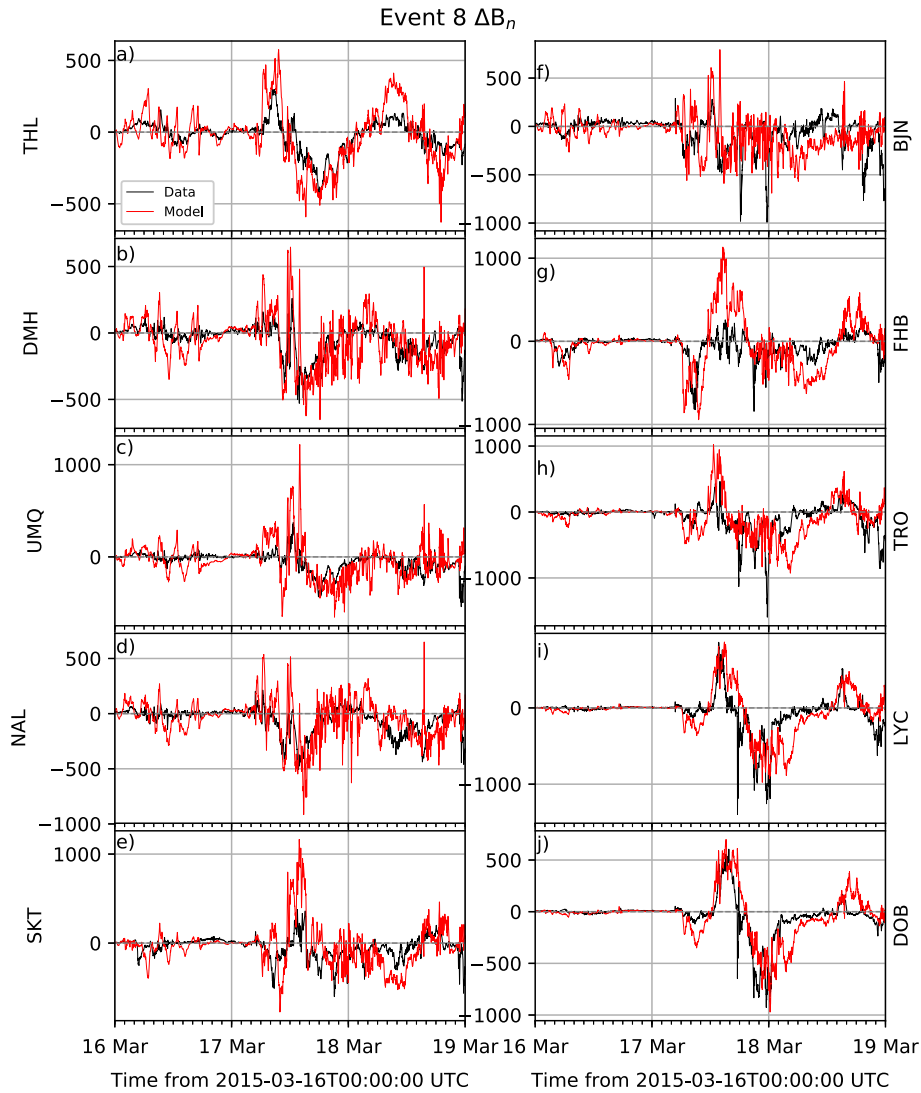
$$\text{FB} = \frac{H + F}{H + M} \quad (6)$$

which measures how well the forecasts correspond to the observations.  $\text{FB} < 1$  indicates that the model predicts events less frequently than nature, whereas  $\text{FB} > 1$  indicates that the model predicts events more frequently than nature. 1 is the perfect score.

## 4 Results

### 4.1 nRMS error and cross correlation coefficient at different magnetic latitudes

[Figures 2a](#) and [2c](#) show the nRMS error and the correlation coefficient, respectively, for the northward component  $\Delta B_n$  of



**Fig. 3.** Ground magnetic northward perturbations,  $\Delta B_n$ , at all ground magnetometer stations for event 8. The black line shows the observed perturbations and the red line shows the modelled ground perturbations.

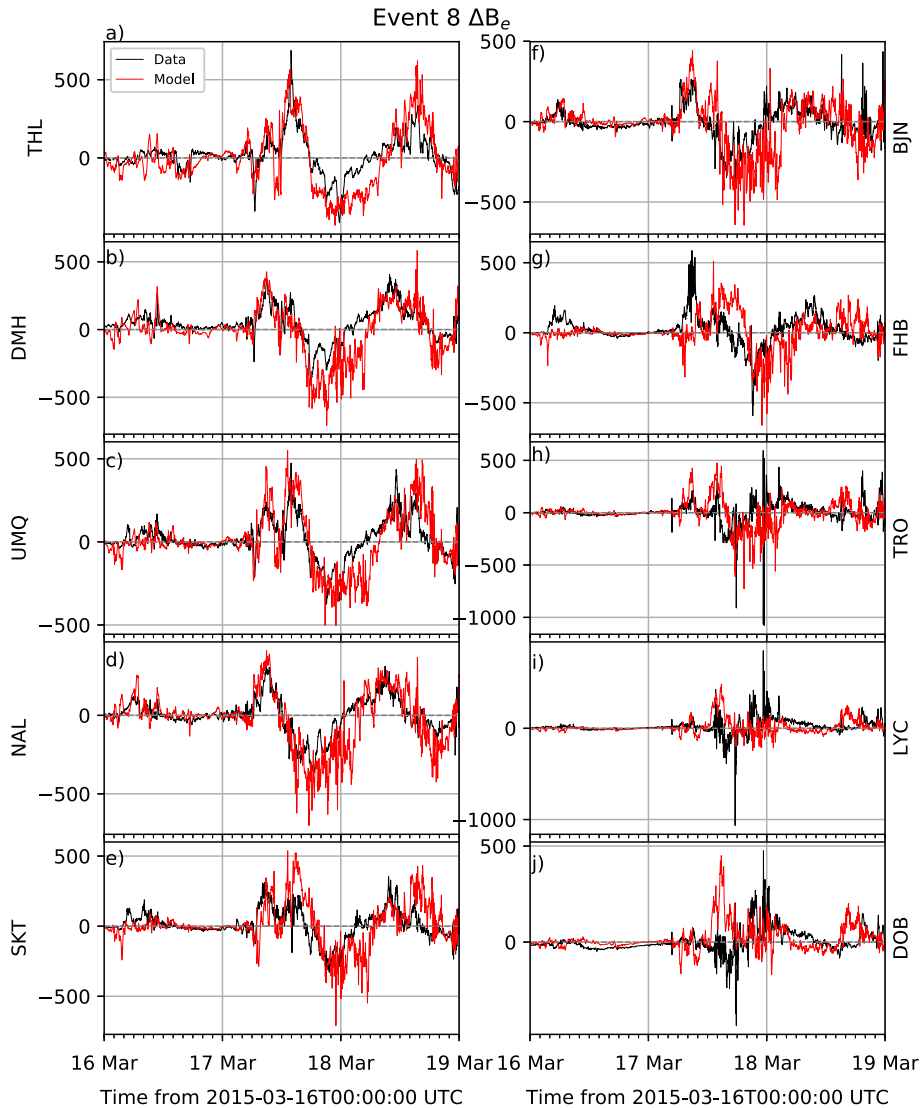
the ground magnetic field perturbations at different magnetic latitudes for the different events 1–8 (Ev1–Ev8). [Figures 2b](#) and [2d](#) show the eastward component  $\Delta B_e$  of the perturbations. The cyan line shows the mean at each magnetometer station while the red line shows the median. A high correlation coefficient indicates that the model predicts similar trends as the observations, whereas nRMS error below 1 indicates that the model predicted magnitudes are comparable to those observed. Generally, our results show that the model tends to predict the northward component perturbations better at lower (sub-auroral) magnetic latitude ( $59^\circ$ – $68^\circ$ ), with mean nRMS  $\leq 1$  and correlation coefficient above 0.5. On the other hand, for the eastern component, the model tends to perform better at higher magnetic latitudes ( $>68^\circ$ ). For the polar cap stations, though, the performance is similar for both components.

One storm event is selected for further analysis. The selected event is the St. Patrick’s Day storm of 2015 (Event 8). The simulation for Event 8 is three days long making it the longest of the eight events simulated in this study. [Figures 3](#) and [4](#) show

the data-model comparison for the northward and eastward components, respectively for Event 8.

Generally, the model seems to capture the large scale trend of the ground perturbations, but sometimes tends to exaggerate the magnitude of the northward component especially at the high magnetic latitudes ([Fig. 3](#)). This overestimation would contribute to the nRMS error  $>1$  at these latitudes. Another striking feature is the large amplitude perturbation that is short lived at lower latitude stations (TRO, LYC, DOB), and which is completely missed by the model at around 18:00 UT on 17 March.

For the eastward component shown in [Figure 4](#), the model generally reproduces the observed perturbations, particularly the longer period variations, to a good degree at the high latitudes both in magnitude and trend. However, the model performs poorly at the lower latitudes. The short duration large amplitude perturbations observed in the northward component, are also observed in the eastward component and the model completely misses them here as well. The fact that these pulses are missed in both components suggest that they likely arise from highly

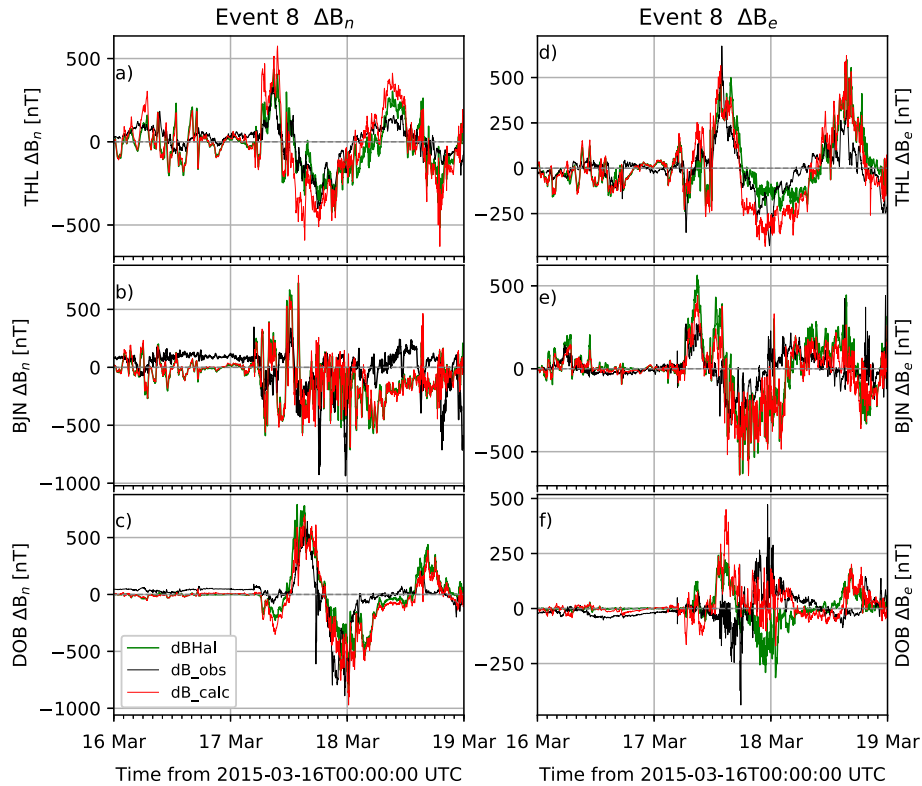


**Fig. 4.** Ground magnetic eastward perturbations,  $\Delta B_e$  at all ground magnetometer stations for event 8. The black line shows the observed perturbations and the red line shows the modelled ground perturbations.

localised enhancements in currents that are not captured by the model (e.g., Yu & Ridley, 2008). The other feature is during the early hours of 18 March, where the model exaggerates the negative eastward component at the high latitude stations.

We further investigate the eastward and northward component perturbations and the modelled contributions from the different current sources. For this analysis, three of the ground magnetometer stations, THL ( $85^\circ$ ) at the highest latitude (polar cap) in the study, BJN ( $71.3^\circ$ ) for auroral latitude, and the lowest latitude (subauroral) DOB ( $59.1^\circ$ ) for Event 8 are selected. Figure 5 shows the contribution from Hall current (green line) compared with the observed (black line) and modelled (red line) ground perturbations. The figure shows that the Hall currents contribute the most to predicted total perturbation for both northward and eastward components, with the exception of the eastward component at the lowest latitude station DOB.

The total contribution from the field aligned currents (FACs) and Pedersen currents is shown by the cyan line in Figure 6. For vertical magnetic field lines, uniform Hall and Pedersen conductances and a horizontal ionosphere, the ground magnetic perturbations due to FACs and Pedersen currents should exactly cancel (e.g., Schield et al., 1969; Ridley et al., 2004). If they did we would expect the total contribution from FACs and Pedersen currents to be zero at the polar cap station THL where the magnetic field lines are most vertical. These ideal conditions clearly do not apply in reality since Figure 6a shows a non-zero contribution from the FACs and Pedersen currents at THL. The total modeled eastward component at the lowest latitude (DOB) is also dominated by the FACs and Pedersen current contributions. On comparing Figures 5 and 6, the magnitude of the exaggeration during the recovery period clearly arises from exaggeration of the Hall currents contribution. The contribution from the other magnetosphere currents (MHD) is shown by the magenta line in Figure 6 and it is negligible at the magnetic latitudes shown.



**Fig. 5.** Two components of ground magnetic perturbations,  $\Delta B_n$  (left side panels) and  $\Delta B_e$  (right side panels), at ground magnetometer stations THL, BJJ, and DOB for event 8. The black line shows the observed perturbations, the red line shows the model total ground perturbations, and the green line shows the modelled perturbation due to Hall current.

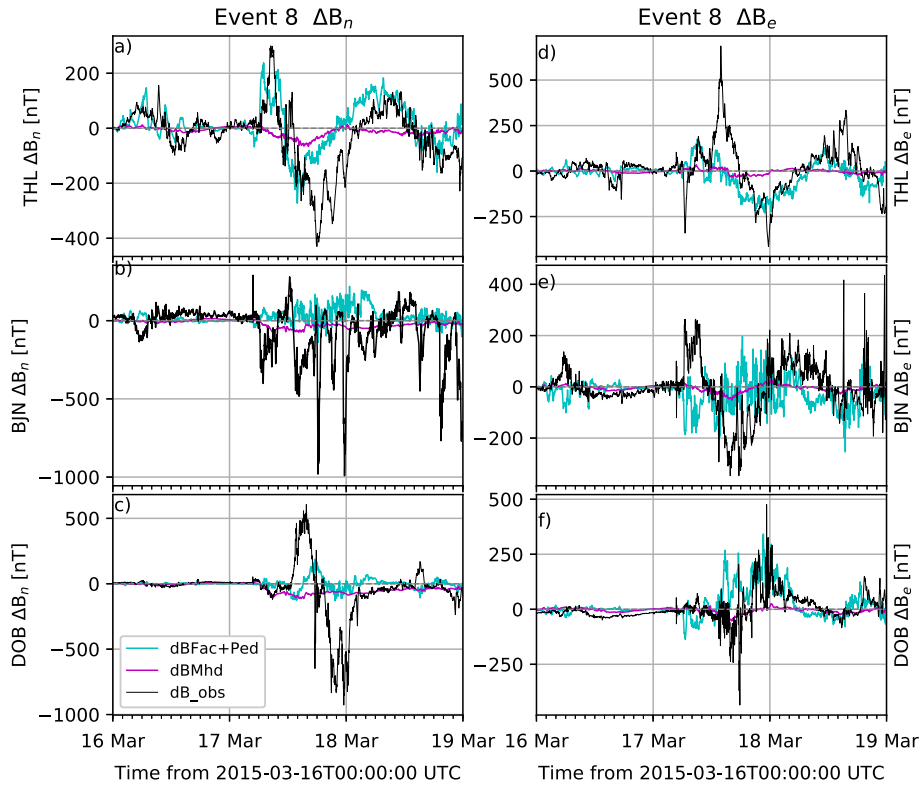
#### 4.2 dB/dt metrics-based analysis at different magnetic latitudes

Figure 7 presents the four dB/dt metrics POD, POFD, HSS, and FB for all the stations at different magnetic latitudes for all events for a dB/dt threshold of 0.3 nT/s. Generally the model performance varies from event to event and for different latitudes. To capture the general performance, we plot the mean (cyan line) and median (red line) at each station as well. The results show that the model generally performs reasonably well with high mean and median POD score of  $\approx 0.75$  and 0.8, respectively, and low POFD (mostly  $< 0.4$ ) and prediction skill better than a random guess i.e.,  $HSS > 0$  although mainly below 0.5. The model also tends to predict events as often as they actually occur in nature at lower latitudes with a mean FB close to one and overestimates the frequency of occurrence at higher latitudes. When the model prediction frequency diverts from nature ( $FB \neq 1$ ), it tends to underestimate the frequency of occurrence of events at the lower latitudes. To summarize the model performance at 0.3 nT/s threshold, we plot POD against POFD in Figure 8a and HSS against FB in Figure 8b. The dashed lines in Figure 8a mark probability 0.5. For a perfect model prediction all the points should be in the left upper corner. This is where the majority of the points are, however some of the high POD scores are associated with high POFD, above 0.5 (right upper corner). The vertical dashed line in Figure 8b marks the perfect FB score ( $FB = 1$ ) and the horizontal line marks the no skill score ( $HSS = 0$ ). Ideally, we want all the points to be along the vertical line

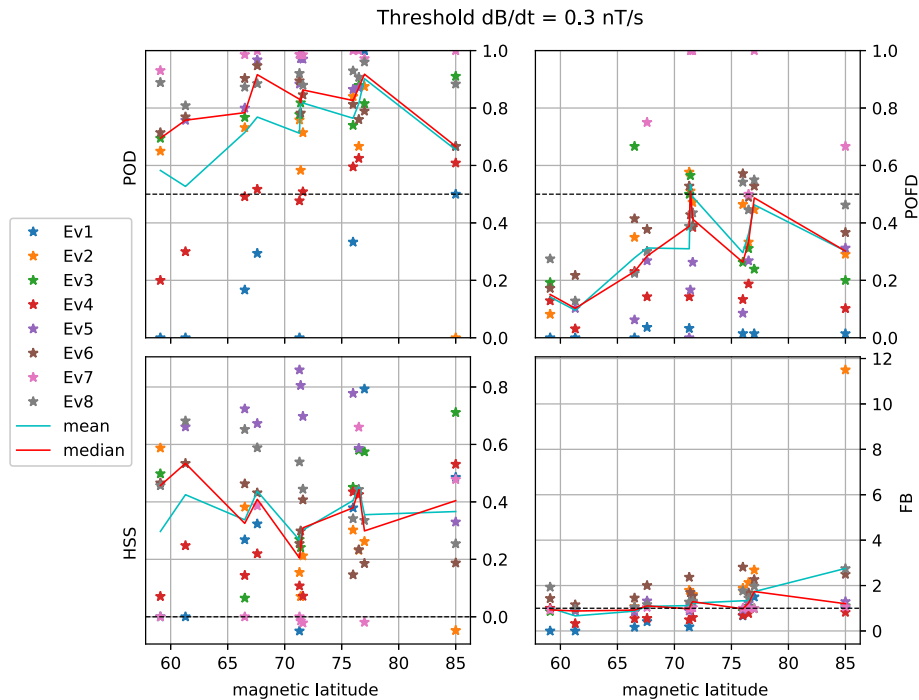
and above the horizontal line and this is where most of the points are observed. Since the above results are based on the 0.3 nT/s threshold, the result is not very meaningful for very intense storm events like Event 7 (Halloween storm 2003). For this reason, the most intense events 3, 7, and 8 are investigated further at all thresholds 0.3, 0.7, 1.1, and 1.5 nT/s.

Figure 9 shows the summary plot similar to Figure 8 for the four dB/dt thresholds (0.3, 0.7, 1.1, and 1.5 nT/s) for Events 3, 7, and 8. Generally, the performance with respect to the dB/dt metrics decreases with increasing threshold, except for the POFD which improves with increasing threshold. This implies that any predicted high threshold crossings are most likely to be true events. The Heidke skill score gradually drops at higher thresholds. The frequency bias FB value shows that the model tends to predict at a frequency closer to that of nature i.e.,  $FB = 1$  at the lowest threshold shown (0.3 nT/s), however, the frequency of prediction diverges from nature as the threshold is increased.

Similar to the previous analysis of the  $\Delta B$  perturbations, we now look at the direct data-model comparison for dB/dt for event 8 at all ground magnetometer stations. The comparison here is done using smoothed data (both model and observations) with a running mean window of 20 min which is the forecast window we use in the dB/dt metrics analysis. Figure 10 shows the comparison. The stations are arranged decreasing in magnetic latitudes from (a) to (j). This shows that for event 8 the model tends to overestimate the amplitudes at higher latitudes and underestimate at the lower latitudes.

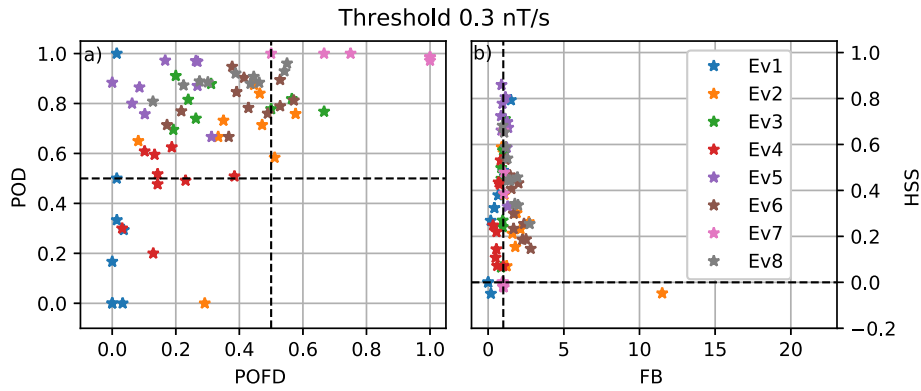


**Fig. 6.** Two components of ground magnetic perturbations,  $\Delta B_n$  (panels a–c) and  $\Delta B_e$  (panels d–f), at ground magnetometer stations THL, BJN, and DOB for event 8. The black line shows the observed perturbations, the cyan line shows the modelled total contribution to ground perturbations due to field aligned currents (FAC) and Pedersen currents and the magenta line is the contribution from the rest of the magnetospheric currents.

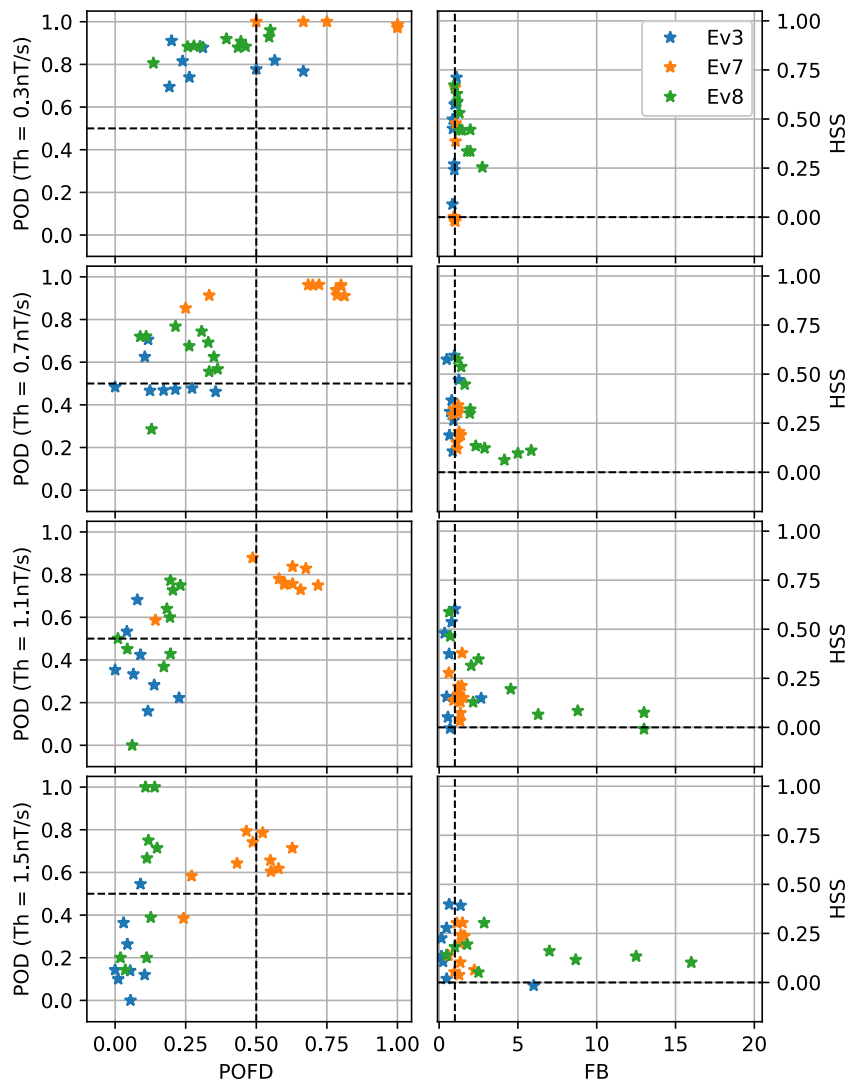


**Fig. 7.**  $dB/dt$  metrics for threshold 0.3 nT/s for all events and magnetometer stations. The cyan and red lines show the mean and median metrics value at each magnetometer station. Metrics shown are probability of detection (POD), Probability of false detection (POFD), Heidke skill score (HSS), and model frequency bias (FB).

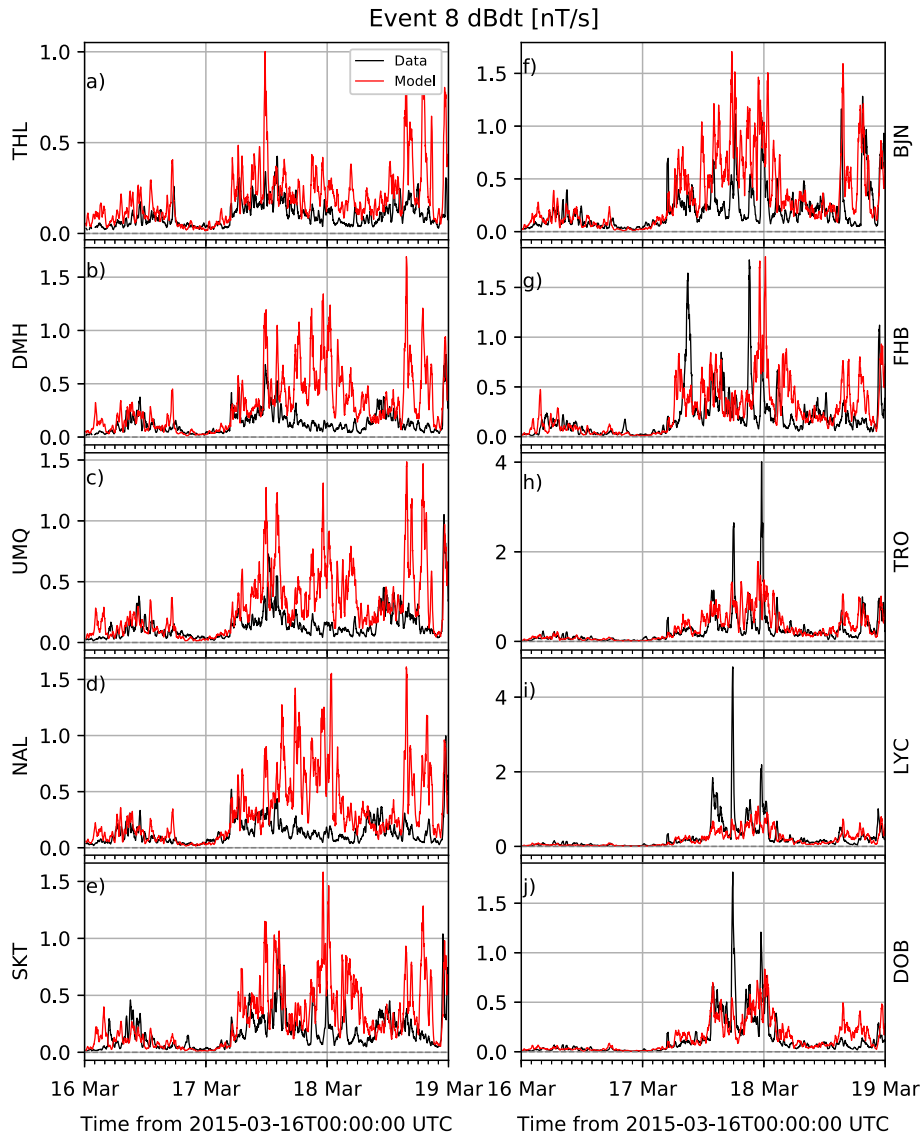




**Fig. 8.** Panel a shows the probability of detection (POD) versus probability of false detection (POFD). The horizontal line in panel a) marks  $POD = 0.5$  while the vertical line marks  $POFD = 0.5$ . Panel b shows the Heidke skill score (HSS) versus model frequency bias (FB). The horizontal line shows no skill score,  $HSS = 0$  while the vertical line marks the perfect score for frequency bias,  $FB = 1$ .



**Fig. 9.**  $dB/dt$  metrics results for thresholds 0.3, 0.7, 1.1, and 1.5 nT/s for the most intense storm events 3, 7, and 8. Probability of detection (POD) versus probability of false detection (POFD) in the left side panels, where dashed lines mark probability 0.5. An ideal panel would have all the points in the left top box in each panel. The panels on the right hand side show Heidke skill score (HSS) versus frequency bias (FB) where the horizontal dashed line marks the no skill score ( $HSS = 0$ ) and the vertical dashed line marks the perfect FB score ( $FB = 1$ ). An ideal model would have all the points above the horizontal line and along the vertical line.



**Fig. 10.** Ground magnetic  $dB/dt$  at all ground magnetometer stations for event 8. The black line shows the observed perturbations and the red line shows the modelled ground perturbations. For this figure both observation and model are smoothed using a moving average window of 20 min.

## 5 Summary and discussion

In this work, we have investigated the performance of the SWMF in predicting ground magnetic perturbations in the northern Europe sector. Both the perturbations  $\Delta B$  and  $dB/dt$  have been investigated. The performance has been evaluated using a metric-based analysis similar to those used in the past validations (e.g., Yu & Ridley, 2008; Pulkkinen et al., 2011, 2013). We have presented the model performance per station per event. It is clear that the model performance varies at each event, however, we have presented the mean and median as a proxy for the general performance (Figs. 2 and 7).

The first part of the analysis focusses on investigating how well the model predicts the horizontal components of the ground magnetic perturbations  $\Delta B_n$  and  $\Delta B_e$ . From the mean and median of the normalized root-mean-square error and the

cross correlation coefficients in Figure 2, the model tends to perform better at lower latitudes (close to subauroral latitudes) in prediction of the northward component. On the other hand, the model tends to perform better in the predicting the eastward component at higher latitudes (auroral and polar cap latitudes). A more detailed analysis of event 8 reveals that the model sometimes tends to exaggerate the amplitude of the northward perturbations at auroral latitudes. This enhancement is likely to contribute to the nRMS error above one at these latitudes. The model captures the trend but sometimes exaggerates the magnitude. For example, Figure 5b at BJJ suggests that the exaggeration arises from the predicted Hall current contribution. Through the Biot-Savart law, the northward ground magnetic perturbations would be induced by an eastward current in the ionosphere (e.g., Yu et al., 2010). This indicates that the model is predicting a higher (Hall) current density in the eastward

direction than what actually exists. A deeper investigation into this is a topic for a follow up study.

There are a number of features which may arise from misplacement of the currents in the SWMF with respect to the magnetometer stations. One of such features is the eastward component of the perturbations that is in the opposite direction to the observed perturbations, for example at the lower latitudes from FHB to DOB (71°–59° Figs. 4g–4j) at 10–18 UT on 17th March. This could explain the poor  $\Delta B_e$  prediction performance at these latitudes. These are directly connected to the northward current density in the ionosphere. Such oppositely directed eastward component perturbations were also reported by Yu & Ridley (2008).

The short lived large amplitude perturbations observed at lower latitude stations FHB, TRO, LYC, and DOB that are greatly underestimated or completely missed (Figs. 3 and 4) also correspond to a large  $dB/dt$  (Fig. 10) lasting just a few minutes. These perturbations could be due to very localised current structure as the signature is only observed from 18.50 to 19.60 MLT between 59° and 66.70° magnetic latitude. We however note that we do not look at all stations in this region to confirm this.

The second part of our analysis focusses on investigating how well the model predicts the rate of change of the ground magnetic perturbations  $dB/dt$ , a parameter crucial in the derivation of the geoelectric field. The geoelectric field is a key contributor to the GIC. With respect to all the four metrics POD, POFD, HSS, and FB, the model performs reasonably well. Unlike the  $\Delta B$ , there is no clear magnetic latitude dependence seen in the  $dB/dt$  metrics analysis mainly because it is the magnitude of horizontal  $dB/dt$  that is investigated. Therefore the errors of the  $\Delta B$  estimates shown in event 8, like the overestimate in magnitude and oppositely directed perturbations, would have no effect, especially for the lower  $dB/dt$  thresholds. However, a direct  $dB/dt$  data-model comparison shown in Figure 10 reveals that the model tends to miss large  $dB/dt$  at the lower latitudes while at the higher latitudes it tends to overestimate the  $dB/dt$  amplitudes. The summary in Figure 8 clearly shows a good performance at the lowest 0.3 nT/s threshold. This performance decreases with increasing threshold except with respect to POFD, which actually improves. This implies that any event prediction made by the model at large thresholds will most-likely be an actual event in reality. The results reported by Pulkkinen et al. (2013), which were obtained by integrating over several stations and all their studied events, showed that both POD and HSS were below 0.5 for threshold 1.5 nT/s. This is in agreement with our results except that our results show POD > 0.5 at this threshold for some stations. The higher POD could also be attributed to the higher grid resolution (highest  $1/8 R_E$ ) in the SWMF configuration we use. Increasing the grid resolution has been reported to improve the prediction performance of global MHD models (e.g., Pulkkinen et al., 2010, 2011).

One of the possible sources of the discrepancy between the SWMF predictions and the observations could arise from the uncertainty in the solar wind data driving the model. There is a possibility that the solar wind seen by the satellite at a point in the L1 position is not necessarily what reaches the magnetopause and in addition, the methods used to propagate the L1 measurement to the bow shock can also introduce errors to the solar wind input (e.g., Morley et al., 2018). Morley et al. (2018) showed that the SWMF prediction skill can be

improved by taking into account the uncertainty in the solar wind input.

Different configurations of SWMF can also affect the discrepancy between the predictions and observations. Some of the configurations or imperfect specifications that could lead to uncertainty in the model include empirical ionospheric conductance (e.g., Welling et al., 2017) and low grid resolution (e.g., Haiducek et al., 2017). A detailed investigation of the different sources of discrepancy will be the topic of a follow-up study.

## 6 Summary

We performed a validation of the SWMF in predicting ground magnetic perturbations for the northern Europe sector. The key results show that the model:

- Performs better at high magnetic latitudes above 67° than at lower latitudes 58°–67° in predicting the eastward component of the magnetic perturbations on the ground and the reverse is true for the northward component.
- Captures the general trend but tends to overestimate the magnitude of northward ground magnetic perturbations particularly at the high magnetic latitudes.
- Generally performs reasonably well for  $dB/dt$  predictions for the 0.3 nT/s threshold.
- Predicts the low threshold (0.3 nT/s) crossings at a rate very close to the frequency of occurrence in reality. This changes when the threshold is increased, in which case the model diverges from the real-world rate underestimating at lower latitudes and sometimes overestimating at higher magnetic latitudes.
- Mostly underestimates the large amplitude short-lived perturbations likely associated with localised current structures.

*Acknowledgements.* We thank Tromsø Geophysical Observatory, UiT the Arctic University of Norway, for providing the data from the Norwegian Magnetometer Line. Magnetometer data from the Greenland Magnetometer Array were provided by the National Space Institute at the Technical University of Denmark (DTU Space). Part of the results presented in this paper rely on data collected at magnetic observatories. We thank the institutes who maintain the IMAGE Magnetometer Array: Troms Geophysical Observatory of UiT the Arctic University of Norway (Norway), Finnish Meteorological Institute (Finland), Institute of Geophysics Polish Academy of Sciences (Poland), GFZ German Research Centre for Geosciences (Germany), Geological Survey of Sweden (Sweden), Swedish Institute of Space Physics (Sweden), Sodankylä Geophysical Observatory of the University of Oulu (Finland), and Polar Geophysical Institute (Russia). We thank the ACE SWEPAM instrument team, the ACE MAG team and the ACE Science Center for providing the ACE data. We thank the WIND Solar Wind Experiment (SWE) and WIND Magnetic Field Investigation (MFI) teams for providing WIND data. The solar index F10.7 were provided by the NASA OMNIWeb service (<https://omniweb.gsfc.nasa.gov/>). We acknowledge the developers of the Space Weather Modeling Framework

and the Center for Space Environment Modeling at the University of Michigan. The Spacepy python library (Morley et al., 2011) was used to read and analyse the SWMF output. It is available at <https://github.com/spacepy/spacepy>. The simulations were performed on resources provided by UNINETT Sigma2 – the National Infrastructure for High Performance Computing and Data Storage in Norway. This project has been funded by the European Space Agency under the PRODEX Experiment contract 4000124903. The editor thanks two anonymous reviewers for their assistance in evaluating this paper.

## References

- Boteler DH, Pirjola RJ, Nevanlinna H. 1998. The effects of geomagnetic disturbances on electrical systems at the Earth's surface. *Adv Space Res* **22**(1): 17–27.
- De Zeeuw DL, Gombosi TI, Groth CPT, Powell KG, Stout QF. 2000. An adaptive MHD method for global space weather simulations. *IEEE Trans Plasma Sci* **28**(6): 1956–1965. <https://doi.org/10.1109/27.902224>.
- De Zeeuw DL, Sazykin S, Wolf RA, Gombosi TI, Ridley AJ, Tóth G. 2004. Coupling of a global MHD code and an inner magnetospheric model: Initial results. *J Geophys Res Space Phys* **109**(A12): A12219. <https://doi.org/10.1029/2003JA010366>.
- Haiducek JD, Welling DT, Ganushkina NY, Morley SK, Ozturk DS. 2017. SWMF global magnetosphere simulations of January 2005: Geomagnetic indices and cross-polar cap potential. *Space Weather* **15**(12): 1567–1587. <https://doi.org/10.1002/2017SW001695>.
- Koren B. 1993. A robust upwind discretization method for advection, diffusion and source terms. In: *Numerical methods for advection-diffusion problems*, Vreugdenhil CB, Koren B (Eds.), Notes on numerical fluid mechanics, **45**, Vieweg, Germany, pp. 117–138.
- Morley SK, Welling DT, Koller J, Larsen BA, Henderson MG, Niehof JT. 2011. Spacepy – A Python-based library of tools for the space science. In: *Proc. 9th Python in Science Conference (SciPy 2010)*, pp. 39–45.
- Morley SK, Welling DT, Woodroffe JR. 2018. Perturbed input ensemble modeling with the space weather modeling framework. *Space Weather* **16**(9): 1330–1347. <https://doi.org/10.1029/2018SW002000>.
- Pirjola R. 2005. Effects of space weather on high-latitude ground systems. *Adv Space Res* **36**: 2231–2240.
- Powell KG, Roe PL, Linde TJ, Gombosi TI, Zeeuw DLD. 1999. A solution-adaptive upwind scheme for ideal magnetohydrodynamics. *J Comput Phys* **154**(2): 284–309. <https://doi.org/10.1006/jcph.1999.6299>.
- Pulkkinen A, Viljanen A, Pirjola R. 2006. Estimation of geomagnetically induced current levels from different input data. *Space Weather* **4**(8): S08005. <https://doi.org/10.1029/2006SW000229>.
- Pulkkinen A, Rasttter L, Kuznetsova M, Hesse M, Ridley A, Raeder J, Singer HJ, Chulaki A. 2010. Systematic evaluation of ground and geostationary magnetic field predictions generated by global magnetohydrodynamic models. *J Geophys Res Space Phys* **115** (A3): A03206. <https://doi.org/10.1029/2009JA014537>.
- Pulkkinen A, Kuznetsova M, Ridley A, Raeder J, Vapirev A, et al. 2011. Geospace environment modeling 2008–2009 challenge: Ground magnetic field perturbations. *Space Weather* **9**(2): S02004. <https://doi.org/10.1029/2010SW000600>.
- Pulkkinen A, Rasttter L, Kuznetsova M, Singer H, Balch C, et al. 2013. Community-wide validation of geospace model ground magnetic field perturbation predictions to support model transition to operations. *Space Weather* **11**(6): 369–385. <https://doi.org/10.1002/swe.20056>.
- Rasttatter L, Kuznetsova MM, Vapirev A, Ridley A, Wiltberger M, Pulkkinen A, Hesse M, Singer HJ. 2011. Geospace Environment Modeling 2008–2009 Challenge: Geosynchronous magnetic field. *Space Weather* **9**(4): S04005. <https://doi.org/10.1029/2010SW000617>.
- Rasttatter L, Kuznetsova MM, Glocer A, Welling D, Meng X, et al. 2013. Geospace environment modeling 2008–2009 challenge: Dst index. *Space Weather* **11**(4): 187–205. <https://doi.org/10.1002/swe.20036>.
- Reay SJ, Allen W, Baillie O, Bowe J, Clarke E, Lesur V, Macmillan S. 2005. Space weather effects on drilling accuracy in the North Sea. *Ann Geophys* **23**(9): 3081–3088. <https://doi.org/10.5194/angeo-23-3081-2005>.
- Ridley AJ, Gombosi TI, DeZeeuw DL. 2004. Ionospheric control of the magnetosphere: Conductance. *Ann Geophys* **22**(2): 567–584. <https://doi.org/10.5194/angeo-22-567-2004>.
- Sazykin S, Wolf RA, Spiro RW, Gombosi TI, De Zeeuw DL, Thomsen MF. 2002. Interchange instability in the inner magnetosphere associated with geosynchronous particle flux decreases. *Geophys Res Lett* **29**(10): 88–1–88–4. <https://doi.org/10.1029/2001GL014416>.
- Schild MA, Freeman JW, Dessler AJ. 1969. A source for field-aligned currents at auroral latitudes. *J Geophys Res (1896–1977)* **74**(1): 247–256. <https://doi.org/10.1029/JA074i001p00247>.
- Sokolov IV, Timofeev EV, Sakai J-I, Takayama K. 2002. Artificial wind – A new framework to construct simple and efficient upwind shock-capturing schemes. *J Comput Phys* **181**(1): 354–393. <https://doi.org/10.1006/jcph.2002.7130>.
- Thébault E, Finlay C, Beggan C, Alken P, Aubert J, et al. 2015. International Geomagnetic Reference Field: The 12th generation. *Earth Planet Space* **67**: 79. <https://doi.org/10.1186/s40623-015-0228-9>.
- Toffoletto F, Sazykin S, Spiro R, Wolf R. 2003. Inner magnetospheric modeling with the Rice Convection Model. *Space Sci Rev* **107**(1): 175–196. <https://doi.org/10.1023/A:1025532008047>.
- Tóth G, Sokolov IV, Gombosi TI, Chesney DR, Clauer CR, et al. 2005. Space Weather Modeling Framework: A new tool for the space science community. *J Geophys Res Space Phys* **110**(A12): A12226. <https://doi.org/10.1029/2005JA011126>.
- Tóth G, Holst BVD, Sokolov IV, De Zeeuw DL, Gombosi TI, et al. 2012. Adaptive numerical algorithms in space weather modeling. *J Comput Phys* **231**(3): 870–903. Special Issue: Computational Plasma Physics, <https://doi.org/10.1016/j.jcp.2011.02.006>.
- Viljanen A, Nevanlinna H, Pajunpää K, Pulkkinen A. 2001. Time derivative of the horizontal geomagnetic field as an activity indicator. *Ann Geophys* **19**(9): 1107–1118. <https://doi.org/10.5194/angeo-19-1107-2001>.
- Viljanen A, Pulkkinen A, Amm O, Pirjola R, Korja T, Group BW. 2004. Fast computation of the geoelectric field using the method of elementary current systems and planar Earth models. *Ann Geophys* **22**(1): 101–113. <https://doi.org/10.5194/angeo-22-101-2004>.
- Welling DT, Ridley AJ. 2010. Validation of SWMF magnetic field and plasma. *Space Weather* **8**(3). <https://doi.org/10.1029/2009SW000494>.
- Welling DT, Anderson BJ, Crowley G, Pulkkinen AA, Rasttatter L. 2017. Exploring predictive performance: A reanalysis of the geospace model transition challenge. *Space Weather* **15**(1): 192–203. <https://doi.org/10.1002/2016SW001505>.

- Welling DT, Ngwira CM, Opgenoorth H, Haiducek JD, Savani NP, et al. 2018. Recommendations for next-generation ground magnetic perturbation validation. *Space Weather* **16(12)**: 1912–1920. <https://doi.org/10.1029/2018SW002064>.
- Wolf RA, Harel M, Spiro RW, Voigt G-H, Reiff PH, Chen C-K. 1982. Computer simulation of inner magnetospheric dynamics for the magnetic storm of July 29, 1977. *J Geophys Res Space Phys* **87(A8)**: 5949–5962. <https://doi.org/10.1029/JA087iA08p05949>.
- Yu Y, Ridley AJ. 2008. Validation of the space weather modeling framework using ground-based magnetometers. *Space Weather* **6(5)**: S05002. <https://doi.org/10.1029/2007SW000345>.
- Yu Y, Ridley AJ, Welling DT, Tlith G. 2010. Including gap region field-aligned currents and magnetospheric currents in the MHD calculation of ground-based magnetic field perturbations. *J Geophys Res Space Phys* **115(A8)**: A08207. <https://doi.org/10.1029/2009JA014869>.

**Cite this article as:** Kwagala NK, Hesse M, Moretto T, Tenfjord P, Norgren C, et al. 2020. Validating the Space Weather Modeling Framework (SWMF) for applications in northern Europe. *J. Space Weather Space Clim.* **10**, 33.