

# Solar wind driven empirical forecast models of the time derivative of the ground magnetic field

Peter Wintoft<sup>1,\*</sup>, Magnus Wik<sup>1</sup>, and Ari Viljanen<sup>2</sup>

<sup>1</sup> Swedish Institute of Space Physics, Scheelevägen 17, 223 70 Lund, Sweden

\*Corresponding author: peter@lund.irf.se

<sup>2</sup> Finnish Meteorological Institute, Erik Palménin Aukio 1, 00560 Helsinki, Finland

Received 29 October 2014 / Accepted 16 February 2015

## ABSTRACT

Empirical models are developed to provide 10–30-min forecasts of the magnitude of the time derivative of local horizontal ground geomagnetic field ( $|d\mathbf{B}_h/dt|$ ) over Europe. The models are driven by ACE solar wind data. A major part of the work has been devoted to the search and selection of datasets to support the model development. To simplify the problem, but at the same time capture sudden changes, 30-min maximum values of  $|d\mathbf{B}_h/dt|$  are forecast with a cadence of 1 min. Models are tested both with and without the use of ACE SWEPAM plasma data. It is shown that the models generally capture sudden increases in  $|d\mathbf{B}_h/dt|$  that are associated with sudden impulses (SI). The SI is the dominant disturbance source for geomagnetic latitudes below 50° N and with minor contribution from substorms. However, at occasions, large disturbances can be seen associated with geomagnetic pulsations. For higher latitudes longer lasting disturbances, associated with substorms, are generally also captured. It is also shown that the models using only solar wind magnetic field as input perform in most cases equally well as models with plasma data. The models have been verified using different approaches including the extremal dependence index which is suitable for rare events.

**Key words.** Solar wind – Ground magnetic field – Forecast

## 1. Introduction

The main factor for space weather-related effects on electrical power grids are geomagnetically induced currents (GIC). The effects of GIC can either be immediate, leading to voltage fluctuations, tripping of relays, heating and possible loss of transformers (Bolduc 2002) or cumulative with the build-up of damaging levels of gasses in the transformer oil (Gaunt 2014). Several studies have been carried out that investigate space weather events, and related GIC and technological effects (Lundstedt 2006; Wik et al. 2009; Schrijver & Mitchell 2013).

The distribution and magnitude of GIC in a power grid are determined by a geophysical component (electric fields, ground conductivity) and a technological component (grid topology, resistances) (Pirjola et al. 2004). Due to limited direct measurement of the electric field, such as measurements carried out at a few locations in the UK (private comm. A. Thomson) and Hungary (Kis et al. 2007), techniques have been developed to compute the E-fields from observed geomagnetic fields such as the complex image method (Boteler & Pirjola 1998), thin-sheet approximation (Beggan et al. 2013), finite element method (Dong et al. 2013). In this work we have used the spherical elementary current system (SECS) approach (Viljanen et al. 2004).

The aim of this work is to develop models predicting a quantity that can be related to electric fields using upstream solar wind data. Empirical and data driven modelling have been applied to many different geomagnetic indices, originating from the work by Burton et al. (1975) for a solar wind – *Dst* model. The data driven approach to modelling assumes

some form of the input-output mapping function, either from basic physical reasoning, or from a generic function, and finds the unknown coefficients from observed data. The target output should capture some essential feature of the phenomena under study, e.g. indices for various geomagnetic and ionospheric processes (Mayaud 1980). Previous studies have shown that there is a good relation between rate-of-change of the magnetic field  $d\mathbf{B}/dt$  and calculated electric field and measured GIC (Viljanen et al. 2001; Wintoft 2005), where the temporal resolution of  $B$  should be 1 min or better to capture the dominant variability in  $d\mathbf{B}/dt$  (Wintoft 2005). It is not feasible to predict the detailed variation of  $d\mathbf{B}/dt$  at 1-min resolution, therefore a summary measure is adopted. The relationship between the envelopes of geomagnetic disturbances and GIC is stronger than between the detailed variations themselves (Trichtchenko & Boteler 2004). In Wintoft (2005) the 10-min root-mean-square (RMS) of the magnitude of the horizontal derivative ( $|d\mathbf{B}_h/dt|$ ) was used. In this work a similar approach is used based on 30-min maximum values driven by 1-min real-time solar wind data.

## 2. Forecast parameter, data selection and models

### 2.1. Datasets

The data consist of ground geomagnetic field measurements, and solar wind plasma and magnetic field. Geomagnetic data, at 1-min resolution, are obtained from World Data Centre for Geomagnetism (Edinburgh).<sup>1</sup> One minute differences are

<sup>1</sup> <http://www.wdc.bgs.ac.uk/catalog/master.html>

formed as an approximation to the time derivative ( $dB_x/dt$ ,  $dB_y/dt$ ,  $dB_z/dt$ ) in the  $XYZ$  coordinate system, where  $X$  points north,  $Y$  east and  $Z$  down. The data quality is generally high, however, clearly erroneous  $d\mathbf{B}/dt$  values, with 1–2 min duration, are replaced with interpolated values. These appear as single spikes in  $d\mathbf{B}/dt$  associated with sudden shift in baseline of  $\mathbf{B}$  from 1 min to the next, and as double spikes with different signs in  $d\mathbf{B}/dt$  associated with a single spike in  $\mathbf{B}$  at a single station during quiet periods.

The solar wind plasma and magnetic field data come from two instruments on the Advanced Composition Explorer (ACE) spacecraft (Stone et al. 1998). The SWEPAM instrument (McComas et al. 1998) provides the bulk density and velocity with 64 s resolution. The magnetic field instrument (Smith et al. 1998) has a resolution of 16 s. The ACE Level 2 data are obtained from The ACE Science Center web site.<sup>2</sup> Both plasma and magnetic field are resampled to 1-min resolution and data gaps of typically less than 15 min are replaced by linearly interpolated values. Data gaps are more common in the density while the magnetic field data have very few data gaps.

## 2.2. Time derivative of horizontal geomagnetic field

We make the approximation that the horizontal electric field  $\mathbf{E}$  is proportional to  $d\mathbf{B}_h/dt$ , i.e.  $(E_x, E_y) \propto (dB_y/dt, dB_x/dt)$ . This step neglects the ground conductivity which affects both the amplitude and power spectrum of the electric field (Cagniard 1953; Viljanen et al. 2012). However, the simplification is still reasonable as shown in Section 2.5. The magnitude of the horizontal rate-of-change, defined here as  $e_h$ , is formed

$$\begin{aligned} e_h &= |d\mathbf{B}_h/dt| = \sqrt{(dB_x/dt)^2 + (dB_y/dt)^2} \propto E_h \\ &= |\mathbf{E}_h|. \end{aligned} \quad (1)$$

## 2.3. Temporal filtering

During events associated with sudden impulses (SI) and geomagnetic storms the changes in  $e_h$  are often abrupt with most of the variability taking place for time scales less than 30 min, and with the peak in the power spectrum around 1 min or less (Wintoft 2005). From 10-second magnetic field data it can be shown that the peak in power spectrum is at, or slightly below, 1 min indicating that a sampling rate of 30 s (Nyquist rate) or better should be used in order to resolve the largest peaks in  $d\mathbf{B}/dt$ . However, using 1-min resolution data in estimating  $\mathbf{E}$  works reasonably well (Pulkkinen et al. 2006) although the amplitude of  $\mathbf{E}$  will be slightly underestimated. A similar variability is seen in measured GIC at a location in southern Sweden. The response at other locations will be different due to geographical differences in ionospheric and magnetospheric processes, and ground conductivity.

Typical prediction lead times range from about 30 min, or less, to a couple of hours depending on the dominant geophysical response. The shortest lead times are associated with solar wind shocks and other sudden increases in pressure which creates SI's, while a longer lead time may be possible depending on local time arrival of the solar wind disturbance and the following geomagnetic storm. In this work a constant prediction lead time of 30 min is used.

From the two considerations above, temporal forward moments are formed

$$e_h^{(a)}(t) = \left( \frac{1}{30} \sum_{t'=t+1}^{t+30} e_h^a(t') \right)^{\frac{1}{a}}, \quad (2)$$

where  $t$  is time in minutes. The first two moments are the arithmetic mean ( $a = 1$ ) and the root-mean-square ( $a = 2$ ), respectively. Higher moments will give higher weight to large values in  $e_h$ . In the extreme case with  $a \rightarrow \infty$  it becomes the 30-min maximum, i.e.  $e_h^{(\infty)}(t) = \max_{t'=t+1}^{t+30} \{e_h(t')\}$ . Applying the average ( $a = 1$ ), used in e.g. Weigel et al. (2002), will result in strong damping of the signal, more strongly for short-lived features like SI's. Another aspect is that for a sudden increase in  $e_h$  there will be a gradual increase in  $e_h^{(1)}$ . At the other extreme, with  $a \rightarrow \infty$ , the response will be immediate to any increases in  $e_h$ . Thus, an SI will be captured by  $e_h^{(\infty)}$ , but it will not resolve at what time within the 30-min interval the maximum occurs. For different types of activities, like SI's, substorms or pulsations,  $e_h^{(\infty)}$  will follow the envelope of the disturbance but the temporal location of individual peaks within 30-min intervals will not be resolved. As mentioned in the Introduction, this work is based on the 30-min maximum of  $|d\mathbf{B}_h/dt|$ , i.e.  $e_h^{(\infty)}$ . The 30-min maximum  $d\mathbf{B}_h/dt$  was also used in the study by Schrijver & Mitchell (2013).

## 2.4. Data selection

Table 1 shows the 90% and 99% percentiles, and the maximum of  $e_h^{(\infty)}$  (30-min maximum  $|d\mathbf{B}_h/dt|$ ) for all available data for stations with good temporal coverage. The stations are ordered with increasing geomagnetic latitude. It is seen that 90% of the time  $e_h^{(\infty)}$  is below 10 nT/min, except for the very high latitude stations. At the 99% level it is still quite small and typically less than 20 nT/min up to geomagnetic latitudes of 55° N. The maximum value for each station range from about 100 nT/min to close to 3000 nT/min. Any detailed statistical comparison between stations is not possible in this table as they cover different time spans and have various degree of data gaps, and therefore are subject to different levels of geomagnetic activity.

In Table 2, a subset of stations is selected for which there are good temporal overlap between stations. Further, only samples for which data exist simultaneously for all stations have been included. Therefore, the statistics formed for this set will include data for which all stations are subject to the same geomagnetic events. Generally,  $e_h^{(\infty)}$  increases with increasing geomagnetic latitude, although there are a few exceptions to the general trend. The table also lists the computed conductances for different integration depths of 40, 80 and 160 km, respectively, where the ground resistivity used are those given in the map by Adam et al. (2012). It is seen that, to some degree, stations at similar geomagnetic latitudes have larger observed  $d\mathbf{B}_h/dt$  when the conductance is higher. Of course, the individual stations may be subject to more localised ground conductance effects than those captured by the map. The final column shows the  $p$  coefficient from Eq. (4) and is described in the next section.

The database is searched for events with increased activity in  $e_h$ . To specify criteria that defines an event is not a trivial task. The start time of an event could be defined as the time of the SI. However, not all storms are associated with a clear SI, and the identification of SI's is complex and their is no real

<sup>2</sup> <http://www.srl.caltech.edu/ACE/ASC/>

**Table 1.** The table lists for each station (Id) the geomagnetic latitude (Lat), the start and end dates for the period when data exist, the number of 30-min samples ( $N$ ), and the 90% and 99% percentiles, and the maximum of  $e_h^{(\infty)}$  for all available data in units of nT/min. The magnetic coordinates are based on IGRF for the year 2010. Note that the statistics is only meaningful per station basis and that comparison between stations cannot be done as the they cover different periods and have varying degree of data gaps.

Id	Lat	Start date	End date	$N$	90%	99%	Max
AQU	42.2	1992-01-01	2009-12-31	310,773	3.6	6.4	130
SUA	42.2	1999-01-01	2009-12-31	160,910	3.2	9.5	206
SPT	42.4	1997-01-01	2010-12-31	245,396	2.2	6.3	125
EBR	42.8	2000-01-01	2009-12-31	174,124	5.1	10.8	313
THY	45.8	1991-01-01	2009-12-31	325,160	3.6	9.4	196
HRB	46.7	1997-01-01	2009-12-31	227,878	3.2	8.5	263
NCK	46.7	1993-01-01	2010-12-31	296,295	3.2	8.6	187
FUR	48.1	1990-01-01	2010-12-31	366,751	3.2	8.5	698
CLF	49.5	1981-01-01	2009-12-31	489,913	4.1	10.4	312
BEL	50.1	1984-01-01	2009-12-31	455,756	4.5	12.0	219
NGK	51.6	1995-01-01	2009-12-31	262,003	4.1	11.0	278
HLP	53.1	1995-01-01	2009-12-31	262,649	5.4	15.6	326
BOX	53.4	2004-04-01	2009-12-31	99,753	5.1	14.0	539
HAD	53.5	1983-01-01	2010-12-31	489,931	4.5	11.0	327
WNG	53.9	1981-01-01	2009-12-31	506,287	5.8	16.1	716
BFE	55.2	1982-01-01	2008-12-31	437,388	7.1	20.6	1995
VAL	55.4	1991-01-01	2010-12-31	337,352	3.6	9.2	358
ESK	57.5	1983-01-01	2010-12-31	490,406	6.4	18.4	1271
LOV	57.8	1982-01-01	2004-04-30	391,194	7.3	25.6	2688
NUR	57.8	1991-01-01	2009-12-31	332,609	7.1	26.0	942
UPS	58.2	2003-01-01	2009-12-31	122,722	5.7	17.3	720
LER	61.7	1983-01-01	2009-12-31	473,156	6.4	28.2	992
DOB	61.8	1999-01-01	2009-12-31	191,686	8.0	37.0	659
SOD	63.9	1980-01-01	2009-12-31	525,924	28.2	107.0	1997
ABK	66.0	1979-01-01	2009-12-31	522,409	42.7	154.4	2248
TRO	67.2	1988-01-01	2009-12-31	331,667	39.8	137.6	1050

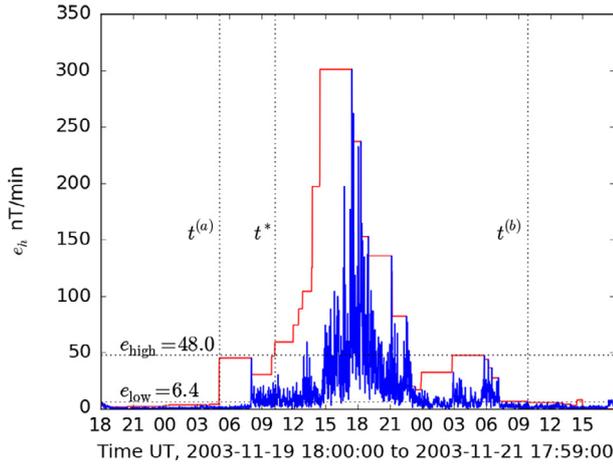
**Table 2.** The table lists for each station (Id) the geomagnetic latitude (Lat), and the 90% and 99% percentiles, and the maximum of  $e_h^{(\infty)}$  for all available data. The computed conductances in log 10(S) are shown using three different integration depths: 40 km (S40), 80 km (S80) and 160 km (S160). The coefficient  $p$  for the linear relation with electric field is also shown in units of (mV/km)/(nT/min). Only data with simultaneous 30-min intervals are chosen. There are  $N = 134,452$  30-min samples covering the period 1999-01-06 to 2008-12-31.

Id	Lat	90%	99%	Max	S40	S80	S160	$p \pm std(p)$
AQU	42.2	3.2	6.7	130	2.1	2.4	3.8	$1.8 \pm 0.1$
SUA	42.2	3.6	10.0	206	3.3	3.6	3.6	$0.6 \pm 0.2$
SPT	42.4	2.8	7.1	125	3.1	3.7	4.1	$0.72 \pm 0.08$
THY	45.8	3.6	10.0	196	2.8	3.4	4.0	$1.1 \pm 0.1$
HRB	46.7	3.6	9.5	263	2.8	3.4	4.0	$1.1 \pm 0.1$
NCK	46.7	3.6	9.8	187	2.8	3.4	4.0	$1.1 \pm 0.1$
FUR	48.1	3.2	8.9	188	1.7	1.7	2.2	$3.8 \pm 0.3$
CLF	49.5	3.6	10.0	206	1.9	2.3	3.1	$2.3 \pm 0.2$
BEL	50.1	4.5	12.7	219	3.1	3.1	3.3	$1.3 \pm 0.3$
NGK	51.6	4.5	12.4	278	3.6	3.6	3.9	$0.6 \pm 0.1$
HLP	53.1	6.1	17.8	326	3.2	3.3	3.4	$0.9 \pm 0.2$
HAD	53.5	4.1	11.4	314	1.9	2.4	2.8	$2.7 \pm 0.3$
WNG	53.9	5.4	15.8	716	3.8	3.9	4.1	$0.36 \pm 0.08$
BFE	55.2	6.3	19.2	628	3.0	3.2	3.3	$1.1 \pm 0.2$
ESK	57.5	6.4	19.6	664	2.4	2.7	2.9	$2.2 \pm 0.2$
NUR	57.8	7.2	26.4	911	2.7	3.4	3.8	$1.12 \pm 0.07$
DOB	61.8	8.5	40.3	659	0.9	1.2	1.5	$5.1 \pm 0.2$
SOD	63.9	28.5	109.8	926	1.7	2.4	2.8	$2.5 \pm 0.1$
ABK	66.0	39.6	141.0	1398	1.7	2.4	2.8	$2.4 \pm 0.1$

consensus on how to automatically identify them (Curto et al. 2007). Instead, we adopt an event-search algorithm as described in the following paragraph.

The first step is to find the first timestamp  $t_1^*$  for which  $e_h$  exceeds a threshold  $e_{high}$ . The timestamps marking the start ( $t_1^{(a)}$ ) and end ( $t_1^{(b)}$ ) of event one are selected as the timestamps before  $t_1^*$  when  $e_h$  is below a threshold  $e_{low}$ . However, as single

1-min  $e_h$  values may fall close to zero even within an active period some temporal filtering should be applied. The autocorrelation of  $e_h$  indicates that there is a quite rapid decay in activity, which is latitude dependent, and has fallen below  $1/e \approx 0.37$  at around 150 min or less. The latitude dependent decay is the result of different dominant processes, at high latitudes the substorms dominate, while at low latitudes the



**Fig. 1.** Illustration of the 3-hour maximum filtering ( $m(t)$ , red curve), thresholds, and the event timestamps  $t_i^*$ ,  $t_i^{(a)}$  and  $t_i^{(b)}$ . The example shows a period in November 2003 for the ESK station (blue curve).

short-lived SI's dominate. Guided by this, temporal maximum filtering with a time window of 3 h = 180 min is applied to  $e_h$ . Therefore, an event is defined as the time interval containing at least one value  $m(t) = \max\{e_h(t), \dots, e_h(t + 180)\}$  exceeding a threshold value  $e_{\text{high}}$  and falling below  $e_{\text{low}}$  outside the event, where  $t$  is in minutes. Thus, the filtering is a 3-h forward maximum. The search algorithm can be summarised as

$$t_i^* = \min \left\{ t; m(t) > e_{\text{high}} \forall t > t_{i-1}^{(b)} \right\},$$

$$t_i^{(a)} = \max \left\{ t; m(t) < e_{\text{low}} \forall t < t_i^* \right\},$$

$$t_i^{(b)} = \min \left\{ t; m(t) < e_{\text{low}} \forall t > t_i^* \right\}, \quad (3)$$

where the subscript  $i$  ( $i \geq 1$ ) refers to event  $i$ ,  $t_i^*$  is the first time the upper threshold is exceeded for event  $i$ , and  $t_i^{(a)}$  and  $t_i^{(b)}$  are the start and end times, respectively. For the first event,  $i = 1$ , the limit  $t_{i-1}^{(b)}$  is set to the first time stamp in the dataset which reduces the first expression to  $t_1^* = \min\{t; m(t) > e_{\text{high}} \forall t\}$ . The start and end times for the first event,  $t_1^{(a)}$  and  $t_1^{(b)}$ , respectively, can then be determined. The other events are then found for  $i > 1$ . The classification of an event is illustrated in Figure 1.

The mid-latitude station ESK is chosen as the reference to define the events. Generally, an event at ESK is accompanied with events at the other locations, although their amplitudes are different. There might also be events that fall below the ESK threshold that would have been selected if another station would be used as reference. Setting the limits to  $e_{\text{low}} = 6.4$  nT/min and  $e_{\text{high}} = 48$  nT/min results in 185 events with an average length of 38 h, ranging from 9 to 96 h. The lower limit and upper limit correspond to the 90% and 99.8% percentiles, respectively.

The target data are based on data from the stations in Table 2. However, the ACE spacecraft further limits the amount of available data due to its shorter temporal coverage and due to data gaps. When only ACE magnetic field data are used the number of available events is reduced to 80. When the density and velocity also are included the number of events decreases to 58. Another aspect is that there are more events

with large  $e_h$  for the ACE magnetic field-only set; e.g. for the magnetic field set there are 19 events with a maximum  $e_h$  above 100 nT/min, while for the plasma set there are only 10 events. The main reason for this is the lack of data from the SWEPAM instrument during strong proton events that in turn are the result of fast CMEs generating the strongest geomagnetic storms.

## 2.5. The geoelectric field

The electric field  $\mathbf{E}$  cannot be directly calculated from  $e_h^{(\infty)}$  using the method by Viljanen et al. (2012). However, an empirical relation can be derived between observed  $e_h^{(\infty)}$  and the calculated 30-min maximum  $|\mathbf{E}_h|$  of the horizontal electric field. We call this quantity  $E_h^{(\infty)}$ .

We applied the EURISGIC ground conductivity model (Adam et al. 2012) to calculate the electric field in 1996–2008 at selected points in Europe using the method by Viljanen et al. (2012). Figure 2 shows the relation between  $e_h^{(\infty)}$  and  $E_h^{(\infty)}$  at four observatories in October 2003. We note that within each 30-min interval the maximum  $|\mathbf{E}_h|$  is not necessarily simultaneous with the maximum  $|\mathbf{dB}_h/\mathbf{dt}|$ . However, there is a clear linear dependence at all sites. Therefore, we express the maximum electric field as

$$E_h^{(\infty)} = p \cdot e_h^{(\infty)}, \quad (4)$$

where  $p$  is an empirical coefficient specific for each site. This coefficient depends on the local ground conductivity model.

The coefficient  $p$  is determined at individual stations for each month using least squares fit. The variation in  $p$  from month to month is rather small, so it is reasonable to use a single value for each site, for which we select the average of monthly values. In Table 2 the coefficients and one standard deviations are shown for the selected stations.

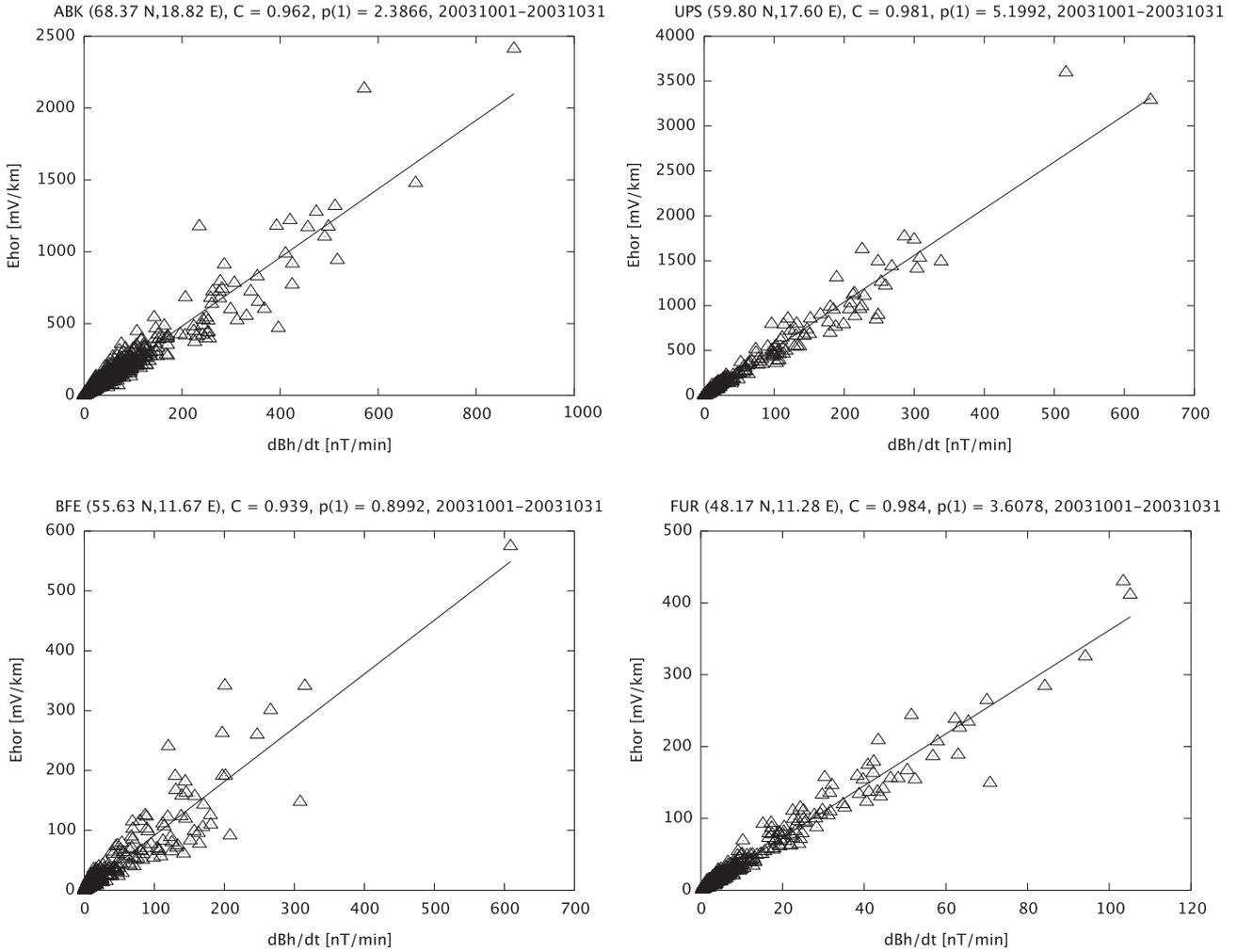
## 2.6. Models

The models mapping from solar wind to  $e_h^{(\infty)}$  consist of Elman neural networks (Elman 1990) and have previously been used for similar problems (Lundstedt et al. 2001; Wintoft et al. 2005). We will not go into the details of the models here, but instead refer to the mentioned papers and references therein. The Elman network can learn not only structures that are spatially related but also structures over time through its internal memory. This is an important feature when the system under study exhibits memory, like the solar wind coupling to geomagnetic storms and substorms. It should also be noted that the temporal resolution of the network memory decreases for more distant features. Therefore, a tapped delay line of the inputs is applied to fully resolve features that are in the near past. The network performs the following mapping

$$\mathbf{u}(t) = (B_y(t), B_z(t), B(t), n(t), V(t)), \quad (5)$$

$$\begin{aligned} y(t) &= f(g_s(t), g_c(t), \mathbf{u}(t), \mathbf{u}(t - \Delta), \dots, \mathbf{u}(t - (n_\Delta - 1)\Delta)) \\ &= f(x_1(t), x_2(t), \dots, x_n(t)), \end{aligned} \quad (6)$$

where  $g_s$  and  $g_c$  are the sine and cosine of local time;  $B_y$ ,  $B_z$ , are the ACE magnetic field  $y$  and  $z$  components in GSM;  $B$  is the magnetic field magnitude; and  $n$  and  $V$  are the plasma density and velocity. The tapped delay line has  $n_\Delta$  terms,



**Fig. 2.** Maximum of the modelled 1-min horizontal electric field versus the maximum of  $|dB_h/dt|$  for 30-min sequences 00:00–00:29, 00:30–00:59, ... UT in October 2003. Top left: ABK, top right: UPS, bottom left: BFE, bottom right: FUR.  $C$  is the linear correlation coefficient,  $p(1)$  is the slope of the fitted linear curve.

each separated with  $\Delta$  minutes. Each input is also labelled as  $x_i(t)$  where the total number of inputs is  $n$ . The choice of values of  $n_\Delta$  and  $\Delta$  is described below. The  $g_s$  and  $g_c$  components are included to capture local time variation. The mapping, with sine and cosine functions, ensures that two instances close in time are numerically close, and that a specific timestamp is uniquely described. Models are also developed without the plasma  $n$  and  $V$ .

The Elman network implements the mapping

$$z_j(t) = \tanh \left( \sum_{i=1}^n w_{j,i} x_i(t) + \sum_{k=1}^m c_{j,k} z_k(t-1) + a_j \right), \quad (7)$$

$$y(t) = \sum_{j=1}^m v_j z_j(t) + b. \quad (8)$$

The free parameters are the number of hidden neurons  $m$ , the weights  $w_{j,i}$ ,  $c_{j,k}$ , and biases  $a_j$  and  $b$ . The optimisation regarding the free parameters are done in the same way as described in Wintoft (2005).

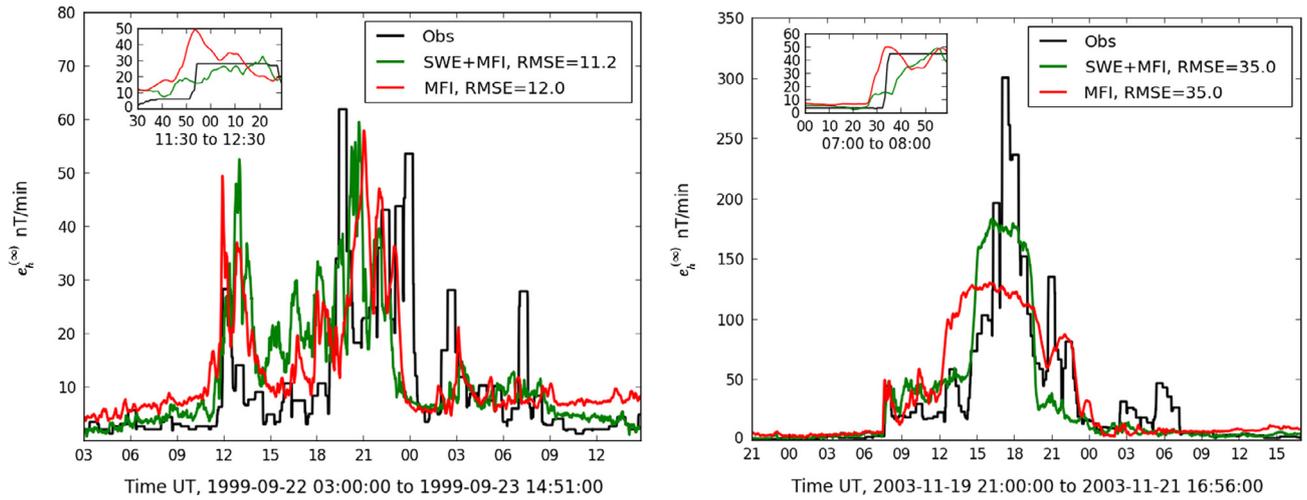
For an average L1-Earth distance of  $1.5 \times 10^6$  km the solar wind bulk flow will have typical travel times ranging from 60 min to 25 min for velocities from 400 km/s to 1000 km/s. However, one should note that for events with

shocks, the bulk velocity and the shock speed are not the same (Viñas & Scudder 1986). There are also extreme cases, like on 29th October 2003 when the solar wind magnetic disturbance (no velocity data were available) preceded the following SI with only 13 min, indicating a shock speed close to 2000 km/s.

The target output for the model is  $e_h^{(\infty)}$  (Eq. (2)). As  $e_h^{(\infty)}$  is a 30-min forward maximum any geomagnetic disturbance at  $t + 30$  min will be reflected at timestamp  $t$  in  $e_h^{(\infty)}(t)$ . For most solar wind velocities a prediction lead time of 30 min is reasonable, therefore, the model output  $y(t)$  is targeted to predict  $e_h^{(\infty)}(t)$ . However, as noted above, longer delays between the solar wind disturbance and the geomagnetic response are also possible and the delay line is set to  $(n_\Delta, \Delta) = (4, 10)$  min, in order to have a detailed description of the solar wind state for the past 30 min. It should also be noted that events with an L1-Earth travel time less than 30 min will be predicted, if the model works, but that the lead time will be shorter, in the extreme cases just 10 min.

## 2.7. Training and validation

The dataset is divided into three sub-sets, each set having similar statistics with respect to  $e_h^{(\infty)}$ . One set is used for training,



**Fig. 3.** Two cases with forecasts using events not in the training set. The black curves are the observed  $e_h^{(\infty)}$ , the green curves the forecast from the SWE + MFI model, and the red curves from the MFI-only model. Note that the two events have very different magnitudes.

i.e. the weights and biases are incrementally adjusted in order to minimise the errors between target and model outputs. A large number of networks with different initial weight values and different numbers of weights are trained. A second set, the validation set, is used to monitor the performance, and the network with the minimum validation error is selected as the optimal network. The third set is used for verification.

### 3. Results

A large number of networks, with different initial weight values and number of hidden units, have been explored through the training-validation process, leading to one optimal network per station. However, there is generally several networks that have mean-square-errors (MSE) that are similar to the MSE of the optimal network. We focus the analysis on the ESK station. We discuss the result for ESK model and explore different approaches to judge the skill of the model.

We start with comparing the model output with observed data for two different events (Fig. 3). For both events, models using solar wind magnetic fields only (MFI), and plasma and magnetic fields (SWE + MFI), are compared.

The first event (1999-09-22) is of moderate magnitude with a peak observed  $e_h^{(\infty)} \approx 60$  nT/min. There is an SI around noon reaching close to 30 nT/min which is followed by the stronger substorm activity in the evening. In this case the MFI-only model quite accurately predicts the time of the SI (see insets in figures) but over-predicts the magnitude by 70%, while the SWE + MFI model shows a slower response, although the forecast is within  $\approx 20\%$  from the observed with a lag of about 10 min. The predicted activity during the afternoon drops slightly but not as much as the observed. The evening and nighttime activity is to some extent captured by the models although individual peaks are offset in time.

The second event (2003-11-19) is much stronger, with peak magnitude of 300 nT/min. The SI in the morning of the 20th reaching close to 50 nT/min is accurately predicted in time and only over-predicted by 10% for the MFI-only model, and slightly under-predicted for the SWE + MFI model. The activity then settles down somewhat and is slightly better traced by the SWE + MFI model compared to the MFI-only model. The onset of activity at 13 UT is probably due to a

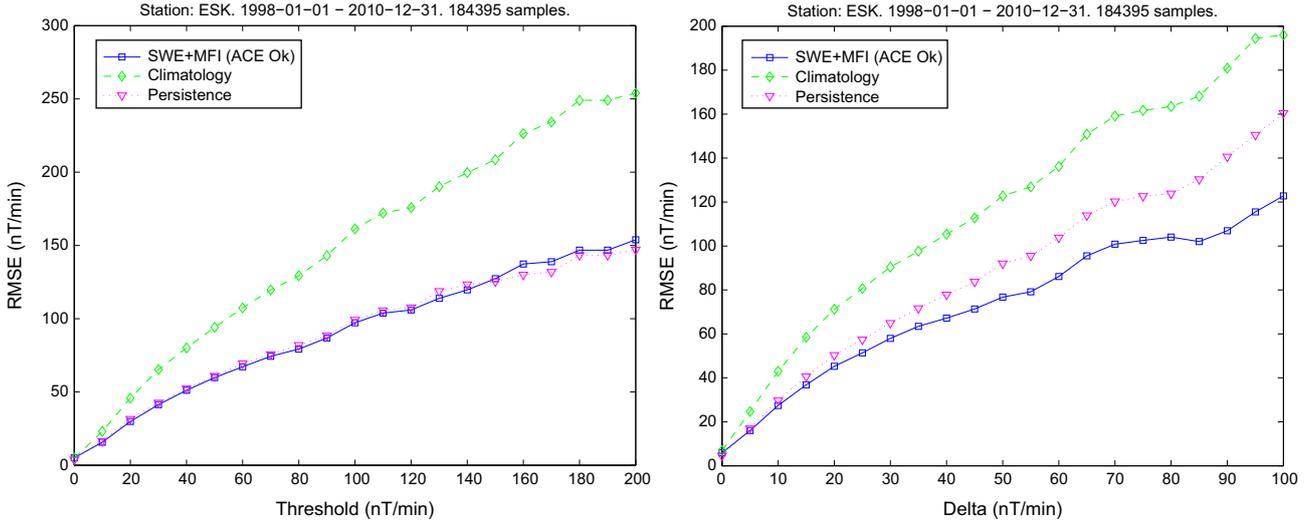
**Table 3.** The RMSE (nT/min) and PE are computed for three different types of models: climatology, persistence and neural network. Each model type is based on either the SWE + MFI set or the combined set. See text for detailed explanation.

Model	SWE + MFI		Combined	
	RMSE	PE	RMSE	PE
Climat.	4.9	0.0	6.0	0.0
Pers.	3.8	0.4	4.3	0.5
N.n.	5.0	-0.1	5.8	0.1

sudden density fluctuation in the solar wind. The SWE + MFI model captures this, although over-forecasting with a factor of two. The MFI-only model also seems to capture this, however, in this case it is a result of the strong solar wind magnetic field. Later in the afternoon/evening the substorm activity sets in which is predicted by both models, although the timing of the SWE + MFI model is better. However, the peak magnitude is under-predicted by a factor of two.

It is interesting to see, for these two cases, that the MFI-only model is able to capture the onset of the SI, as the SI is related to the plasma pressure pulse in the solar wind. However, a solar wind pressure pulse is almost always associated with a sharp increase in the solar wind magnetic field, and this correlation is picked up by the model. Another aspect is that the MFI data are generally of higher quality and with fewer datagaps compared to the plasma data.

We now turn to the problem of verifying the models in the general case. We run the models for all available data for the period January 1998–December 2010. We do not only run the models on the selected events, as there always are some subjective assumptions that will affect the verification. The root-mean-square-error RMSE and the prediction efficiency PE (which is a MSE-based skill score using observed variance as reference model) are computed for the SWE + MFI model and a combined model, shown in Table 3. The combined model uses SWE + MFI when the SWE data exist, and the MFI model when only MFI data exist. For comparison the measures are also computed for two simple models: climatology model, i.e. a constant forecast using the mean value; persistence model, i.e. the next value is equal to



**Fig. 4.** The RMSE as function of threshold ( $\tau$ , left) and as function of positive difference ( $\delta$ , right). The RMSEs are calculated for the SWE + MAG, climatology, and persistence models. Only data for which SWE/PAM data exist have been used.

the current value. Of course, neither of these two models provide any forecast capability. However, the persistence model will often score very high for many measures as the output values have the same statistics as the observed values. Looking at Table 3 it is seen that the persistence model scores best on these measures, and that the neural network model is on par with the climatology model.

In order to explore if there is a gain in using a more complex model over very simplistic models, such as climatology or persistence, or any other model, we study different approaches. Clearly, from the two examples in Figure 3, the neural network model performs better than the climatology model. Further, the persistence model is by its definition not capable of producing a forecast, but it will always trace the observed values and therefore will have small errors.

In Figure 4 we show the RMSE as function of a threshold value ( $\tau$ , left) and positive difference value ( $\delta$ , right), respectively. We define  $\tau$  and  $\delta$  in the following two paragraphs.

In the first case (left figure), only pairs of forecast and observed values are included in the RMSE calculation for which the observed values are above the threshold  $\tau$ , i.e.  $\forall t$  for which  $e_h^{(\infty)}(t) > \tau$ . The points at  $\tau = 0$  corresponds to the values in Table 3. It is seen that the NN and persistence models have similar RMSE ( $\tau$ ), but that the climatology model has more rapidly increasing RMSE. Thus, the NN model now shows a better performance than the climatology model, and similar performance as the persistence model.

In the second case (right figure) pairs are selected when the observed increase is above  $\delta$ , i.e.  $\forall t$  for which  $e_h^{(\infty)}(t) - e_h^{(\infty)}(t-1) > \delta$ . Such a selection will penalise persistence, and any other model that has forecast potential should perform better. The figure shows that indeed the NN model performs better than the persistence model, thus indicating that it actually produces forecasts.

The most interesting events are the extreme events, where extreme can be defined based on different criteria depending on the system at study. However, extreme events are often synonym with rare events. Studying forecasting of dichotomous events, a contingency table can be formed by counting the number of outcomes in the four classes:  $a$  = number of hits;  $b$  = number of false alarms;  $c$  = number of misses;

$d$  = number of correct rejections. The total number of outcomes is  $n = a + b + c + d$ . The event rarity, also known as base rate, then becomes  $p = (a + c)/n$ . Extreme events can thus equivalently be defined as events occurring with small base rate  $p$ . The forecast rate is similarly defined as  $q = (a + b)/n$ . Further, the hit rate and false alarm rate are defined as  $H = a/(a + c)$  and  $F = b/(b + d)$ , respectively. Under certain assumptions on the behaviour of the base rate  $p$  and the forecast rate  $q$ , it can be shown that both  $\log H$  and  $\log F$  are proportional to  $\log p$  and can be expressed in terms of the extremal dependence index (Ferro & Stephenson 2012)

$$\text{EDI} = \frac{\log F - \log H}{\log F + \log H}. \quad (9)$$

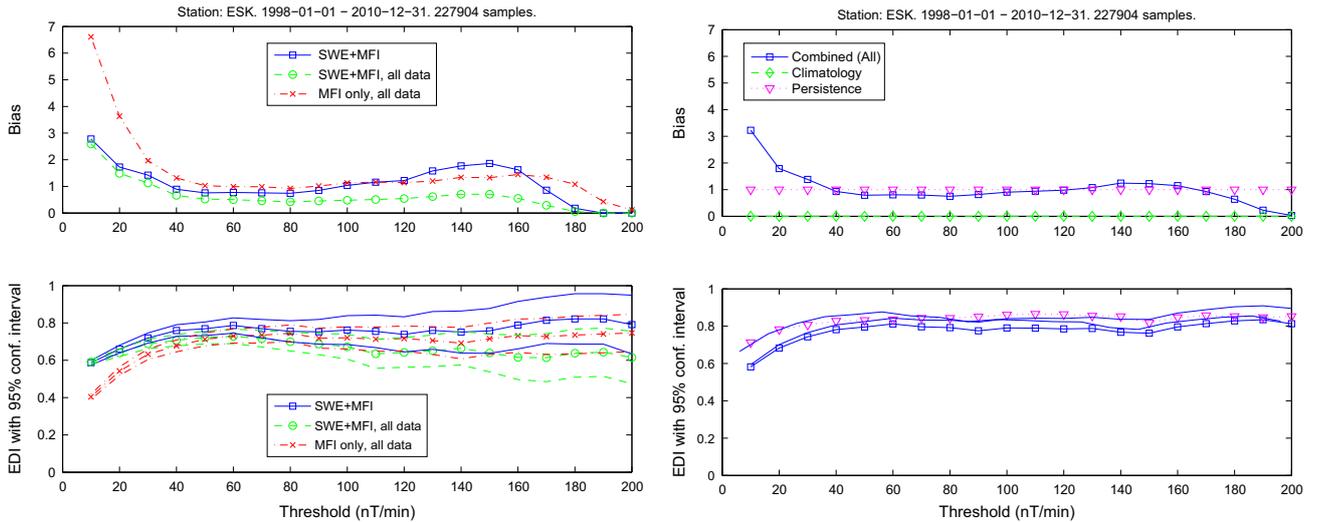
The EDI describes the performance in relation to random forecasts with varying  $p$ . EDI lies in the range  $[-1, +1]$ , where  $\text{EDI} = 0$  corresponds to random forecasts, and  $\text{EDI} > 0$  corresponds to forecasts that are better than random forecasts. Some interesting properties of EDI are that it is base rate independent and asymptotically equitable. The EDI requires that  $p = q$ , which can be achieved by recalibrating  $q$  using different thresholds on the forecasts and observations, respectively. To complement EDI the frequency bias is calculated based on the uncalibrated  $q$  as

$$\text{Bias} = q/p. \quad (10)$$

The Bias indicates whether the model over-forecasts ( $\text{Bias} > 1$ ) or under-forecasts ( $\text{Bias} < 1$ ) at a certain base rate.

Figure 5 shows EDI and Bias for the ESK model as function of observed threshold  $\tau$ . The base rate decreases as the threshold increases. The top left plot shows the Bias for two different models using two different datasets. The ‘‘SWE + MFI’’ label corresponds to the SWE + MFI model evaluated on data for which SWE data exist. The ‘‘SWE + MFI, all data’’ label corresponds to the same model evaluated on all data, including events with missing SWE data during which the forecasts will be zero. Finally, the ‘‘MFI only, all data’’ is the MFI model evaluated on all data.

Generally, all models over-forecast at the lower levels ( $\lesssim 30$  nT/min), e.g. the number of times the ‘‘MFI only’’



**Fig. 5.** The Bias (top row) and EDI (bottom row) computed for the ESK model. Plots at left show the SWE + MFI model evaluated on data for which SWE data exist (SWE + MFI legend) and for all data (SWE + MFI, all data legend), and the MAG only model. Plots at right show the combined model, the climatology model, and the persistence model evaluated for all data. Lines without symbols indicate the 95% confidence levels.

model forecasts values above 20 nT/min is  $\approx 3.6$  times that observed. For a large range ( $40 \lesssim \tau \lesssim 120$  nT/min) the “SWE + MFI” model has  $\text{Bias} \in [0.8, 1.2]$ , then increases to 1.9 at  $\tau = 150$  nT/min, and drops to zero for  $\tau \gtrsim 190$  nT/min. This means that the forecast rate  $q = 0$ , thus there are no forecasts above that threshold. The “SWE + MFI, all data” Bias values are consistently lower, a natural consequence of the zero-forecast for missing SWE data. Finally, the “MFI only” model has Bias closer to one for a large range of thresholds compared to the “SWE + MFI” model, except for small  $\tau$ . The EDIs are similar for both models, within 95% confidence limits, for  $\tau \gtrsim 60$  nT/min. As EDI is asymptotically equitable the models can be compared even though they are based on different datasets.

The EDI stays at a high level of  $\text{EDI} \approx 0.8$  for large  $e_h^{(\infty)}$ . As stated above, the forecasts are re-calibrated to force  $q = p$  before EDI is computed. This is achieved by lowering the forecast threshold which will increase the number of hits ( $a$ ) and false alarms ( $b$ ). If the model would produce random forecasts, then both the hit rate and false alarm rate would increase equally much, with increasing threshold, leading to  $\text{EDI} \approx 0$ . The EDI thus indicates that the model provides actual forecasts but, taken together with the Bias, the forecast values are below observed values at the high end range.

#### 4. Conclusions

Although the transformation of 1-min  $dB/dt$  to 30-min maxima  $e_h^{(\infty)}$  inhibits the calculation of the geoelectric field, it has been shown that there is a reasonably linear relationship between  $e_h^{(\infty)}$  and the 30-min maxima of the computed geoelectric field (Fig. 2). Therefore,  $e_h^{(\infty)}$  can be used as a proxy of the electric field using the  $p$ -coefficients in Table 2. The  $e_h^{(\infty)}$  is simple and well defined, and depends only on the measured  $B$ -field. At the same time, if the electric field model or ground conductivity model are updated, the  $p$ -coefficients can easily be recalculated.

The verification of forecast models is an important and difficult subject. Case studies, such as those in Figure 3, are useful but difficult to summarise. Measures, as given in Table 3,

summarise the verification into a single number, but are hard to interpret. In this special case, we know that the climatology and persistence models cannot provide any forecasts, but still they score better with these simple measures. A more careful selection of data, over which the measures are calculated, can provide further insight to the performance. As Figure 4 shows, the neural network model performs better than the climatology model for large values, and better than the persistence model for large increases.

From Tables 1 and 2 it is seen that the most interesting events are rare. To verify the forecasts of rare events we applied the EDI method. The analysis indicates that the models have a stable EDI around  $\approx 0.8$  for large and rare events (Fig. 5, bottom left). Within the 95% confidence limits it is not possible to decide which model performs best. However, the “MFI” model has a Bias closer to one, and especially it provides forecasts also for the largest values. In principle, a combined model could be used depending on whether the plasma data are reliable or not. For the combined model (Fig. 5, right) the EDI and Bias are very similar to that of the persistence model. Thus, in this case EDI suffers from the same problem as RMSE (Table 3) and RMSE ( $\tau$ ) (Fig. 4, left) in its ability to distinguish between the persistence model and the neural network models.

In a real-time situation it can be difficult to determine when to switch model. Therefore, a real-time forecast system could rely on the “MFI only” model and ignore forecasts below  $\lesssim 30$  nT/min. A related issue for real-time forecasting is that the models generally over-forecast for the low values (Fig. 5, top left), most likely a result of the event selection that is necessary for training.

The models have been implemented for real-time operation and are available at <http://corona.lund.irf.se/eurisig/>.

*Acknowledgements.* The research leading to these results has received funding from the European Community’s Seventh Framework Programme (FP7/2007-2013) under Grant Agreement No. 260330. We thank WDC-Edinburgh and IMAGE, and the contributing institutes, for providing geomagnetic observatory data in Europe, and the ACE SWEPAM and MFI instrument teams and the ACE Science Center for providing ACE data. The editor thanks D. Danskin and an anonymous referee for their assistance in evaluating this paper.

## References

- Adam, A., E. Pracsér, and V. Wetztergom. Estimation of the electric resistivity distribution (EURHOM) in the European lithosphere in the frame of the EURISGIC WP2 project. *Acta Geod. Geoph. Hung.*, **47**, 377–387, 2012.
- Beggan, C., D. Beamish, A. Richards, G. Kelly, and A. Thomson. Prediction of extreme geomagnetically induced currents in the UK high-voltage network. *Space Weather*, **11**, 407–419, 2013.
- Bolduc, L. GIC observations and studies in the Hydro-Québec power system. *J. Atmos. Sol. Terr. Phys.*, **64**, 1793–1802, 2002.
- Boteler, D.H., and R.J. Pirjola. The complex-image method for calculating the magnetic and electric fields produced at the surface of the earth by the auroral electrojet. *Geophys. J. Int.*, **132**, 31–40, 1998.
- Burton, R.K., R.L. McPherron, and C.T. Russell. An empirical relationship between interplanetary conditions and *Dst*. *J. Geophys. Res.*, **80** (31), 4204–4214, 1975.
- Cagniard, L. Basic theory of the magneto-telluric method of geophysical prospecting. *Geophysics*, **18**, 605–635, 1953.
- Curto, J.J., T. Araki, and L.F. Alberca. Evolution of the concept of sudden storm commencements and their operative identification. *Earth Planets Space*, **59**, i–xii, 2007.
- Dong, B., D.W. Danskin, R.J. Pirjola, D.H. Boteler, and Z.Z. Wang. Evaluating the applicability of the finite element method for modelling of geoelectric fields. *Ann. Geophys.*, **31** (10), 1689–1698, 2013, <http://www.ann-geophys.net/31/1689/2013/>.
- Elman, J.L. Finding structure in time. *Cognitive Science*, **14**, 179–211, 1990.
- Ferro, C.A.T., and D.B. Stephenson. *Forecast verification: A practitioner's guide in atmospheric science*, 2nd edn., John Wiley and Sons Ltd, Chapter 10 Deterministic forecasts of extreme events and warnings, 2012.
- Gaunt, C.T. Reducing uncertainty – responses for electricity utilities to severe solar storms. *J. Space Weather Space Clim.*, **4**, A01, 2014.
- Kis, A., A. Koppan, I. Lemperger, T. Prodan, J. Szendrői, J. Verő, and V. Wetztergom. Long-term variation of the geoelectric activity index. *Publs. Inst. Geophys. Pol. Acad. Sc.*, **C-99** (398), 353–360, 2007.
- Lundstedt, H. The sun, space weather and GIC effects in Sweden. *Adv. Space Res.*, **37**, 1182–1191, 2006.
- Lundstedt, H., H. Gleisner, and P. Wintoft. Operational forecasts of the geomagnetic *Dst* index. *Geophys. Res. Lett.*, **106** (A6), 1–4, 2001.
- Mayaud, P.N. *Derivation, meaning, and use of geomagnetic indices*, American Geophysical Union, Vol. 22 of Geophysical monograph, 1980.
- McComas, D.J., S.J. Bame, P. Barker, W.C. Feldman, J.L. Phillips, P. Riley, and J.W. Griffee. Solar wind electron proton alpha monitor (SWEPAM) for the Advanced Composition Explorer. *Space Sci. Rev.*, **86**, 563–612, 1998.
- Pirjola, R., A. Viljanen, A. Pulkkinen, S. Kilpua, and O. Amm. *Ground effects of space weather*, NATO Science Series, Kluwer Academic Publishers, Chapter Space weather effects on electric power transmission grids and pipelines, 2004.
- Pulkkinen, A., A. Viljanen, and R. Pirjola. Estimation of geomagnetically induced current levels from different input data. *Space Weather*, **4**, S08005, 2006.
- Schrijver, C.J., and S.D. Mitchell. Disturbances in the US electric grid associated with geomagnetic activity. *J. Space Weather Space Clim.*, **3**, A19, 2013.
- Smith, C.W., J. L'Heureux, N.F. Ness, M.H. Acuña, L.F. Burlaga, and J. Scheifele. The ACE magnetic fields experiment. *Space Sci. Rev.*, **86**, 611, 1998.
- Stone, E., A. Frandsen, R. Mewaldt, E. Christian, D. Margolies, J. Ormes, and F. Snow. The advanced composition explorer. *Space Sci. Rev.*, **86**, 1–22, 1998, DOI: [10.1023/A:1005082526237](https://doi.org/10.1023/A:1005082526237).
- Trichtchenko, L., and D.H. Boteler. Modeling geomagnetically induced currents using geomagnetic indices and data. *IEEE Trans. Plasma Sci.*, **32** (4), 1459–1467, 2004.
- Viñas, A.F., and J.D. Scudder. Fast and optimal solution to the “Rankine-Hugoniot Problem”. *J. Geophys. Res.*, **91** (A1), 39–58, 1986.
- Viljanen, A., H. Nevanlinna, K. Pajunpää, and A. Pulkkinen. Time derivative of the horizontal geomagnetic field as an activity indicator. *Ann. Geophys.*, **19**, 1108–1118, 2001.
- Viljanen, A., R. Pirjola, M. Wik, A. Ádám, E. Prácsér, Y. Sakharov, and Y. Katkalov. Continental scale modelling of geomagnetically induced currents. *J. Space Weather Space Clim.*, **2**, A17, 2012, DOI: [10.1051/swsc/2012017](https://doi.org/10.1051/swsc/2012017).
- Viljanen, A., A. Pulkkinen, O. Amm, R. Pirjola, T. Korja, and B.W. Group. Fast computation of the geoelectric field using the method of elementary current systems and planar earth models. *Ann. Geophys.*, **22**, 101–113, 2004.
- Weigel, R.S., D. Vassiliadis, and A.J. Klimas. Coupling of the solar wind to temporal fluctuations in ground magnetic fields. *Geophys. Res. Lett.*, **29** (19), 1915, 2002.
- Wik, M., R. Pirjola, H. Lundstedt, A. Viljanen, P. Wintoft, and A. Pulkkinen. Space weather events in July 1982 and October 2003 and the effects of geomagnetically induced currents on Swedish technical systems. *Ann. Geophys.*, **27**, 1775–1787, 2009.
- Wintoft, P. Study of the solar wind coupling to the time difference horizontal geomagnetic field. *Ann. Geophys.*, **23**, 1949–1957, 2005.
- Wintoft, P., M. Wik, H. Lundstedt, and L. Eliasson. Predictions of local ground geomagnetic field fluctuations during the 7–10 November 2004 events studied with solar wind driven models. *Ann. Geophys.*, **23**, 3095–3101, 2005.

**Cite this article as:** Wintoft P., M. Wik & A. Viljanen. Solar wind driven empirical forecast models of the time derivative of the ground magnetic field. *J. Space Weather Space Clim.*, **5**, A7, 2015, DOI: [10.1051/swsc/2015008](https://doi.org/10.1051/swsc/2015008).