Extreme value analysis of the time derivative of the horizontal magnetic field and computed electric field

Peter Wintoft1, Ari Viljanen2, and Magnus Wik1

1Swedish Institute of Space Physics, Scheelevägen 17, 22370 Lund, Sweden
2Finnish Meteorological Institute, Erik Palménin aukio 1, 00560 Helsinki, Finland

Correspondence to: Peter Wintoft (peter@lund.irf.se)

Received: 21 December 2015 – Revised: 11 April 2016 – Accepted: 18 April 2016 – Published: 2 May 2016

Abstract. High-frequency (≈ minutes) variability of ground magnetic fields is caused by ionospheric and magnetospheric processes driven by the changing solar wind. The varying magnetic fields induce electrical fields that cause currents to flow in man-made conductors like power grids and pipelines. Under extreme conditions the geomagnetically induced currents (GIC) may be harmful to the power grids. Increasing our understanding of the extreme events is thus important for solar-terrestrial science and space weather. In this work 1-min resolution of the time derivative of measured local magnetic fields (\(|dB_h/dt|\)) and computed electrical fields (\(E_h\)), for locations in Europe, have been analysed with extreme value analysis (EVA). The EVA results in an estimate of the generalized extreme value probability distribution that is described by three parameters: location, width, and shape. The shape parameter controls the extreme behaviour. The stations cover geomagnetic latitudes from 40 to 70° N. All stations included in the study have contiguous coverage of 18 years or more with 1-min resolution data. As expected, the EVA shows that the higher latitude stations have higher probability of large \(|dB_h/dt|\) and \(|E_h|\) compared to stations further south. However, the EVA also shows that the shape of the distribution changes with magnetic latitude. The high latitudes have distributions that fall off faster to zero than the low latitudes, and upward bounded distributions can not be ruled out. The transition occurs around 59–61° N magnetic latitudes. Thus, the EVA shows that the observed series north of ≈ 60° N have already measured values that are close to the expected maxima values, while stations south of ≈ 60° N will measure larger values in the future.

Keywords. Magnetospheric physics (auroral phenomena; magnetospheric configuration and dynamics; storms and substorms)

1 Introduction

Understanding the extreme conditions for the solar-terrestrial system is a scientific challenge and of importance to the society. Extreme conditions on the Sun result in powerful flares and fast coronal mass ejections (CMEs) (Cliver and Svalgaard, 2004) that may cause large disturbances on the geomagnetic field (Echer et al., 2008) with possible technological effects (Allen et al., 1989; Boteler and van Beek, 1993; Daglis, 2005; Gaunt, 2014).

The term extreme means that something far from the average or normal takes place, thus some level needs to be defined to be able to classify an event as extreme. The usual approach is to look at the distribution of events and set a limit at e.g. 0.1 % percentiles and assuming that extreme events is the same as rare events.

The strongest geomagnetic storms are driven by fast CMEs with strong southward pointing magnetic field (\(B_z\)), and particularly when multiple CMEs arrive at Earth. Fast Earth directed multiple halo CMEs produce the strongest geomagnetic storms as characterized by the Dst index (Gopalswamy et al., 2007).

Geomagnetic indices have an important role for solar-terrestrial studies as they summarize some aspect of geomagnetic activity and they are generally global (Mayaud, 1980). Geomagnetic storms, in the general sense, display different temporal evolution depending on the choice of index. However, common to all storms are that they show a great range of variability in the geomagnetic field from periods less than a minute to several hours (Wintoft, 2005).

The driver of GIC is the geoelectric field (\(E\)) (Pirjola et al., 2004) which in turn is driven by the variations in the geomagnetic field (Viljanen et al., 2004). The high-frequency
components of the geomagnetic field (≤ minutes) are the most important (Pulkkinen et al., 2006), but in cases with a less conductive layer above a highly conductive geological layer in the ground, the low-frequency components will have a large influence (Pirjola, 2010; Pulkkinen et al., 2010). However, for European conditions the magnitude of the calculated $E$-field follows closely the magnitude of $dB/dt$ (Wintoft et al., 2015). Thus, the total variation of the geomagnetic field is not the most important factor, but rather the time derivative $dB/dt$ (Viljanen et al., 2001; Wintoft, 2005). The maximum geoelectric field can also be estimated from forecasts of maximum $dB/dt$ (Wintoft et al., 2015).

The relevance of $E$ or $dB/dt$ to GIC has led to several studies related to their extreme behaviour (Wik et al., 2009; Myllys et al., 2014; Pulkkinen et al., 2015). In Thomson et al. (2011) extreme event analysis (EVA) was performed on $dB/dt$ from stations covering similar latitudes as in this study, and 100- and 200-year return levels were estimated. The return level is obtained by extrapolating the data-fitted distribution to a small probability $p$ and interpreting that the corresponding level will be exceeded once every $1/p$ years. The 100-year estimated return levels were found to lie in the range 1000–2000 nT min$^{-1}$ for geomagnetic latitudes above 60° N and typically below 1000 nT min$^{-1}$ south of 53° N. However, in the intermediate latitudes large values of 4000–5000 nT min$^{-1}$ were estimated. Sharp increases in amplitude of both $dB/dt$ and calculated $E$ around 50° N geomagnetic latitude were found by Pulkkinen et al. (2012) using data for the great storms of 13–15 March 1989 and 29–31 October 2003.

In this work 1-min resolution geomagnetic data and modelled electric field have been analysed using the generalized extreme value (GEV) distribution. As pointed out in Thomson et al. (2011) the generalized Pareto distribution (GPD) is appropriate when complete high-resolution measurement series are available as more data are used for the estimates. In GEV only one value is used out of the 500 000 values in the year, but in GPD all values above a threshold are used. However, a threshold needs to be defined in the GPD and care need also to be taken for clustering so that only a single value from an event is included in the estimate. So in this respect the GEV is more general but at the same time fewer data points are considered. Another point is that the GEV used in this study forms an independent analysis against the GPD. Furthermore, we also apply the GEV analysis to the calculated electrical field at each site with measured magnetic fields.

2 Analysis

2.1 Geomagnetic data and calculated electric field

Geomagnetic data, at 1 min resolution, are obtained from World Data Centre for Geomagnetism (Edinburgh; http://www.wdc.bgs.ac.uk/catalog/master.html). The 1-min differences are formed as an approximation to the time derivative ($dB_x/dt$, $dB_y/dt$, $dB_z/dt$) in the geographic $X$–$Y$–$Z$ coordinate system, where $X$ points north, $Y$ east, and $Z$ vertically down. The 15 stations included are shown on the map in Fig. 1. The criterium to include a station is that there should be at least 18 years of contiguous 1-min data and the number of years of available data is shown in parenthesis. The red curves mark the geomagnetic latitudes, using the 11th generation IGRF coefficients, for epoch 1 January 2000.

![Figure 1](https://www.wdc.bgs.ac.uk/catalog/master.html) The locations of the magnetic observatories used. Each observatory has at least 18 years of contiguous 1-min data and the number of years of available data is shown in parenthesis. The red curves mark the geomagnetic latitudes, using the 11th generation IGRF coefficients, for epoch 1 January 2000.
2.2 Extreme value analysis – background

We follow the approach described in Coles (2001) for our analysis. A time series \( x(t) \) is divided into blocks of equal duration and then the maximum in each block is determined thereby producing a new series with block-maxima \( z(t) \). If the number of values within each block-maxima is large, independent, and identically distributed then \( z(t) \) will be distributed according to the generalized extreme value distribution

\[
G(z) = e^{-\left[1 + \xi \left( \frac{z - \mu}{\sigma} \right) \right]^{-1/\xi}},
\]

where the probability that \( z < z^* \) for some \( z^* \) is \( G(z^*) \), i.e. \( \Pr(z < z^*) = G(z^*) \). The parameters \( \mu, \sigma, \) and \( \xi \) are known as the location, width, and shape, respectively. The corresponding probability density function (PDF) is

\[
g(z) = \left[ 1 + \xi \left( \frac{z - \mu}{\sigma} \right) \right]^{-1/(1+\xi)} e^{-\left[1 + \xi \left( \frac{z - \mu}{\sigma} \right) \right]^{-1/\xi}},
\]

when \( 1 + \xi \left( \frac{z - \mu}{\sigma} \right) > 0 \), else \( g(z) = 0 \). In the special case with \( \xi = 0 \) the PDF is

\[
g_0(z) = \frac{1}{\sigma} e^{-\frac{z - \mu}{\sigma}}.
\]

The shape \( \xi \) is of special interest. If \( \xi < 0 \), the PDF is zero for \( z \geq \mu - \sigma/\xi \), i.e. \( z \) has an upper limit. Note that \( \mu \) and \( \sigma \) are positive, therefore \( \mu - \sigma/\xi \) is also positive for negative \( \xi \). If instead \( \xi \geq 0 \), then \( z \) has no upper limit, although \( g(z) \to 0 \) and \( g_0(z) \to 0 \) when \( z \to \infty \). The larger the \( \xi \), the slower \( g(z) \) falls off to zero.

2.3 Extreme value analysis – applied to \( dB/dt \) and \( E \)

From the observed horizontal magnetic field \( \{B_x(t; s), B_y(t; s)\} \) at time \( t \) and location \( s \), the magnitude of its derivative is defined as

\[
x_h^{(s)}(t) = \sqrt{\left(dB_x(t; s)/dt\right)^2 + \left(dB_y(t; s)/dt\right)^2}.
\]

As we are interested in the magnitude of the horizontal electric field, \( |E| = \sqrt{E_x^2 + E_y^2} \), and that to a first approximation \( (E_x, E_y) \propto (dB_y/dt, -dB_x/dt) \) we use the expression

\[
x_e^{(s)}(t) = \sqrt{\frac{dB_y(t; s)}{dt}^2 + \frac{dB_x(t; s)}{dt}^2}.
\]

Similarly the magnitude of the calculated horizontal electric field vector \( \{E_x(t; s), E_y(t; s)\} \) is

\[
x_e^{(s)}(t) = \sqrt{E_x^2(t; s) + E_y^2(t; s)}.
\]

Block-maxima \( z_h^{(s)}(t) \) and \( z_e^{(s)}(t) \) are calculated from \( x_h^{(s)}(t) \) and \( x_e^{(s)}(t) \), respectively. The length of each block should be long enough to contain many samples, but at the same time the block-maxima series \( z \) should also be as long as possible for the statistical interpretation. From a theoretical viewpoint the observations within a block should be independent and identically distributed (i.i.d.). Clearly, it is difficult to show that i.i.d. holds in the strict sense, however, we believe it is a reasonable assumption according to the following. We choose yearly block-maxima as they contain a large number of values. Although 1 year consists of about 526,000 1-min values, they are not independent. Considering individual substorms one may argue independence on timescales of hours. However, substorms may also be part of a storm and then the timescales are of the order of hours to a few days. Thus, each yearly block maxima contains at least \( \sim 100 \) days timescale) to thousands (hours timescale) of independent values. Finally, we assume that they are independently distributed (i.d.), which naturally can be questioned. However, the i.d. assumption is to some degree supported by the result by Weigel and Baker (2003) who found that the form of the distribution of large \( dB/dt \) for a high-latitude station (SOD) was independent of the solar wind forcing, time of day, and day of year.

The three parameters \( \mu, \sigma, \) and \( \xi \) are found by maximizing the log-likelihood function using the block-maxima series \( z(t) \). The results are shown for each station in Table 1. As expected, the location \( \mu \) and width \( \sigma \) generally increase with increasing magnetic latitude. However, the shape \( \xi \) is generally smaller, and even negative, at high magnetic latitudes. The estimated distributions for the stations TRO, ABK, ESK, and FUR are shown in Fig. 2. Notice that the estimated distribution at TRO has an upper limit of the block-maxima \( z \) due to the negative \( \xi \). It is also seen that the ABK distribution falls off faster then the ESK and FUR distributions.

As the shape parameter \( \xi \) is especially interesting for the extreme behaviour the parameter interval at 95 % confidence level is estimated. As a first approximation it can be assumed that \( \xi \) is normally distributed leading to a straightforward es-
Table 1. The magnetic latitude, number of years (N), maximum of block-maxima z (nT min^{-1}), the estimated parameters \( \mu \) (nT min^{-1}), \( \sigma \) (nT min^{-1}), and \( \xi \) for each analysed station. Note that UPV is the combined series from the LOV and UPS stations. The magnetic latitudes have been calculated using the 11th generation IGRF coefficients for year 2000.

<table>
<thead>
<tr>
<th>Station</th>
<th>M. lat.</th>
<th>N</th>
<th>dB/dt (nT min^{-1})</th>
<th>E (mV km^{-1})</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>maxz ( \mu ) ( \sigma ) ( \xi )</td>
<td>maxz ( \mu ) ( \sigma ) ( \xi )</td>
</tr>
<tr>
<td>TRO</td>
<td>67.0</td>
<td>22</td>
<td>1050 589 174 -0.26</td>
<td>1994 1398 383 -0.60</td>
</tr>
<tr>
<td>ABK</td>
<td>65.9</td>
<td>32</td>
<td>2248 640 272 0.12</td>
<td>4081 1409 570 0.09</td>
</tr>
<tr>
<td>SOD</td>
<td>63.7</td>
<td>31</td>
<td>1997 462 192 0.18</td>
<td>3297 1082 419 0.11</td>
</tr>
<tr>
<td>LER</td>
<td>61.9</td>
<td>30</td>
<td>992 204 142 0.05</td>
<td>2430 530 361 0.12</td>
</tr>
<tr>
<td>UPV</td>
<td>58.0</td>
<td>30</td>
<td>2688 133 133 0.69</td>
<td>1938 160 153 0.60</td>
</tr>
<tr>
<td>ESK</td>
<td>57.7</td>
<td>30</td>
<td>1271 81 62 0.97</td>
<td>2412 174 126 0.80</td>
</tr>
<tr>
<td>VAL</td>
<td>55.7</td>
<td>19</td>
<td>358 34 23 0.70</td>
<td>645 89 53 0.63</td>
</tr>
<tr>
<td>BFE</td>
<td>55.3</td>
<td>25</td>
<td>1995 96 95 1.10</td>
<td>1562 121 95 0.82</td>
</tr>
<tr>
<td>WNG</td>
<td>54.0</td>
<td>32</td>
<td>716 61 50 0.90</td>
<td>172 23 15 0.54</td>
</tr>
<tr>
<td>HAD</td>
<td>53.8</td>
<td>30</td>
<td>327 40 26 0.91</td>
<td>703 121 67 0.46</td>
</tr>
<tr>
<td>HLP</td>
<td>53.0</td>
<td>18</td>
<td>326 56 37 0.73</td>
<td>198 52 29 0.31</td>
</tr>
<tr>
<td>DOU</td>
<td>51.3</td>
<td>18</td>
<td>261 35 21 0.69</td>
<td>416 64 36 0.74</td>
</tr>
<tr>
<td>BEL</td>
<td>50.0</td>
<td>29</td>
<td>219 48 31 0.57</td>
<td>277 59 30 0.39</td>
</tr>
<tr>
<td>CLF</td>
<td>49.7</td>
<td>29</td>
<td>312 37 24 0.62</td>
<td>480 85 49 0.54</td>
</tr>
<tr>
<td>FUR</td>
<td>48.2</td>
<td>21</td>
<td>196 32 21 0.65</td>
<td>703 121 76 0.58</td>
</tr>
</tbody>
</table>

The maximum of \( \ell \), denoted by \( \ell^* \), is obtained for each station by the parameters in Table 1. For the ESK station the maximum is \( \ell^* = -186.9 \) as seen in Fig. 3. The profile of the log-likelihood function is achieved by iteratively changing \( \xi \) away from its optimal value by small amounts and finding new optimal values for \( \mu \) and \( \sigma \) for which \( \ell \) is locally maximized. The 95 % level is reached when \( 2(\ell^* - \ell(\mu, \sigma, \xi)) < c_{0.05} \), where \( c_{0.05} \) is the 1 - 0.05 = 0.95 quantile of the \( \chi^2_1 \) distribution (Coles, 2001). Figure 3 shows the profile for the ESK station, and the limits obtained both from the profile and from assuming a normal distribution. It is seen that the profile limits are asymmetric around the log-likelihood maximum and shifted towards larger values as compared to the normal limits.

The shapes \( \xi \), determined from \( |dB/dt| \) and \( |E| \), at the different locations are shown in Fig. 4 with respect to the magnetic latitude. The profile estimates of the confidence limits at 95 % confidence level are indicated with the vertical bars. All stations below magnetic 58° N have positive values for \( \xi \) determined from \( |dB/dt| \) at the 95 % confidence level, while stations above 61° N contain both positive and negative \( \xi \). The variation of the widths of the confidence limits is caused by a couple of factors. The length of each series has some influence, however, in this case not large, as the number of samples ranges from 18 to 32 and the confidence limits are approximately inversely proportional to the square root of the sample size. There is a also a “natural” variation as samples are randomly drawn from the underlying distribution. Finally, samples from a distribution with positive \( \xi \) will
be more spread out than samples coming from a distribution with $\xi$ around zero, see e.g. Fig. 2. Therefore, there will generally be a larger uncertainty associated with positive $\xi$ than with $\xi$ around zero, hence the narrower confidence intervals at higher latitudes.

The $\xi$ parameter estimated from the calculated electric field shows a similar variation with latitude. Stations above $\sim 60^\circ$ N lie around $\xi = 0$, while stations south of that have positive $\xi$, with one exception, namely the HLP station. However, it should be noted that there are only 18 years of data from the HLP station (Fig. 1).

Although all stations selected have data extending over 18 years or more, they cover slightly different time intervals. For example, the UPV station includes the year 1982 with an extreme of 2688 nT min$^{-1}$, while some of the other stations do not include that year. To test the effect on the estimate of the distribution parameters we remove the data for that year and redo the analysis for UPV. The result yields $(\mu, \sigma, \xi) = (137, 125, 0.49)$ which is similar to that in Table 1. The new confidence limits also puts $\xi$ above zero, with lower and upper 95% confidence limits of $\xi$ equal to (0.09, 1.19).

From the distributions the 100-year return levels can be estimated. We find that the return levels are similar in magnitude and latitudinal variation as found by Thomson et al. (2011). However, in this analysis the 95% confidence limits are much larger making it hard to come to a strong conclusion.

We also turned the question around and asked what is the return period for some high level. Selecting a value of 3000 nT min$^{-1}$ it is above all observed maxima (Table 1) but not to a great extent from that observed at ABK and UPV. The fitted distribution at TRO will never reach 3000 nT min$^{-1}$ as $\xi$ is less than zero. The ABK, SOD, and LER stations have an average $\bar{\xi} = 0.1$ and their corresponding 3000 nT min$^{-1}$ return periods are much larger than the lengths of the corresponding measurements series as shown in Table 2. The mid-latitude stations (UPV, ESK, BFE, WNG) have $\bar{\xi} = 0.9$ and 3000 nT min$^{-1}$ return periods in the range 30–80 years, except the VAL station. The lower latitude stations (HAD, HLP, DOU, BEL, CLF, FUR) have $\bar{\xi} = 0.7$ and return periods in the range 200–1000 years. The VAL station has a very low maximum $|dB/dt|$, in accordance with Thomson et al. (2011), in that both predicted and observed extremes are significantly less compared to stations at similar magnetic latitudes. However, the $\xi$ parameter is still positive at the 95% significance level.

Clearly, with return periods largely exceeding the lengths of the analysed time series the uncertainty is very large. E.g. the LER station, with 10$^6$ year return period at 3000 nT min$^{-1}$, is caused by a small $\xi$ and that the observed maximum of 992 nT min$^{-1}$ is far from the 3000 nT min$^{-1}$ limit. In particular, the LER 3000 nT min$^{-1}$ return period can be brought down to below 800 years by changing its $\xi$ within the 95% confidence limits. This is also in line with the 200-year return levels shown in Fig. 6 by Thomson et al. (2011) when comparing with ABK and SOD.

Further, we may also relate the return periods to limits determined from the individual station maxima in Table 2. The high latitude stations have return periods longer than the extent of the measured series for reaching the currently observed maxima. The more southerly stations generally have return periods similar to or shorter than the observed series lengths.

We may also explore the effect on $\xi$ on the chosen conductivity model. From Table 1 it is seen that $\xi$ for all stations,
except two, have lower values for $|E|$ than for $|dB/dt|$, i.e. $\xi_E < \xi dB/dt$. A simple frequency-independent scaling is not sufficient to change the values of $\xi$. This is an effect of that in $dB/dt$ only the high-frequency components of $B$ are considered, while in $E$ there is also some contribution from the lower-frequency components in $B$. To test this we chose the SOD station and changed the conductivity of the deep layer. The original model has two layers: a top layer of 30 km depth with resistivity 5000 $\Omega$ m and an infinitely deep bottom layer with 200 $\Omega$ m (Adam et al., 2012). The bottom layer resistivity is then changed to 20 $\Omega$ m (same as UPV) and 2 $\Omega$ m and $\xi$ recalculated. The damping of the high-frequency components results in that $\xi_E$ is decreased from 0.11 to 0.099 and 0.091, respectively.

3 Conclusions

Extreme value analysis (EVA) has been applied to 1-min resolution time series of locally measured $dB/dt$ for 14 stations in Europe (Fig. 1), and for locally calculated $E$. The higher latitude stations have on average higher levels of disturbance as shown by the location ($\mu$) and width ($\sigma$) (Table 1). This is expected due to the stations proximity to the auroral oval.

For the shape parameter $\xi$ the results show that, at 95 % confidence level, stations south of $\approx 60^\circ$ N have positive shape parameter $\xi$, indicating upwardly unbounded distributions. On the other hand, stations north of $\approx 60^\circ$ N have $\xi$ values close to zero, and negative values can not be ruled out. If the results hold, it has implications on what levels of disturbance can be achieved in the extreme cases. Assuming that $\xi$ is negative for the northern stations (TRO, ABK, SOD, LER) then the maximum $|dB/dt|$ or $|E|$ have upper limits and values not much higher than what has already been observed are anticipated (see maximum values in Table 1). For the other stations the $\xi$ parameters are larger and definitively positive and thus the maximum observed values will be exceeded in the future. But even if it is assumed that the northerly stations have positive $\xi$, the distribution still falls off quicker than for more lower latitude stations (Fig. 2).

From Table 1 it is seen that $\xi$ for all stations, except two, have lower values for $|E|$ than for $|dB/dt|$, i.e. $\xi_E < \xi dB/dt$. an effect of that in $dB/dt$ only the high-frequency components of $B$ are considered, while in $E$ there is also some contribution from the lower-frequency components in $B$. This was also confirmed at one location (SOD). In future work it would be interesting to further explore the $\xi$ dependency on the conductivity model, and also especially try to understand why $\xi_E$ is larger than $\xi dB/dt$ for some locations.

The return periods for a 3000 nT min$^{-1}$ disturbance were computed from the fitted distributions, except for TRO which can never reach that level. The three high latitude stations showed return periods of several hundred years or more, similar to the most southerly stations, while the mid-latitude stations showed 30–80 years return periods. It is generally stated that geomagnetic variations are most intense at high latitudes. This is true when considering, for example, long-term averages of field variations. However, it is does not automatically follow that the same holds for extreme events. As observations of the July 1982, March 1989 and October 2003 storms indicate, magnetic variations and $dB/dt$ reached values at subauroral latitudes that were comparable or even larger than at the average auroral region (Kappenman, 2005, 2006; Pulkkinen et al., 2005; Wik et al., 2009). Observations also show that ionospheric electrojets move southwards when their intensity increases (Ahn et al., 2005; Rostoker and Duc Phan, 1986). Thus, the physical interpretation could be as follows. During moderate-size storms the auroral oval extends over higher latitudes and therefore the high latitude stations will observe larger $dB/dt$ compared to stations at lower latitudes. However, during the extreme storms the oval will move further southward with the result that the more extreme $dB/dt$ will be observed further south.

Acknowledgements. This work has been funded by the following grants: European Community’s Seventh Framework Programme (FP7/2007–2013) under Grant Agreement No. 260330; and Swedish Civil Contingencies Agency, contract MSB-2011-4186. We thank WDC-Edinburgh and IMAGE, and the contributing institutes, for providing geomagnetic observatory data in Europe. We would also like to thank the reviewers (J. M. Torta and an anonymous) for improving the manuscript with valuable comments.

The topical editor, G. Balasis, thanks J. M. Torta and one anonymous referee for help in evaluating this paper.

References


Daglis, I. A.: Effects of Space Weather on Technology Infrastructure, NATO Science Series II: Mathematics, Physics and Chem-


