British Geological Survey
NATURAL ENVIRONMENT RESEARCH COUNCIL

Geomagnetic Extreme Statistics for Europe

Alan Thomson (awpt@bgs.ac.uk), Sarah J Reay and Ewan Dawson
British Geological Survey, West Mains Road, Edinburgh EH9 3LA, United Kingdom

Abstract

Rapidly changing geomagnetic field variations constitute a natural hazard, for example to grounded power grids and pipeline networks. To understand this hazard we have continuous magnetic measurements across the world for typically less than 100 years. Much of the older data is also in analogue form, or is only available digitally as hourly or daily magnetic indices or mean levels. So it may not yet be clear what the true extremes in geomagnetic variation are, particularly on time scales - seconds to minutes - that are relevant for estimating the hazard to technological systems.

We therefore use a number of decades of one minute samples of magnetic data from observatories in Europe, together with the technique of 'extreme value statistics', to explore estimated maxima in field variations in the horizontal strength and in the declination of the field. These maxima are expressed, for example, in terms of the variations that might be observed on time scales of 100 and 200 years. We also examine the extremes in the one-minute rate of change of these field components on similar time scales.

The results should find application in hazard assessment and navigation applications.

Extreme Value Statistics - Theory

For our analysis we use a *Generalised Pareto Distribution* (GPD) to describe the tail of the distribution of geomagnetic activity (see e.g. Coles, 2004). The GPD is a unifying description of the Gumbel, Frechet and Weibul distributions, widely used in the scientific literature when examining extremes. These three distributions can be shown firstly to be combined in a single *Generalised Extreme Value* (GEV) function of three parameters, describing the location (L), scale (S) and shape (P) of the distribution. GEV statistics are commonly used where only block maxima are available, e.g. annual maxima of daily temperatures.

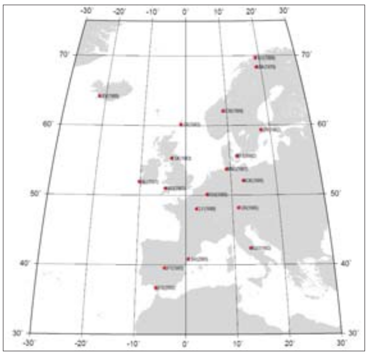
The GPD is more general still and is applicable to our data, as we have individual one-minute samples and some idea of an appropriate threshold of extreme activity. This is known as the 'point over threshold' approach. One example of a relevant geomagnetic threshold is the Space Weather Prediction Centre's 'severe storm', defined for Ap>100.

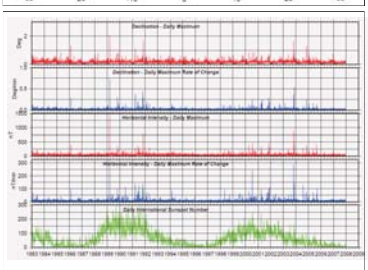
The GPD function $\mathbf{G(x;S',P,u) = 1 - [1+P(x-u)/S']^{(-1/P)}}$ where x-u>0, 1+P(x-u)/S'>0 and S'=S+P(u-L), is given in terms of equivalent parameters, L, S, P, from the GEV distribution.

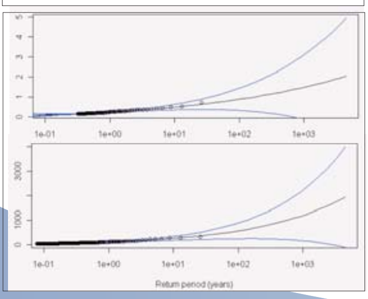
G gives the probability of the random variable (here a field variation or residual), X, exceeding a high value, x, given that it already exceeds a high threshold, u, i.e. Pr[X>x|X>u].

There are subtleties in applying extreme value statistics to geomagnetic data, e.g. the need to de-cluster sequences of magnetic data (close-in-time storms or sub-storms following from one or more related coronal mass ejections) and, in general, the non-stationarity of geomagnetic data. Though these can be dealt with, to leave stationary, independent random estimates for GPD analysis, these points are not considered further in this preliminary work.

Extreme Value Statistics - Data & Methods







One minute geomagnetic data in H (horizontal field), D (declination), dH/dt and dD/dt (the latter two computed as one-minute differences) were obtained from the Edinburgh World Data Centre (www.wdc.bgs.ac.uk), for 19 European observatories (Figure 1).

These observatories were chosen to provide a representative spread of measurement sites across the continent, to include a range of magnetic latitudes and for which there are continuous data over a number of years.

Figure 1. European geomagnetic observatory data used in the analysis, annotated with the start year (from 1st January) of one-minute data. End date is either (December 31) 2007 or 2008, depending on observatory.

Variations, i.e. residual data for each observatory, representing the external field only, were constructed by removing a quiet mean level, established for each month from the five 'International Quiet Days', as determined by the International Service of Geomagnetic Indices. Absolute values of these variations were then computed (Figure 2).

Figure 2. H (nT), D (degrees) residuals and one-minute rates-of-change for Hartland mid-latitude observatory (one minute data, 1983-2008). Daily maxima of the absolute residuals are shown, as is sunspot number.

The absolute variation data were analysed using the *eXtremes* software toolkit (Gilleland and Katz, 2005) that runs on the *R* statistical analysis package (R Development Core Team, 2008). Maximum one-minute values observed per 15-minute time block were used as our basic data set, providing a manageable reduction in database size, whilst permitting a reasonably 'fine-grained' analysis.

Component maxima were determined both within the time-span of data and from the projected GPD distribution, for periods of 100 and 200 years. 95% confidence levels were also determined. An example of an output from the *eXtremes* toolkit is given in Figure 3.

Figure 3. Return periods for observed Hartland dD/dt (upper) and dH/dt (lower) residuals (circles) and the fitted and extrapolated GPD (line) to each of dD /dt and dH/dt. Vertical scales are deg/min and nT/min respectively; horizontal scale is time in years. Blue lines are the estimated +/-95% confidence limits from the fit of model to data.

Extreme Value Statistics - Results

For each observatory and each component we have extracted the peak variation (i.e. residual) and rate-of-change predicted via the GPD to be exceeded every return period of 100 and 200 years. To do this appropriate geomagnetic activity thresholds (u) were determined for each observatory, these thresholds increasing with geomagnetic latitude and typically around 10% of the maximum observed. The results are summarised in Figure 4.

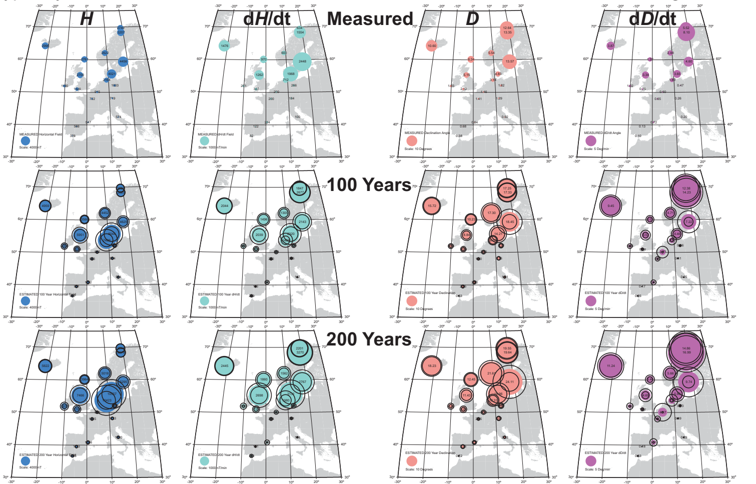


Figure 4. Top/Middle/Bottom - respectively the measured/100-year-return-level/200-year-return-level, for H (far left column), D (second column), dH/dt (3rd column) and dD/dt (right/4th column). Units are nT, nT/min, degrees, degrees/min (left to right). 95% confidence levels are the inner and outer rings around each coloured mean level.

Return levels are of course strongly influenced by the measured levels found in each observatory data set. There is, predictably, a dependence on geomagnetic latitude but also some 'regional structure' found around the North Sea, particularly in H and dH/dt.

Does the Carrington Storm Fit?

The Carrington magnetic storm of 1st-5th September 1859 is not included in our data set. Therefore we can ask: "Do geomagnetic data measured during the Carrington storm support the analysis presented here?". As an event, one *might* expect that Carrington levels of activity should 'sit' between the 100 and 200 year return levels given in Figure 4. That is, the Carrington event has not been exceeded in the 150 years since it occurred, so that the estimated return level for such an event is at least 150 years. It could be much more.

In Figure 5 we show paper magnetograms from the Greenwich (London) observatory recorded during the Carrington event (see poster V02-0715, by Clarke *et al*, for more information on a project to scan and ultimately digitise British observatory paper records).

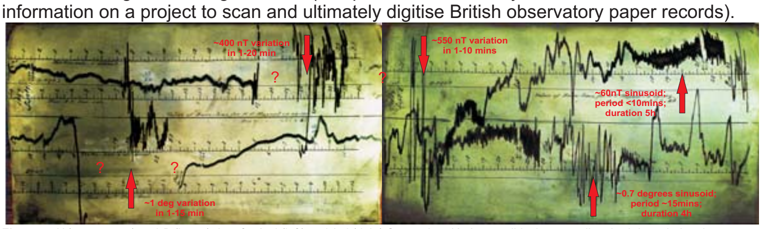


Figure 5. H (upper trace) and D (lower) data for 2nd (left) and 3rd (right) September 1859 - possibly the most disturbed days during the event for which a substantial photographic record survives. Annotation is used to highlight interesting episodes of activity. Rates-of-change prove easiest to estimate: absolute variations await future baseline level confirmation. Major gaps in the record are indicated by '?'.

From examining the Carrington storm magnetograms we tentatively estimate that **measurable** variations were no more than about **500 nT/min in d*H*/dt** and **1.0 deg/min in d*D*/dt**, *assuming* that changes occurred on time scales around one minute, rather than the 5-10 minutes that we estimate is the best that can be inferred from these photographs.

From Figure 4 and interpolating between Hartland and Chambon-la-Foret observatories, one would have anticipated **170(200) < d*H*/dt <700(900) nT/min**, and **0.6(0.7) <d*D*/dt <2.0(2.7) deg/min** in **100 (200) years**. The Carrington data therefore provide modest support, given the assumptions, to the extrapolation of the GPD fit. We note that there were many periods during the full event for which the data have not survived (off-scale, paper & ink degradation, etc). It is therefore very likely that greater variations did, in fact, occur.

Conclusions and Future Work

The results of this rather preliminary analysis are summarised in Figure 4. We find that predicted return magnitudes increase with geomagnetic latitude, although there is some other, as yet unexplained, structure in the data found around the North Sea.

The Carrington data are interesting, but hardly conclusive in support of our analysis, insofar as many data are off-scale and are lost. There is also some uncertainty over the time resolution we can infer and the instrument's likely sensitivity to the fastest changes (magnetometer frequency response and mid-1800s photographic paper sensitivity).

Future work will involve checking the assumptions made, e.g. the appropriateness of the GPD for extreme geomagnetic data, 'block-averaging' versus 'point-over-threshold' statistics, de-clustering of storm data (Coles, 2004) and extending the analysis to include other observatories, in order to provide a more global view of the extreme geomagnetic hazard to technology and to navigation.

Acknowledgements

The authors would like to acknowledge the provision of geomagnetic data by scientific institutes in Europe and the added value provided by INTERMAGNET and the World Data Centres. Colleagues at BGS are also thanked for their comments on this work.

References

Coles, S., 2004 *An Introduction to Statistical Modeling of Extreme Values*, Springer-Verlag, London, ISBN 1-852334-59-2.

Gilleland, E and Katz, R. W., 2005. *Tutorial for the Extremes Toolkit: Weather and Climate Applications of Extreme Value Statistics*, http://www.assessment.ucar.edu/toolkit.

R Development Core Team, 2008. *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0, www.R-project.org.

IAGA 11th Scientific Assembly
Sopron, Hungary August 27th, 2009
Symposium: V04 Poster: 0567

IAGA 2009 SOPRON HUNGARY