Quantifying extreme behavior in geomagnetic activity

Alan W. P. Thomson,[1] Ewan B. Dawson,[1] and Sarah J. Reay[1]

[1] Understanding the extremes in geomagnetic activity is an important component in understanding just how severe conditions can become in the terrestrial space environment. Extreme activity also has consequences for technological systems. On the ground, extreme geomagnetic behavior has an impact on navigation and position accuracy and the operation of power grids and pipeline networks. We therefore use a number of decades of one-minute mean magnetic data from magnetic observatories in Europe, together with the technique of extreme value statistics, to provide a preliminary exploration of the extremes in magnetic field variations and their one-minute rates of change. These extremes are expressed in terms of the variations that might be observed every 100 and 200 years in the horizontal strength and in the declination of the field. We find that both measured and extrapolated extreme values generally increase with geomagnetic latitude (as might be expected), though there is a marked maximum in estimated extreme levels between about 53 and 62 degrees north. At typical midlatitude European observatories (55–60 degrees geomagnetic latitude), compass variations may reach approximately 3–8 degrees/minute, and horizontal field changes may reach 1000–4000 nT/minute, in one magnetic storm once every 100 years. For storm return periods of 200 years the equivalent figures are 4–11 degrees/minute and 1000–6000 nT/minute.

Citation: Thomson, A. W. P., E. B. Dawson, and S. J. Reay (2011), Quantifying extreme behavior in geomagnetic activity, *Space Weather*, *9*, S10001, doi:10.1029/2011SW000696.

1. Introduction

[2] A rapidly changing geomagnetic field constitutes a natural hazard, for example to grounded power grids and pipeline networks [e.g., *Lanzerotti*, 2001; *Pirjola et al.*, 2005] and to precise navigation [e.g., *Reay et al.*, 2005]. To understand this hazard we have continuous magnetic measurements across the world for typically less than 100 years. Much of the older data is in analog form, as paper records, or is only available digitally as hourly or daily means, or as magnetic indices. So it is not yet clear what the true extremes in geomagnetic variation are, particularly on time scales, seconds to minutes, that are relevant in estimating the hazard to technological systems from space weather.

[3] Within Europe many magnetic observatories have produced digital data at a sampling rate of one minute or better, in many instances for over two decades. In this time a number of severe magnetic storms have occurred, for example on 13 March 1989 and 30 October 2003. However, we know that larger storms have occurred in the past, most notably the Carrington event of September 1859

[e.g., *Cliver and Svalgaard*, 2004] and for which only limited information, in the form of analog paper records, is available.

[4] Extreme value statistics (EVS) is a technique that has been used in a wide variety of geophysical and environmental contexts, particularly in studies into extreme weather and climate (see, e.g., *Beirlant et al.* [2004], *Coles* [2004], and *Gilleland and Katz* [2006] for an introduction). EVS relies on the "extremal types theorem" [e.g., *Coles*, 2004], which shows that extreme data follow one of three types of distribution and that inference may therefore be made of likely extremes outside the data set by fitting observed data to a generalized distribution function.

[5] EVS has previously been applied to geomagnetic data, but only in the form of magnetic indices (*Dst* index [*Tsubouchi and Omura*, 2007; *Silbergleit*, 1996]; daily *Aa* index [*Silbergleit*, 1999]; "half-daily" *aa* index [*Siscoe*, 1976]). However, geomagnetic activity, as measured by indices, misses important fine-scale detail that may be relevant. For example, in terms of the geomagnetic hazard to power systems, the time rate-of-change of the field is widely regarded as the most relevant quantity [e.g., *Thomson*, 2007; *Thomson et al.*, 2010]. Other applications of EVS in the context of space weather (energetic electron

---

[1]Geomagnetism, British Geological Survey, Edinburgh, UK.

fluxes) are analyzed by *Koons* [2001] and *O'Brien et al.* [2007].

[6] In this paper we apply EVS to measured geomagnetic data at a temporal resolution that is arguably more relevant, as compared to hourly *Dst*, three-hourly *aa* or daily *Aa*, to the assessment of the space weather hazard to technology. We use EVS and digital one-minute data from 28 European observatories to assess the range of variations that may be observed up to a return period of 200 years. In the following section we present a brief overview of the EVS methodology. We then discuss the European observatory data set and the results obtained from applying EVS. Our results are given as "preliminary," in the sense that other observatory data sets may be added in time and that some of the assumptions need to be further tested.

2. Extreme Statistics Methodology

[7] The Gumbel, Frechet and Weibull distributions are widely used in the scientific literature when examining the tails of distributions of physical quantities. However, these three distributions can be combined in a single generalized extreme value (GEV) function of three parameters, describing the location (L), scale (S) and shape (K) of the distribution. GEV statistics are most appropriate where only block averages or block maxima are available, e.g., annual maxima of daily atmospheric temperatures (method described further by *Coles* [2004]). For our analysis however we use a generalized Pareto distribution (GPD) to describe the tail of the distribution of geomagnetic activity (see, e.g., *Coles* [2004] for full details and *Tsubouchi and Omura* [2007] for a summary as applied to magnetic index data). The GPD is, in principle, more appropriate for our data, as we have individual one minute samples and some idea of an appropriate threshold of extreme activity (although in practice we choose a threshold from statistical considerations; see section 3). This approach is known as "point over threshold," where the GPD function is fitted to all data over the threshold.

[8] The GPD is defined as (for large u)

$$G_{u,K,S'}(x) = 1 - \left[1 + K(x - u)S'\right]^{(-1/K)} \tag{1}$$

where $x\text{-}u > 0$, $1 + K(x - u)/S' > 0$ and $S' = S + K(u - L)$, given in terms of the equivalent parameters, L, S, K, from the GEV distribution. (1-G) gives the probability that the random variable (below we use a magnetic field residual or its rate of change as the random variable), X, exceeds some value, x, given that it already exceeds a threshold, u, i.e., $Pr[X > x \mid X > u]$ [*Gilleland and Katz*, 2006].

[9] There is a duality between the GPD and GEV functions such that K has the same meaning for both distributions. Thus in equation (1) $K = 0$, $K > 0$ and $K < 0$ correspond to the Gumbel, Frechet and Weibull distributions respectively. The Weibull distribution is bounded above, i.e., there is a maximum that cannot be exceeded (and should therefore be applicable to some

magnetic indices, such as *Kp*, *Ap* and *aa*). The Gumbel distribution is light-tailed, in that it is unbounded, although the probability of increasingly higher values decays exponentially. The Frechet distribution decays polynomially with the size of the random variable.

[10] For a given threshold selection the parameters of the GPD (through equation (1)) can be determined by maximum likelihood estimation. In our analysis we used the 'R' statistics package [*R Foundation for Statistical Computing*, 2008] and the eXtremes toolkit of Gilleland and Katz (Tutorial for the Extremes Toolkit: Weather and Climate Applications of extreme value statistics, http://www.assessment.ucar.edu/toolkit, as described by *Stephenson and Gilleland* [2006]).

[11] By fitting separate GPDs to individual observatory data sets, we can estimate the probability of geomagnetic conditions more extreme than those contained within those data sets, and thus provide estimates of the most likely extremes expected to be observed within a given period. To do this we construct a "return level" plot [*Coles*, 2004], which expresses how the most likely extreme value increases with the return period, together with an estimate of an appropriate confidence interval for that estimate. (An example of a return level plot is given later in Figure 4.)

[12] There are subtleties in applying extreme value statistics to geomagnetic data, e.g., clusters of extreme values may occur during a single magnetic storm or group of related sub-storms, meaning each sample is not statistically independent; and, in general, the non-stationary nature of geomagnetic data, tied as it is to the solar activity cycle. Though these can be dealt with, to leave stationary and independent random estimates for analysis, only de-clustering and threshold selection is considered in this analysis.

3. How Extreme Can Geomagnetic Storms Become?

[13] One-minute geomagnetic time series of *H* (horizontal field) and *D* (declination) were obtained from the World Data Centre for Geomagnetism in Edinburgh (www.wdc. bgs.ac.uk) for 28 European observatories. Europe was chosen for this exercise because of the density of measurement sites and the length of digital sampling at those sites. The particular observatories were chosen (see Figures 1 and 2) to represent a spread of sampling points across the continent, covering a range of magnetic latitudes and for which there exist digital data spanning a number of years. From *H* and *D* we computed time series of the residuals with respect to quiet levels, to isolate the external field influenced by space weather. These residuals were determined for each observatory by removing the quiet mean level, established for each month and for each observatory from the five "International Quiet Days," as determined by the International Service of Geomagnetic Indices. A further two time series, d*H*/dt and d*D*/dt, were constructed for each observatory by computing first-differences of minute-means. Individual GPD analyses
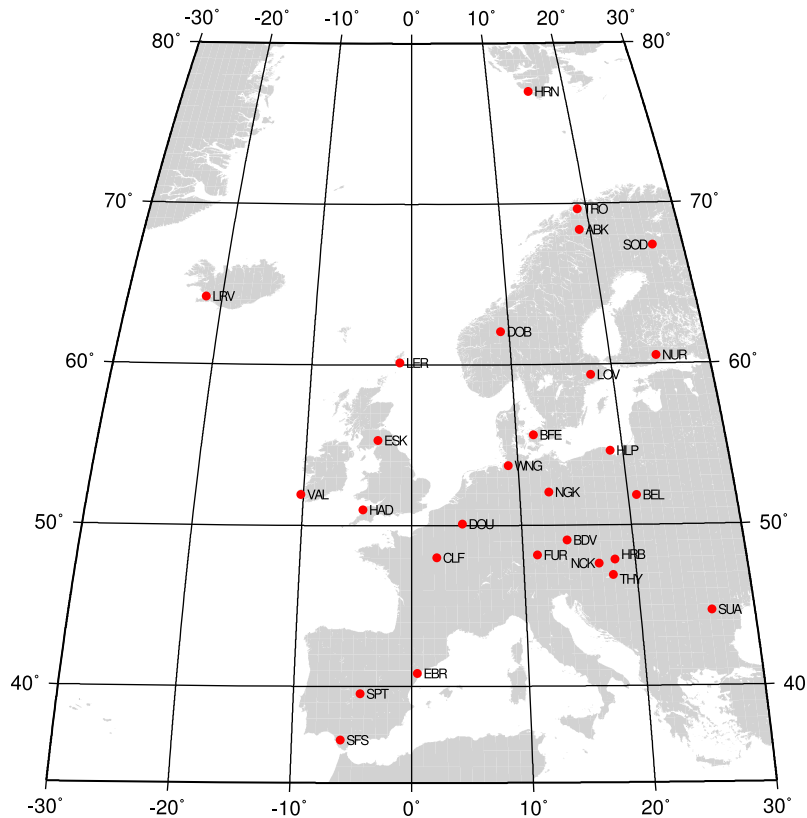
Figure 1. The geographical distribution of European observatories used in the study.

were performed on these four time series from each of the 28 observatories.

[14] Before fitting each GPD to the data, we first determined a threshold to mark the onset of extreme behavior. All data points below this threshold were then discarded. An appropriate threshold for each of the variables, and for each observatory, can be determined by plotting the scale (S) and shape (K) parameters of the resulting GPD for a range of thresholds (using techniques described by *Coles* [2004]). The ideal threshold should be low enough to allow for a meaningful number of samples, but high enough that the modified scale parameter is approximately constant and the shape parameter approximately linear (within error-margins), above the chosen threshold. We found that setting the threshold at the 99.97th percentile proved reasonable for each variable at most observatories.

[15] In Figure 3 we show the effect of this choice of threshold in terms of the "mean residual life" (MRL) curve for a typical midlatitude observatory (Eskdalemuir, UK). The MRL for a data set is a function of the threshold level, u, and measures the mean difference between the exceedances above each threshold and the threshold itself [*Coles*, 2004, section 4.3]. That is, the MRL is a plot of $(u, \sum (X_i - u)/N_i)$, where u is some threshold, and the X_i are all the N_i samples above that threshold.

[16] The key feature of the MRL is that an appropriate threshold for extreme behavior in a data set is marked by the onset of a linear trend. In Figure 3 this occurs around the 99.97 percentile, consistent with the scale and shape parameter threshold estimation method. For other observatories we found a broadly similar behavior and thus chose 99.97% as a common threshold to ensure homogeneity of analysis and results across the continent. However, for some observatory data sets, this choice of threshold may produce few threshold exceedances and hence reduce the confidence in the EVS fit and subsequent extrapolation.

[17] Clusters of extreme values occur during geomagnetic storms. These can be where the storm maximum is accompanied closely in time with other near-maxima. Including these near-maxima in the data skews the extreme value statistics in the following sense: we wish to identify the return period for each major event, i.e., single magnetic storm, in the sense that one would describe 13 March 1989 or 30 October 2003 as a single event, even though there was much complex structure in the activity that occurred on each of these days.

[18] Clustering results in statistical dependency in the data and that should be eliminated to meet the assumptions of the model. We identified clusters by looking for extreme values that were separated by at least 12 h. Only
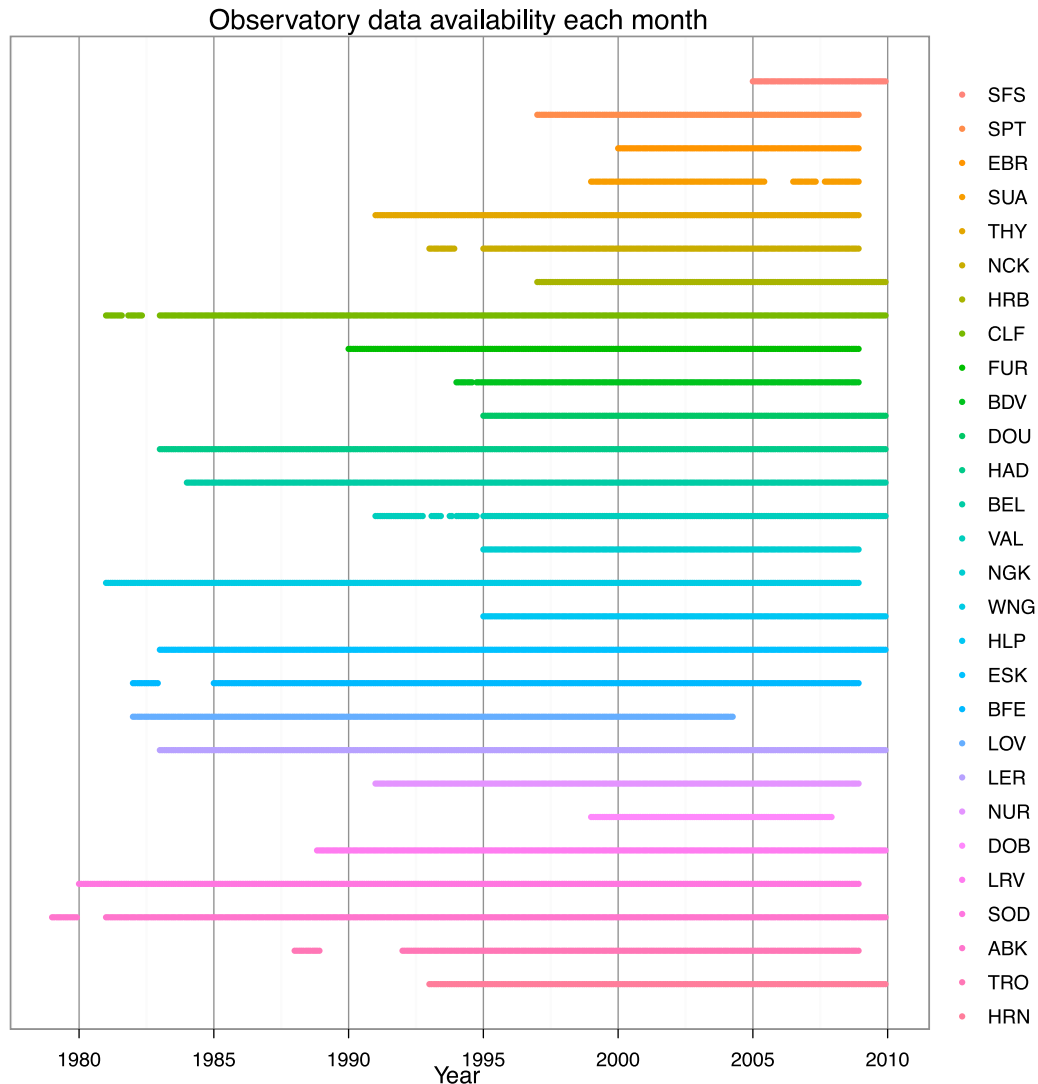
Figure 2. The digital data availability of the observatories in Figure 1. Color is used to identify each observatory but has no other significance.

the peak value from each cluster was retained. In Figure 4 we show the effect of combining thresholding with de-clustering at Eskdalemuir observatory, in a "return level" plot (compare with description given in section 2). Figure 4 shows how the magnitude of the extreme increases as a function of return time, depending on the threshold and the cluster length applied to the data. Applying de-clustering widens the estimated 95% confidence limits, as does increasing the threshold. Comparison with observed maxima (small circles) suggests that a threshold of 99.97% and a de-cluster length of 12 h is a reasonable compromise and one that matches the observations at low return periods (the bottom center plot in Figure 4).

[19] Experimenting with the de-cluster length (from three hours to one week) showed that the return level is generally only weakly dependent on the de-cluster length for most observatories, but that de-clustering certainly increases the extrapolated return level, compared to not applying de-clustering. However, for some observatories, such as Chambon-la-Foret (CLF), Hurbanovo (HRB) and San Fernando (SFS), varying the de-clustering length strongly affected the estimated return level. Whether there is a reason for this difference between some observatories and the majority in our sample is not known. For example, we note that San Fernando has a short data span, while the span of Hurbanovo data is among the longest in the sample.

[20] For each observatory we computed the peak residual and peak rate-of-change predicted by the observatory GPD to be exceeded once within periods of 100 and 200 years, from examination of the return-level statistics (e.g., Figure 4). The results for the four time series are summarized in Figures 5 (H residual), 6 (dH/dt), 7 (D residual) and 8 (dD/dt). In these figures we do not consider any local time,
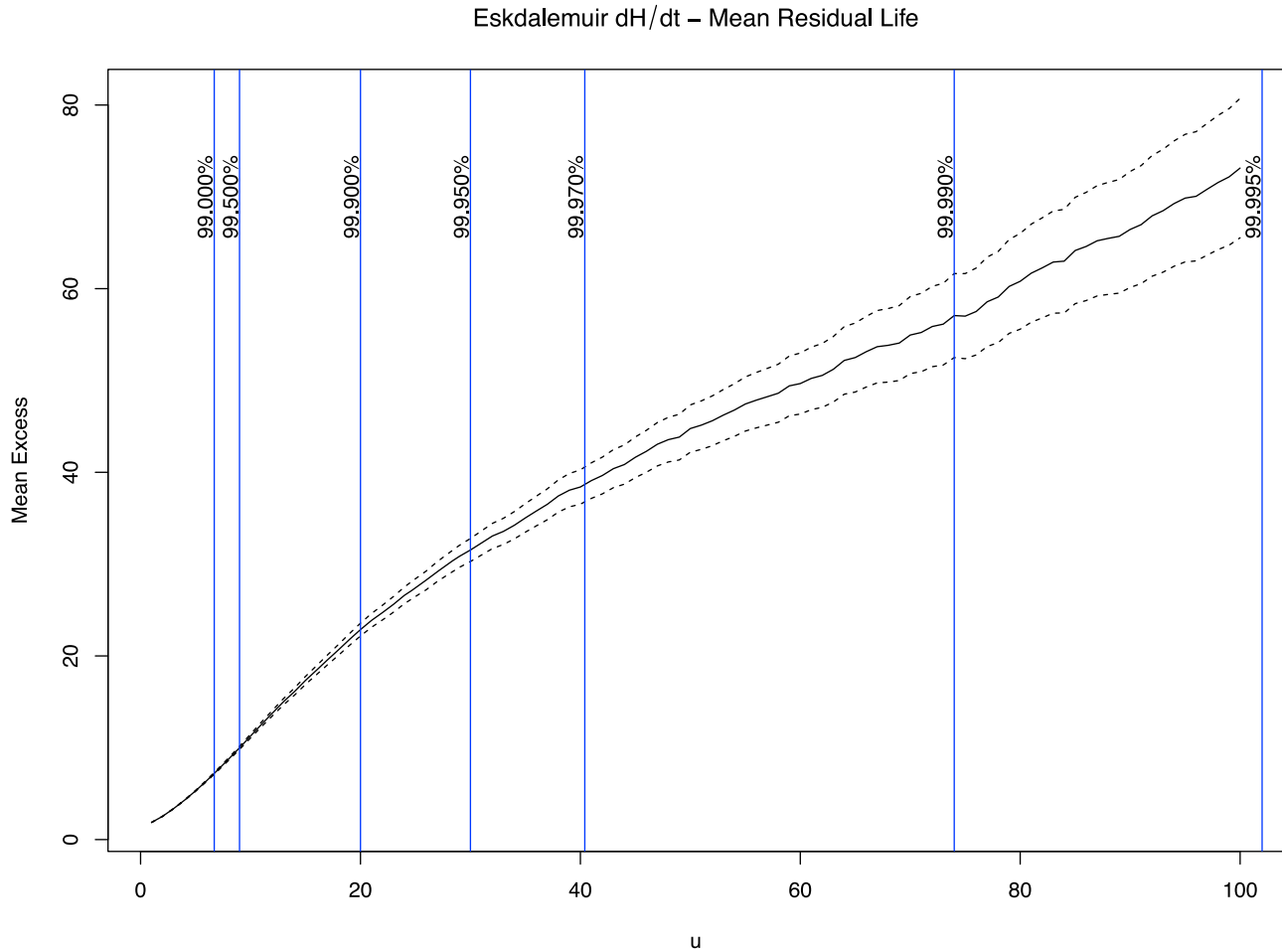
Figure 3. The mean residual life plot for Eskdalemuir observatory (57.8°N geomagnetic latitude). The plot shows the "mean excess" as a function of the excess threshold u. Vertical percentiles are labeled to indicate, e.g., that 99.500% of the data lie to the left of each line, and are <10 nT/min. The quantity shown here is dH/dt, and units for u and the mean excess are nT/min.

or longitude, dependence as significant as the latitude dependence. During any individual storm, when the peak electrojet activity occurs will define where the extreme is localized in longitude (if it is localized). However, at the time scales of years between single storms and with de-clustering of 12 h or more, such small time dependency of a few hours is probably not significant (discussed more in the concluding section).

[21] In all four figures there is a clear increase in measured and therefore predicted extreme levels as a function of geomagnetic latitude. This is perhaps more marked for the declination data. There is also a distinct local maximum around 53–62 degrees. We can attribute this to an enhanced auroral electrojet, which will move south under strong forcing from the solar wind.

[22] Geomagnetic data residuals are typically non-Gaussian in nature and one standard deviation uncertainties in the prediction may give a false picture of the range of

uncertainty in the prediction. For that reason we use 95% confidence limits in Figures 5–8. It is also clear from Figures 5–8 that there can be a distinct spread in behavior (in both observed and predicted) from observatories close in latitude; this being particularly true in the proximity of the auroral electrojet. Such variability will partly reflect the differing time spans of each observatory series (as in Figure 1). We note, for example, that the 1982 severe storm is represented in observed data from the Lovö and Brorfelde observatories, but not from others, such as Eskdalemuir, whose series starts in 1983. Interestingly the 100 and 200 year extrapolations made from the Lovo and Brorfelde data set are less extreme (e.g., Figure 6) than those estimated from Eskdalemuir. This gives some insight into the uncertainties underlying the method. In Figures 5–8 the data from Valentia observatory seem "anomalous," in the sense that both the measured and predicted extremes are significantly less than from other observatories at similar
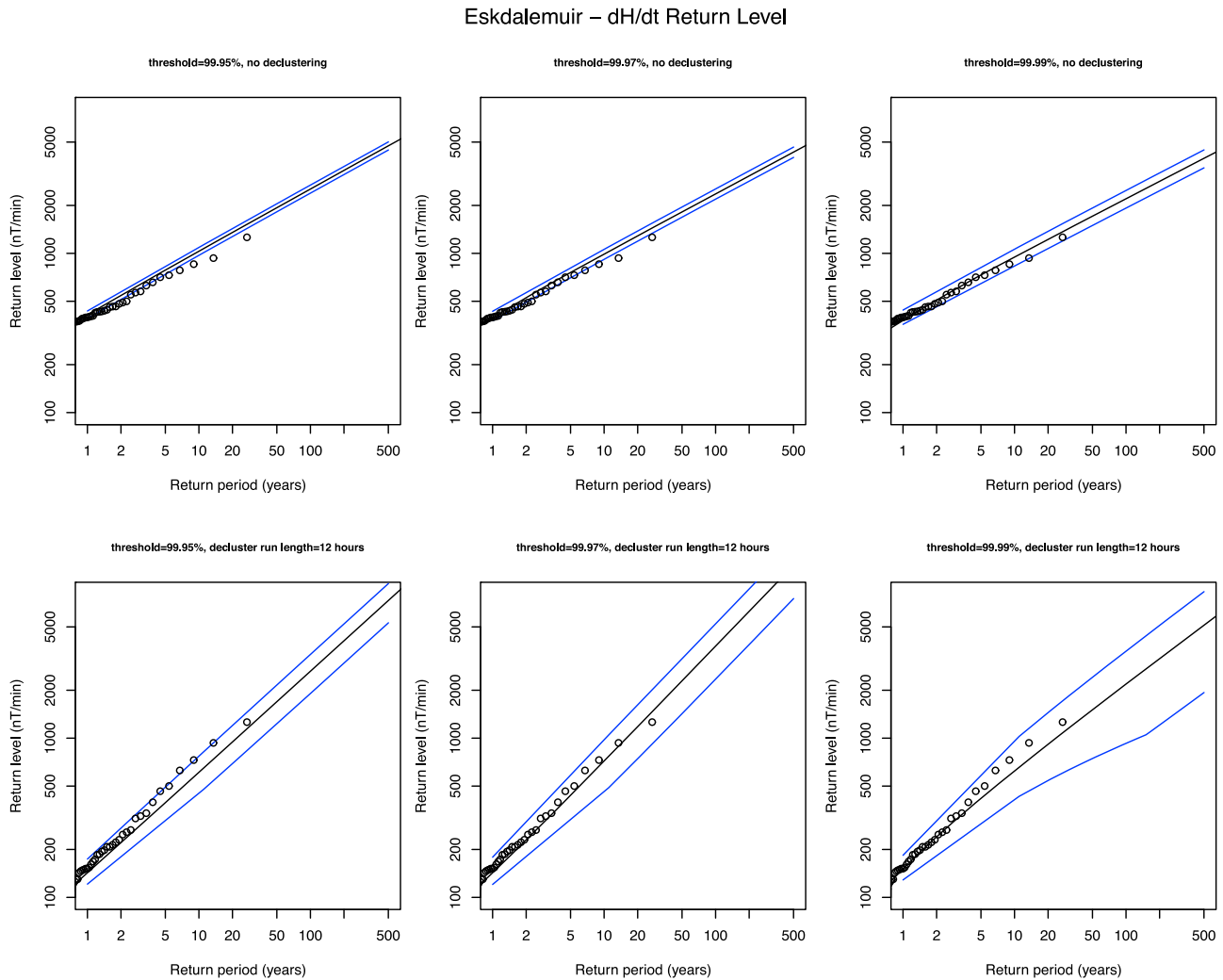
Figure 4. The Eskdalemuir d*H*/dt return level (in nT/min) as (top) a function of threshold, where no declustering is used, and (bottom) as a function of threshold level, where a decluster length of 12 h is used. Measured data are shown by small circles. Asymmetric 95% confidence levels in the EVS fit are also given. The preferred model of the data is the plot at bottom center.

latitudes. Why this should be the case, and whether it is significant, is not known, though it may be related to geology (crustal anomaly fields), geography, deep ocean proximity and/or instrumentation. We note that Hartland observatory data may also be anomalous, among the group of stations just south of 55 degrees.

[23] In Table 1 we provide a broad summary of the likely maximum observed in each of the four components in the "interesting" range of latitudes between about 55 and 60 degrees north (geomagnetic). These figures are simply a guide to the variations that could be observed. For other latitudes, particularly where there is less fluctuation between observatories, typical values and their uncertainties can be read directly from Figures 5–8. In Table 1, if one excludes Valentia observatory as being "anomalous," then the figures in parentheses apply. Excluding Valentia data

serves only to increase the lower bound on the estimated maxima.

[24] From the GPD fits to each observatory data set we find that, for nearly all observatories and data types, the shape parameter *K* is positive, but generally less than one, and exceeds its standard error. This suggests that extreme geomagnetic activity is not bounded by some maximum, but follows either a Gumbel or Frechet distribution, i.e., the probability of increasingly higher values diminishes exponentially or polynomially with value. The consequence of this is that over even longer return periods, compared for example with 200 years, higher extreme values can occur. The scale parameter *S* for all fits is positive and always exceeds the formal standard error on *S*. Finally, all GPD fits show these characteristics regardless
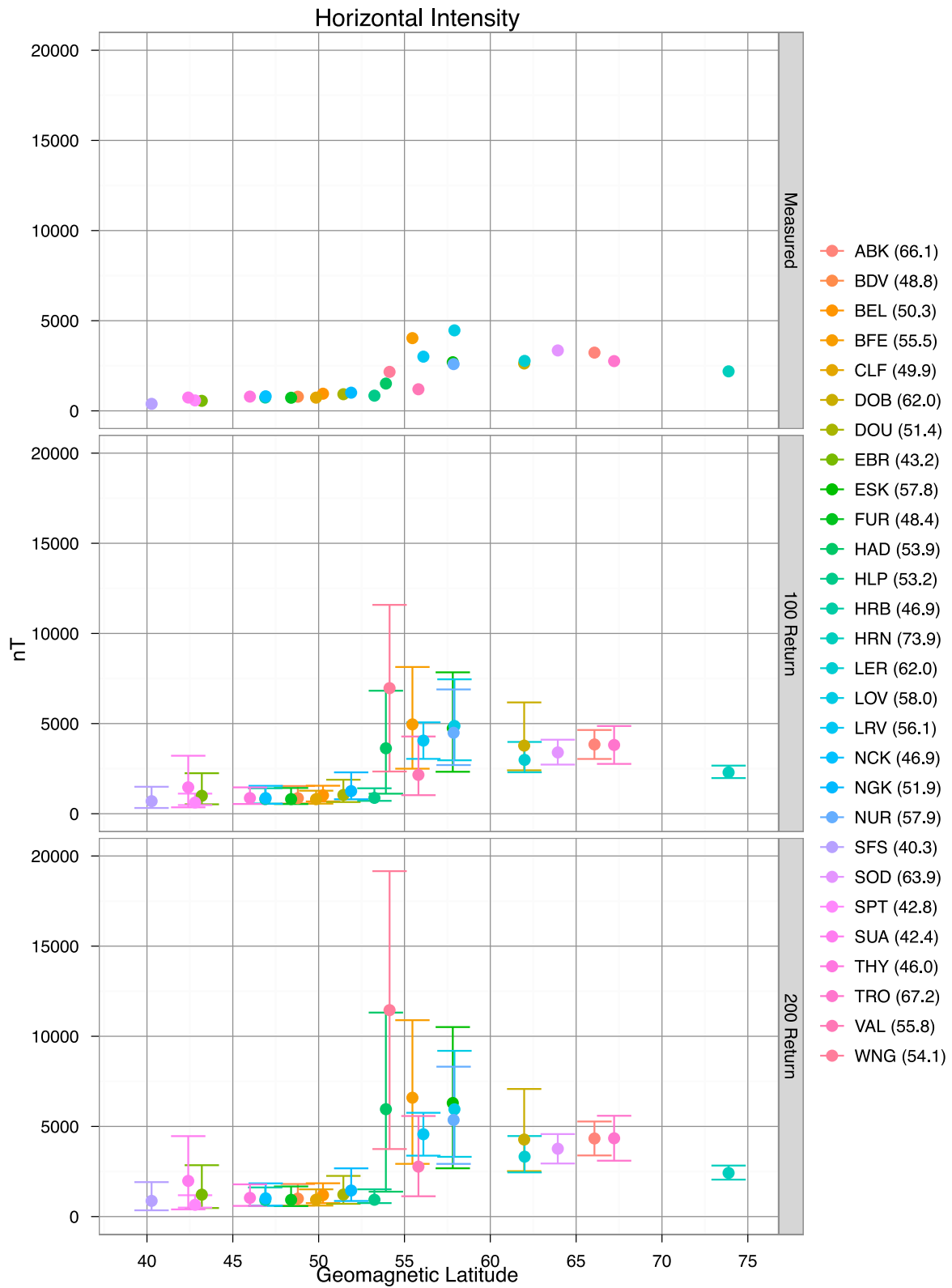
Figure 5. (top) The measured maximum for each observatory and estimated (middle) 100-year and (bottom) 200-year return levels, for H, in nT, as a function of geomagnetic latitude. The 95% confidence limits are shown.
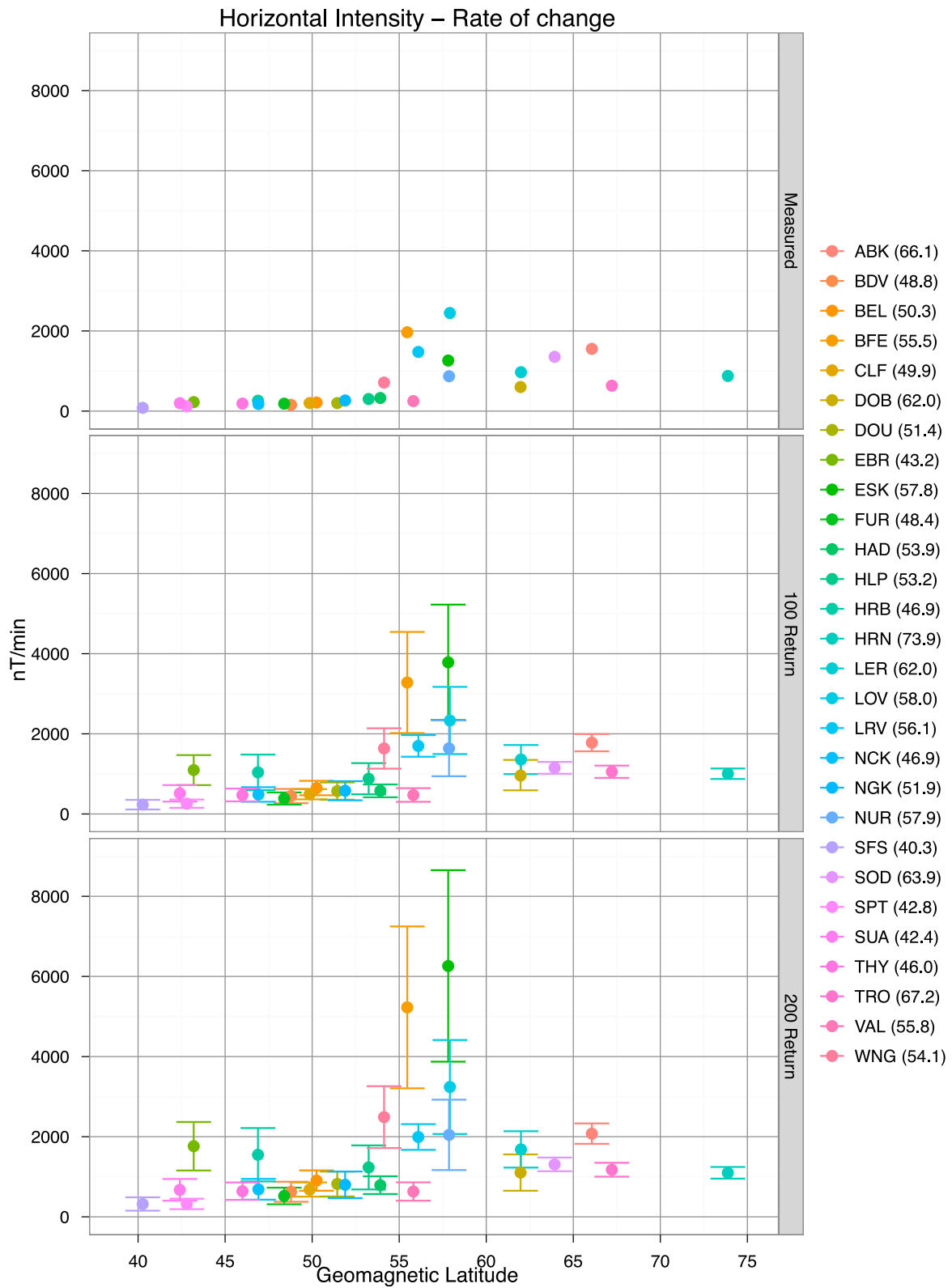
Figure 6. (top) The measured maximum for each observatory and estimated (middle) 100-year and (bottom) 200-year return levels, for dH/dt, in nT/min, as a function of geomagnetic latitude. The 95% confidence limits are shown.
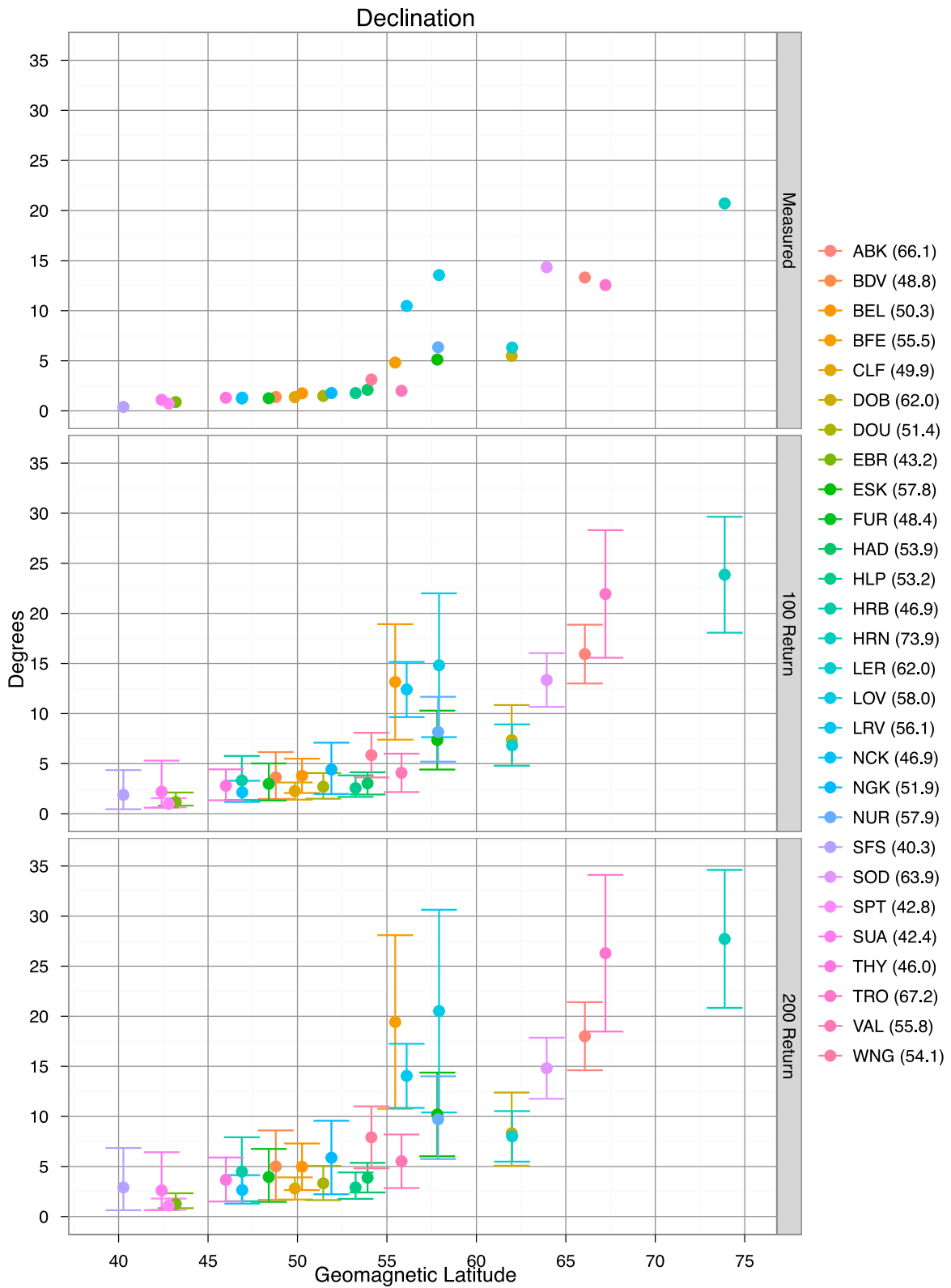
Figure 7. (top) The measured maximum for each observatory and estimated (middle) 100-year and (bottom) 200-year return levels, for D, in degrees, as a function of geomagnetic latitude. The 95% confidence limits are shown.
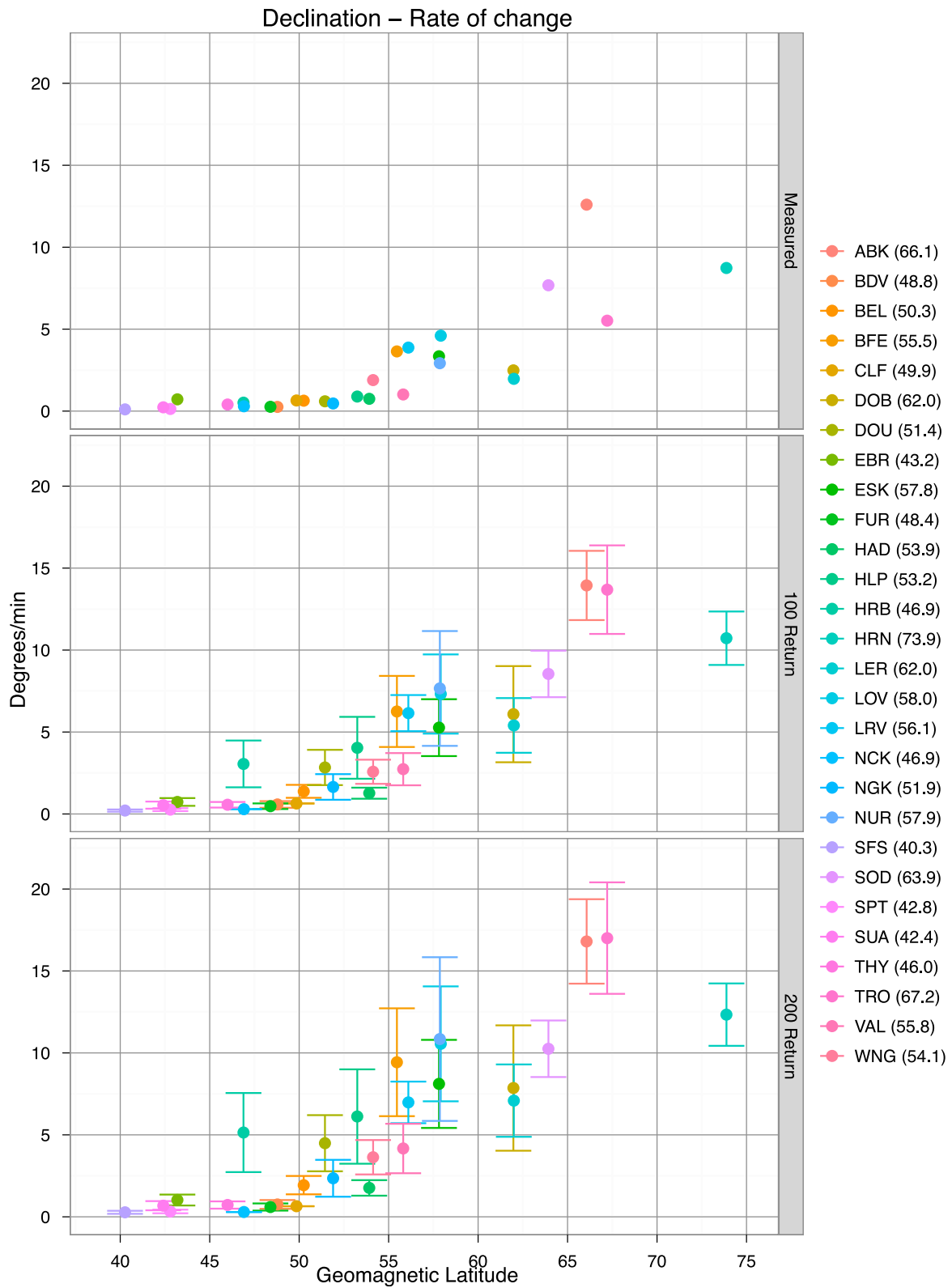
Figure 8. (top) The measured maximum for each observatory and estimated (middle) 100-year and (bottom) 200-year return levels, for d*D*/dt, in degrees/min, as a function of geomagnetic latitude. The 95% confidence limits are shown.

Table 1. Estimated 100 and 200 Year Maxima in Each of *H*, d*H*/dt, *D*, and d*D*/dt Between 55 and 60 Geomagnetic Degrees North, Summarized From Figures 5–8[a]

	H (nT)	d*H*/dt (nT/min)	*D* (deg)	d*D*/dt (deg/min)
100 year return	2000–5000 (4000–5000)	1000–4000 (1500–4000)	4–15 (7–15)	3–8 (5–8)
200 year return	3000–6500 (4500–6500)	1000–6000 (2000–6000)	5–20 (10–20)	4–11 (7–11)

[a]Estimated maxima for other latitudes can be obtained directly from Figure 5–8. These are only guide estimates as the spread in results from Figures 5–8 demonstrates clearly. Values in parentheses apply where Valentia observatory data are excluded.

of whether residuals or rates-of-change in residuals are analyzed.

4. Conclusions

[25] We regard our analysis as preliminary, partly because it describes only European data but partly also because of the simplifying assumptions we have made. However, we find that predicted return magnitudes increase with latitude, as would be expected, although there is also some interesting structure in the data between about 53 and 62 degrees north (geomagnetic), as the auroral electrojet moves south from its more 'normal' position between 65 and 75 degrees. (Thus also explaining why the more northerly observatories can experience smaller extremes than those predicted around 55–60 degrees.)

[26] Figures 5–8 detail, and Table 1 summarizes, the range of extremes found in the analysis (based on 95% confidence limits). For example at mid European latitudes (55–60 degrees geomagnetic), once every 100 (200) years the change in compass variation is estimated to exceed 3–8 (4–11) degrees per minute, and horizontal field changes may exceed 1000–4000 (1000–6000) nT/min. These figures can only be considered as guide values, as the ranges observed within Figures 5–8 demonstrate. The data for Valentia observatory also seem anomalous, in comparison with other midlatitude observatories, and for reasons unknown. Table 1 also includes a statistical summary, where Valentia is excluded. Both Table 1 and Figures 5–8 apply under the assumption that the local time or longitude of any enhanced auroral currents driving the extreme conditions is not significant, in comparison with the time scales (100–200 years) between extreme storms.

[27] The statistical analysis might be improved by, for example, treating variables *D* (or *H*) and d*D*/dt (or d*H*/dt) as components of the same multivariate statistic. It should also be possible to extend the application of this technique to look at magnetic extremes across the world, as seen in global observatory data, dependent on data availability. This would help confirm whether the results obtained here from the European continent are applicable elsewhere. It is also feasible to examine higher extremes that may return over longer time spans (e.g., as in Figure 3,

at 500 years), although the confidence limits widen considerably. Even in this paper we have looked at predictions on time scales up to ten times (200 years) that of the duration of the original digital data (~20–30 years).

[28] The non-stationary nature of geomagnetic data implies that our analysis probably needs refinement. There is a solar cycle dependent variability seen in residuals, and perhaps de-trending by smoothed sunspot number or by monthly mean numbers of coronal mass ejections is required. We may return to this in future work. However, in our results we look at likely trends over many solar cycles and the significance of such "short" period variations may be less. One may also consider the appropriateness of the "block-averaging" versus "point-over-threshold" methods and perhaps re-examine the threshold and de-clustering choices we have made. These may all have some impact on the robustness of our results. However, the results already given here should find application in hazard assessment and in magnetic navigation applications, not least in helping to assess the risk to power systems and to magnetic navigation activities within Europe and beyond.

References

Beirlant, J., Y. Goegebeur, J. Segers, and J. Teugels (2004), *Statistics of Extremes*, John Wiley, Chichester, U. K.

Cliver, E., and L. Svalgaard (2004), The 1859 solar-terrestrial disturbance and the current limits of extreme space weather activity, *Sol. Phys.*, *224*(1–2), 407–422, doi:10.1007/s11207-005-4980-z.

Coles, S. (2004), *An Introduction to Statistical Modelling of Extreme Values*, Springer, London.

Gilleland, E., and R. W. Katz (2006), Analyzing seasonal to interannual extreme weather and climate variability with the extremes toolkit (extRemes), paper presented at the 18th Conference on Climate Variability and Change, 86th American Meteorological Society (AMS) Annual Meeting, Atlanta, Ga., 29 Jan. to 2 Feb. [Available at http://www.ral.ucar.edu/staff/ericg/.]

Koons, H. C. (2001), Statistical analysis of extreme values in space science, *J. Geophys. Res.*, *106*, 10,915–10,921, doi:10.1029/2000JA000234.

Lanzerotti, L. J. (2001), Space weather effects on technologies, in *Space Weather*, *Geophys. Monogr. Ser.*, vol. *125*, edited by P. Song, H. J. Singer, and G. L. Siscoe, pp. 11–22, AGU, Washington, D. C.

O'Brien, T. P., J. F. Fennell, J. L. Roeder, and G. D. Reeves (2007), Extreme electron fluxes in the outer zone, *Space Weather*, *5*, S01001, doi:10.1029/2006SW000240.

Pirjola, R., K. Kauristie, H. Lappalainen, and A. Viljanen (2005), Space weather risk, *Space Weather*, *3*, S02A02, doi:10.1029/2004SW000112.

Reay, S. J., W. Allen, O. Baillie, J. Bowe, E. Clarke, V. Lesur, and S. Macmillan (2005), Space weather effects on drilling accuracy in the North Sea, *Ann. Geophys.*, *23*, 3081–3088, doi:10.5194/angeo-23-3081-2005.

R Foundation for Statistical Computing (2008), *R: A Language and Environment for Statistical Computing*, Vienna. [Available at http://www.r-project.org.]

Silbergleit, V. M. (1996), On the occurrence of geomagnetic storms with sudden commencements, *J. Geomagn. Geoelectr.*, *48*, 1011–1016, doi:10.5636/jgg.48.1011.

Silbergleit, V. M. (1999), Forecast of the most geomagnetically disturbed days, *Earth Planets Space*, *51*, 19–22.

Siscoe, G. L. (1976), On the statistics of the largest geomagnetic storms per solar cycle, *J. Geophys. Res.*, *81*, 4782–4784, doi:10.1029/JA081i025p04782.

Stephenson, A., and E. Gilleland (2006), Software for the analysis of extreme events: The current state and future directions, *Extremes*, *8*, 87–109, doi:10.1007/s10687-006-7962-0.

Thomson, A. W. P. (2007), Geomagnetic hazard, in *Encyclopaedia of Geomagnetism and Paleomagnetism*, edited by D. Gubbins and E. Herrero-Bervera, pp. 316–319, Springer, Dordrecht, Netherlands, doi:10.1007/978-1-4020-4423-6_117.

Thomson, A. W. P., C. T. Gaunt, P. Cilliers, J. A. Wild, B. Opperman, L.-A. McKinnell, P. Kotze, C. M. Ngwira, and S. I. Lotz (2010), Present day challenges in understanding the geomagnetic hazard to national power grids, *Adv. Space Res.*, *45*(9), 1182–1190, doi:10.1016/j.asr.2009.11.023.

Tsubouchi, K., and Y. Omura (2007), Long-term occurrence probabilities of intense geomagnetic storm events, *Space Weather*, *5*, S12003, doi:10.1029/2007SW000329.

E. B. Dawson, S. J. Reay, and A. W. P. Thomson, Geomagnetism, British Geological Survey, West Mains Road, Edinburgh EH9 3LA, UK. (awpt@bgs.ac.uk)