Modelling directionality, seasonality, and local time dependences in extreme geomagnetic field fluctuations

Neil C. Rogers[1], Jim A. Wild[1] and Emma F. Eastoe[2]

1: Space & Planetary Physics group, Lancaster University, UK 2: Dept. of Mathematics and Statistics, Lancaster University, UK
n.rogers1@lancaster.ac.uk

1: Introduction

Statistics of extreme temporal changes in the horizontal component of the geomagnetic field ($\boldsymbol{dB_H/dt}$) may be used to assess the risk of damaging geomagnetically induced currents (GIC). By fitting Generalised Pareto (GP) distributions [1] to measurements of $|dB_H/dt|$ above a high threshold, we have determined return levels (RL) of $|dB_H/dt|$ expected over periods of 100 years or more. Large fluctuations are driven by diverse magneto-ionospheric driving processes including substorm expansions, Pc5 ULF waves, and sudden commencements (see Fig. 1). The occurrence rate and magnitude of large $|dB_H/dt|$ therefore vary with geomagnetic latitude, magnetic local time (MLT), season, and with the compass direction of the fluctuation $\boldsymbol{dB_H}$. Occurrence rates are also dependent on the interplanetary magnetic field (IMF) orientation and vary across the solar activity cycle.

In applying extreme value theory it is assumed that extreme events are (i) independent, and (ii) identically distributed (IID). In this poster we describe how we satisfy (i) by declustering, and (ii) by combining GP distributions from subsets of the data discretely sectored by direction, season, or local time sector.

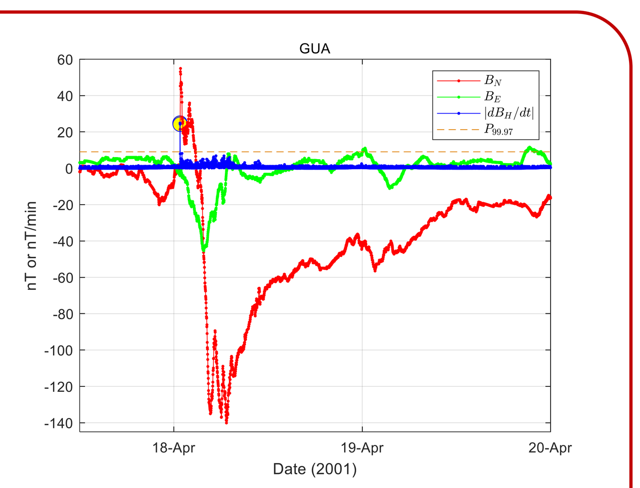

Fig. 1. Magnetic field fluctuations at a low-latitude magnetometer (Guam (GUA)). The rapid fluctuation $|dB_H/dt|$ (blue line) highlighted exceeds the 99.97th percentile (dashed line) early on 18 April 2001 and is associated with a Storm Sudden Commencement.

2: SuperMAG Magnetometer Data

From the SuperMAG database [2,3] we selected geomagnetic measurements from 125 stations (see **Fig. 2**), each with between 20 and 48 years of geomagnetic field measurements (1-minute averages). These were baselined to remove yearly and diurnal trends, rotated into local geomagnetic coordinates, and manually inspected to remove artefacts. For each site, only values of $|dB_H/dt|$ above the 99.97th percentile threshold $P_{99.97}$ were retained. The values of $P_{99.97}$ are indicated by the colours in **Fig. 2** and are greatest in the auroral zone (55–80° corrected geomagnetic (CGM) latitude), which is subject to intense enhancements of the auroral electrojets during geomagnetic substorms. To remove temporal dependence in the points over threshold, declustering was applied by discounting contiguous data above the $P_{99.97}$ threshold and requiring that values are below threshold for ≥ 12 hours to be considered independent events. Only the largest point in each cluster is retained.

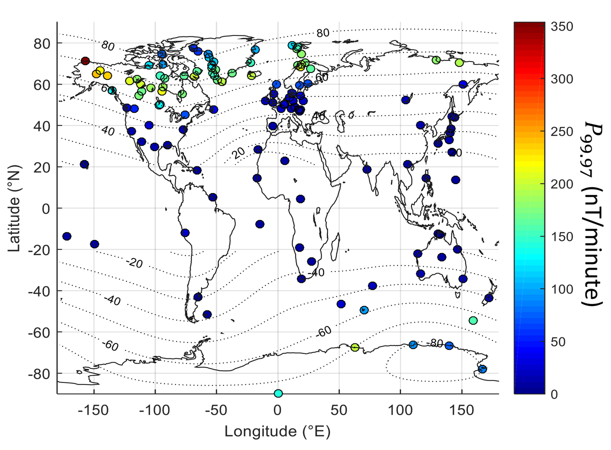

Fig. 2. The 99.97th percentile of $|dB_H/dt|$ for each site. Contours indicate corrected geomagnetic (CGM) latitudes.

3: Modelling Extremes of $|dB_H/dt|$

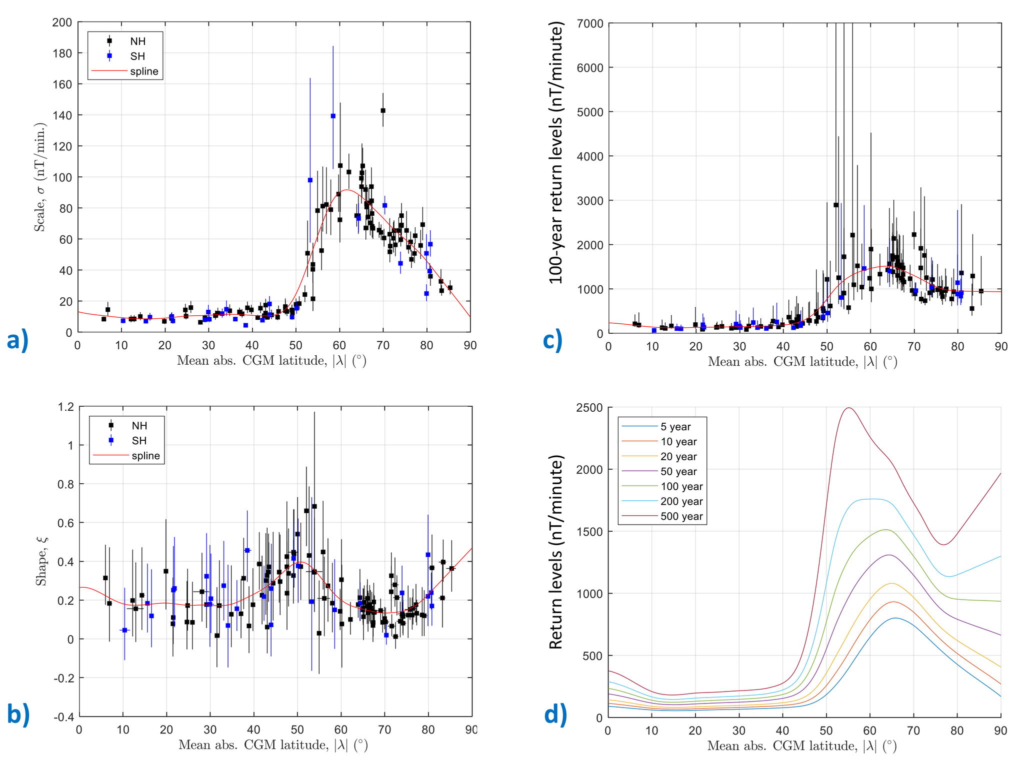

Fig. 3. Absolute geomagnetic latitude profiles of fitted GP parameters a) scale, σ, and b) shape, ξ (Eqn 1). c) 100-year return levels of $|dB_H/dt|$ (Eqn 2), and d) smoothed spline fits for a range of return periods. Error bars indicate 95% confidence intervals.

To model the probability of extremely large (and rare) values that may not yet have been observed, we fitted a GP distribution (Eqn. 1) to the declustered values of X $=|dB_H/dt|$ above the 99.97th percentile ($P_{99.97}$). Thus, assuming IID,

$$Pr(X \le \boldsymbol{x} \mid X > u) = 1 - \left[1 + \frac{\xi(\boldsymbol{x} - u)}{\sigma}\right]_+^{-1/\xi}$$

Equation 1

where $[\cdot]_+ = \max(\cdot, 0)$, and $\sigma > 0$. Maximum likelihood estimates (MLE) for the GP parameters ξ and σ were determined numerically for each site and are presented in **Fig. 3a** and **b** as a function of absolute CGM latitude, with error bars indicating 95% confidence intervals (CI). The N-year return level is

$$x_N = u + \frac{\sigma}{\xi}\left[\left(N\, n_y\, \zeta_u\right)^{\xi} - 1\right]$$

Equation 2

where n_y is the number of samples per year (= 365.25 × 24 × 60), and $\zeta_u = Pr(X > u)$ is the occurrence rate of exceedances.

100-year return levels of $|dB_H/dt|$ are presented in **Fig. 3c**. A smoothing spline is fitted to the magnetic latitude profile of MLEs and repeated for various return periods to produce the RL curves in **Fig. 3d**. This shows that for higher return periods, the return values peak at lower auroral latitudes (~ 53°). This reflects the fact that the most extreme (and rare) substorm currents occur at lower auroral latitudes, after a pronounced expansion phase. Also note the secondary peak in return levels above 75° CGM latitude for return periods greater than 100 years.

4: Latitude, Season, and Local Time Dependences

Fig 4a presents the probability of cluster peak occurrences, $Pr(|dB_H/dt| > P_{99.97})$ as a function of absolute corrected geomagnetic latitude, $|\lambda|$ and MLT, which may be substituted for ζ_u (Eqn 2) if $u = P_{99.97}$. The principal magneto-ionospheric drivers for extreme $|dB_H/dt|$ are labelled. **Fig 4b.** presents a 19th order spherical harmonic fit to the probability surface, which provides a smooth interpolating function for modelling.

Fig. 5 presents, for three different northern hemisphere latitudinal regions, the occurrence probabilities Pr($|dB_H/dt| > P_{99.97}$) *vs* (month, MLT) (top row) and *vs* (direction, MLT) (bottom row). The strong northward directional preference for low-latitude stations (left panels) is associated with Sudden Commencement activity. For auroral locations (centre panels) the strong equinoctial maxima in the 20-24 MLT sector is associated with substorm expansions in which the predominantly southward direction results from strong westward auroral electrojet currents. In the polar region (right panels) the increased occurrence near summer noon may result from increased tail-lobe reconnection events, which occur under conditions of high dipole tilt.

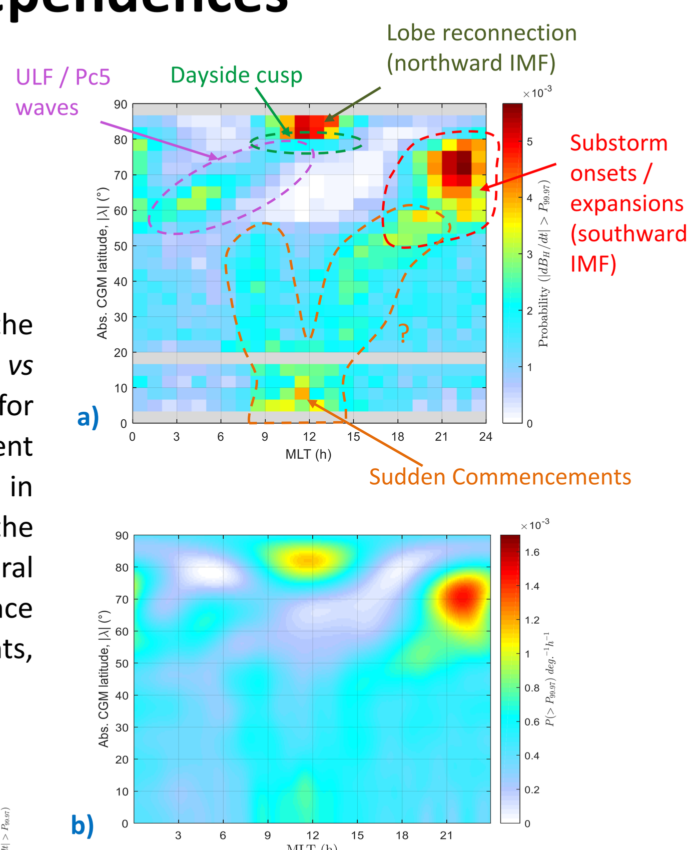

Fig. 4 a) CGM Latitude vs MLT distribution of cluster peak occurrence probabilities ($Pr(|dB_H/dt| > P_{99.97})$, indicating regions dominated by Sudden Commencements, Pc5 waves, Substorm expansions, and lobe reconnection poleward of the dayside cusp region.

b) A 19th order spherical harmonic series fit to the surface in a).

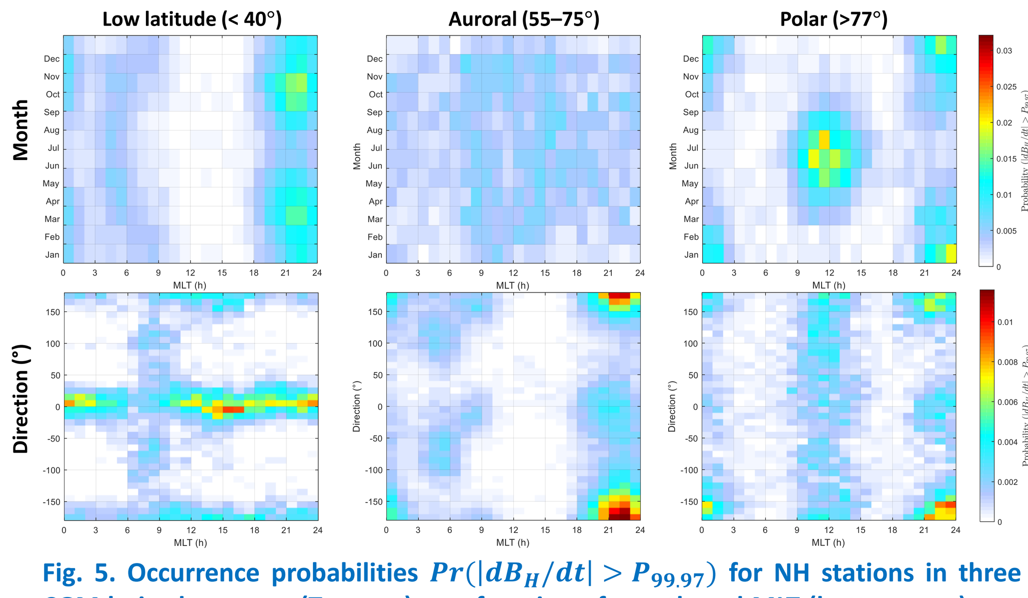

Fig. 5. Occurrence probabilities $Pr(|dB_H/dt| > P_{99.97})$ for NH stations in three CGM latitude ranges. (Top row) as a function of month and MLT (bottom row) as a function of compass direction of dB_H (° clockwise from N) and MLT.

5: Predicting Extreme $|dB_H/dt|$ from Discretely Sectored Directions, Seasons, or Magnetic Local Times

The strong dependence of occurrence likelihood on covariates (direction, MLT, month, etc.) violates the assumption of an identical distribution in Eqn 1. Models that take into account covariate effects may yield more accurate results since they explain more of the variability in the data by setting thresholds that vary with the covariate (see **Fig. 6**). The following example – based on an ocean wave height analysis method [4] – applies to peaks of $|dB_H/dt|$ in discrete directional sectors, but is easily adapted for seasonal and local time sectors.

Given k non-intersecting directional sectors (or seasons, etc.), not necessarily of equal size, we determine a sufficiently high threshold, u_i in each sector, i. (We have initially tried $u_i = P_{99.97}$).

In each sector, we fit (by ML estimation) parameters of the complementary GP distribution of $\boldsymbol{x} = |dB_H/dt|$, conditional on exceeding u_i, given by

$$Pr(X_i > \boldsymbol{x} \mid X_i > u_i) = \left[1 + \frac{\xi_i(\boldsymbol{x} - u_i)}{\sigma_i}\right]_+^{-1/\xi_i}$$

Equation 3

where X_i denotes a measurement in sector *i*, and $\sigma_i > 0$. The variation of u_i, ξ_i, and σ_i for each site is plotted in **Fig. 7** *vs* direction sector and CGM latitude. In each sector, we determine the complementary GP distribution, conditional on exceeding the highest threshold, $u_{max} = \max(u_i : i = 1, \ldots, k)$,

$$Pr(X_i > \boldsymbol{x} \mid X_i > u_{max}) = \left[1 + \frac{\xi_i(\boldsymbol{x} - u_{max})}{\tilde{\sigma}_i}\right]_+^{-1/\xi_i}$$

Equation 4

where $\tilde{\sigma}_i = \sigma_i + \xi_i(u_{max} - u_i)$ is the 'modified' scale parameter (defined for $\tilde{\sigma}_i > 0$). The omnidirectional distribution may then be reconstructed as a weighted sum of complementary GP distributions in each sector. Return periods, $\boldsymbol{N}$ (years) may then be calculated for a range of return levels $\boldsymbol{x_N}$ from

$$\frac{1}{\boldsymbol{N}\, n_y} = \sum_{i=1}^{k} Pr(X_i > \boldsymbol{x_N})$$

$$= \sum_{i=1}^{k} Pr(X_i > \boldsymbol{x_N} \mid X_i > u_{max})\, Pr(X_i > u_{max} \mid X_i > u_i)\ Pr(X_i > u_i)$$

$$= \sum_{i=1}^{k} \left[1 + \frac{\xi_i(\boldsymbol{x_N} - u_{max})}{\tilde{\sigma}_i}\right]_+^{-1/\xi_i} \left[1 + \frac{\xi_i(u_{max} - u_i)}{\sigma_i}\right]_+^{-1/\xi_i} \frac{n_{c,i}}{n}$$

Equation 5

where $n_{c,i}$ is the total number of cluster peaks in the i^{th} sector, and n is the total number of samples. The resulting return period *vs* RL plot for GUA is presented in **Fig. 8**, where the bold line (—) represents the omnidirectional profile (Eqn 5) reconstructed from directional sector distributions (dot-dashed lines) (from Eqn 3). The return levels are significantly smaller than those obtained by Eqn 2 fitting to all data regardless of direction (- - -). This is typical for sites below 40° CGM latitude as illustrated by **Fig. 9a**, which compares 100-year return levels for all sites using the directional sector reconstruction method (Eqn 5) (blue □) with those obtained using Eqn 2 ignoring direction (+). The differences, presented in **Fig 9b**, indicate that reconstructing the RL from directional sectors often leads to higher RL value estimates at mid- to high latitudes.

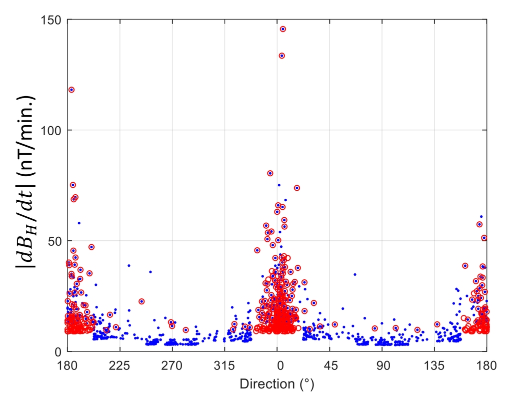

Fig. 6. Illustrating the effect of varying threshold by directional sector at a low latitude site (GUA).
o = cluster peaks ignoring direction;
· = cluster peaks in discrete sectors of width 45°.

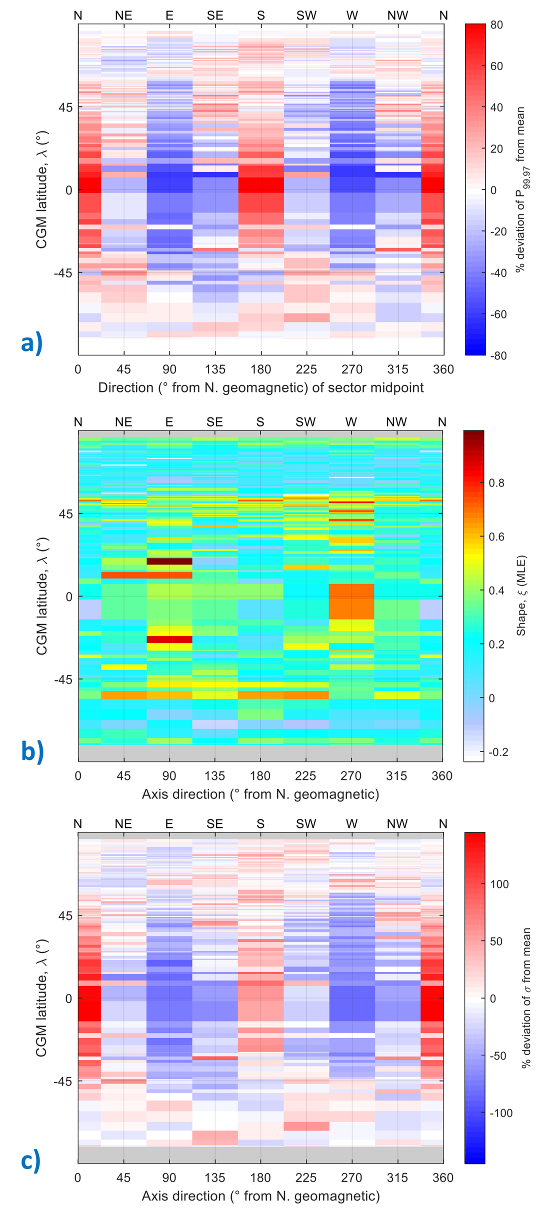

Fig. 7. Fitted GP parameters *vs* CGM latitude and direction sector. a) % deviation of $P_{99.97}$ from the mean (for each site), b) shape, ξ_i, c) % deviation of scale, σ_i.

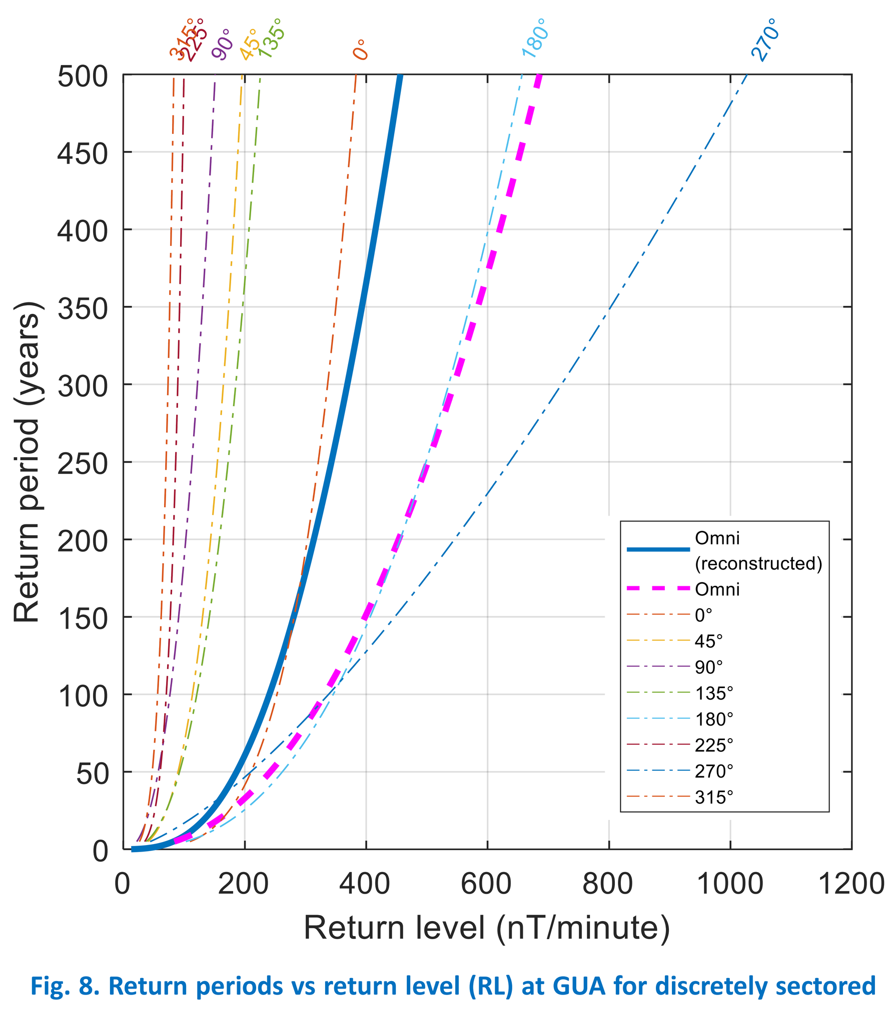

Fig. 8. Return periods vs return level (RL) at GUA for discretely sectored data (· – lines). The solid line indicates the omnidirectional profile reconstructed from discrete directional sectors. The magenta dashed line shows the MLE RLs for the GPD ignoring directionality.

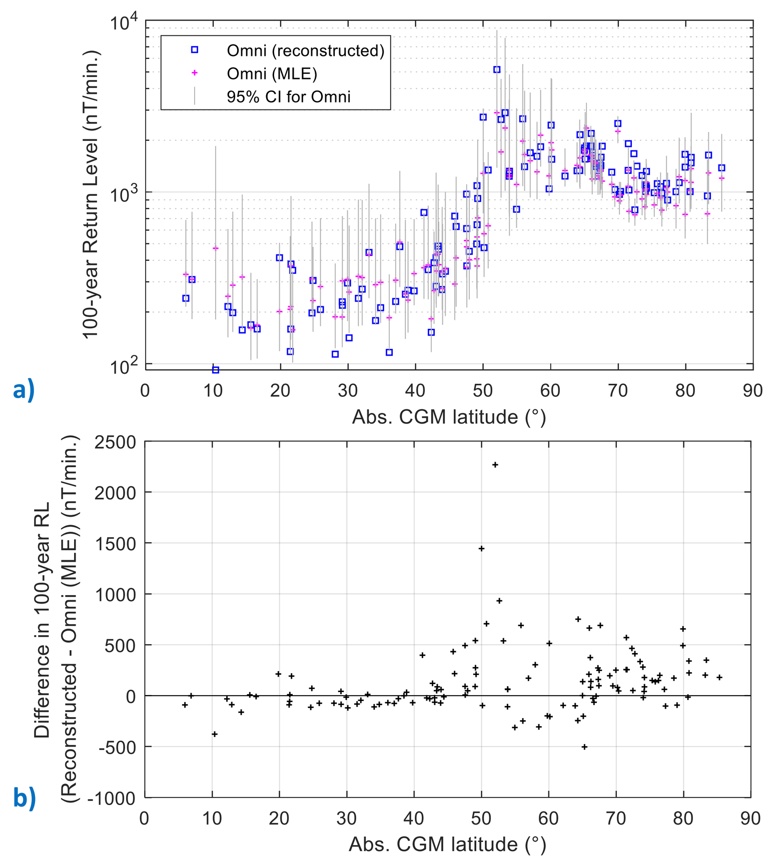

Fig. 9. a) 100-year return level predictions vs absolute CGM latitude comparing results reconstructed from eight 45° directional bins (blue) (Eqn 5), with those ignoring direction (magenta) (Eqn 2). b) Differences in RL estimates for the two methods.

6: Summary

- Predictions of extreme geomagnetic fluctuations are an important indicator of GIC risk.
- Using an archive of magnetograms from 125 sites worldwide, we predict the largest return levels for $|dB_H/dt|$ occur at 53° CGM latitude (N and S) with a secondary peak near the geomagnetic poles.
- Occurrence probability strongly depends on latitude, local time, month and direction, and is influenced by IMF orientation.
- Occurrence patterns match the known patterns of substorm expansions, Pc5 waves, sudden commencements, and high-latitude lobe reconnection events.
- The poor assumption of identically distributed peaks is addressed by fitting GP distributions to data discretely sectored by direction (or month, MLT, etc.).
- Return levels for GP distributions reconstructed from directionally sectored data yields lower return levels for absolute CGM latitudes below 40° where there are strong directional anisotropies, but higher return levels above 40°.

7: Future Work

- An alternative to the discrete sectoring approach requires fitting GP parameters as continuous functions of covariates (direction, month, MLT, etc.). The predictive performance of these methods will be compared.
- We have observed that $|dB_H/dt|$ occurrence rates and magnitudes are strongly dependent on the time-scale (dt), particularly at lower latitudes. Future models will incorporate the full spectrum of temporal fluctuations of interest to GIC risk modelling.

References

[1] Coles S. (2001), An Introduction to Statistical Modeling of Extreme Values. London: Springer.
[2] Gjerloev, J. W. (2009), A Global Ground-Based Magnetometer Initiative, *Eos, 90*, 230-231.
[3] Gjerloev, J. W. (2012), The SuperMAG data processing technique, *J. Geophys. Res., 117*, A09213.
[4] MacKay et al. (2010), On the use of discrete seasonal and directional models for the estimation of extreme wave conditions, Ocean Eng. 37, 425-442.

Acknowledgements

Magnetometer data and the substorm list were provided by SuperMAG and we gratefully acknowledge contributions from the SuperMAG collaborators: (Listed at http://supermag.jhuapl.edu/info/?page=acknowledgement.)

This work was funded by the **Natural Environment Research Council**, UK grant number NE/P016715/1.