British Geological Survey
NATURAL ENVIRONMENT RESEARCH COUNCIL

EXTREME VALUE STATISTICS APPLIED TO GEOELECTRIC ACTIVITY IN EUROPE *... a first look*

ORSI BAILLIE[1] [orba@bgs.ac.uk], Alan Thomson[1], Tamás Nagy[2], Ewan Dawson[1] Gemma Kelly[1], Viktor Wesztergom[2] and Ellen Clarke[1]

[1]British Geological Survey, West Mains Road, Edinburgh, United Kingdom
[2]Geodetic and Geophysical Institute, Sopron, Hungary

1. Introduction

Since 1957 the Nagycenk Széchenyi István Geophysical Observatory (IAGA code: NCK), Hungary, has provided continuous Earth current and geomagnetic observations. NCK lies on thick conductive sediment and is situated within a National Park, which helps reduce the effects of man-made electromagnetic noise.

Instrument house at Nagycenk observatory

Potential differences are measured using 2m deep, low polarization, lead electrodes in the North-South (*Ex*) and East-West (*Ey*) directions, with an electrode spacing of 500 m and recorded at 1 sec and 10 sec sampling intervals. Data resolution is about 6.1 μV/km.

Given such a long continuous measurement record, the NCK data are ideal for estimating electric field extremes that could be observed at ground level due to space weather. Here we provide an initial assessment of these extreme values using extreme value statistical analysis.

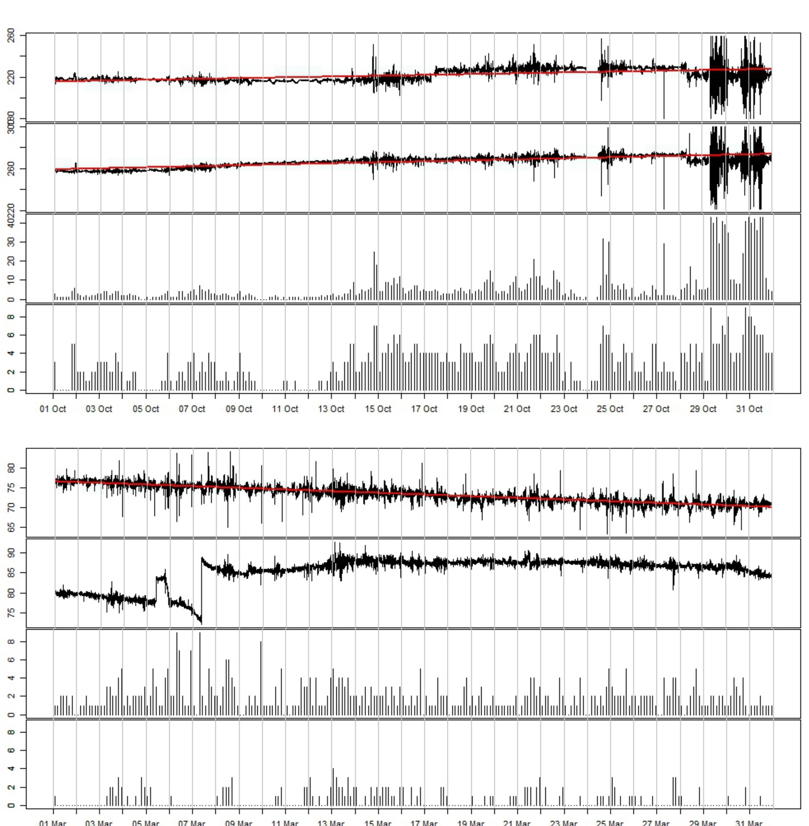
2. Extreme Value Statistics - Theory

We use a Generalised Pareto Distribution (GPD) to describe the tail of the distribution of geoelectric activity (e.g. Coles, 2004). An equivalent analysis (including method description) for European observatory magnetic data is given in Thomson *et al.* (2011). There are subtleties in applying extreme value statistics to geoelectric data, e.g. the need to de-cluster sequences of increased electric field activity relating to the same magnetic storms, and, in general, seasonal variations, jumps and steps in the geoelectric data. These subtleties are all taken into account in this study.

3. EVS - Data & Method

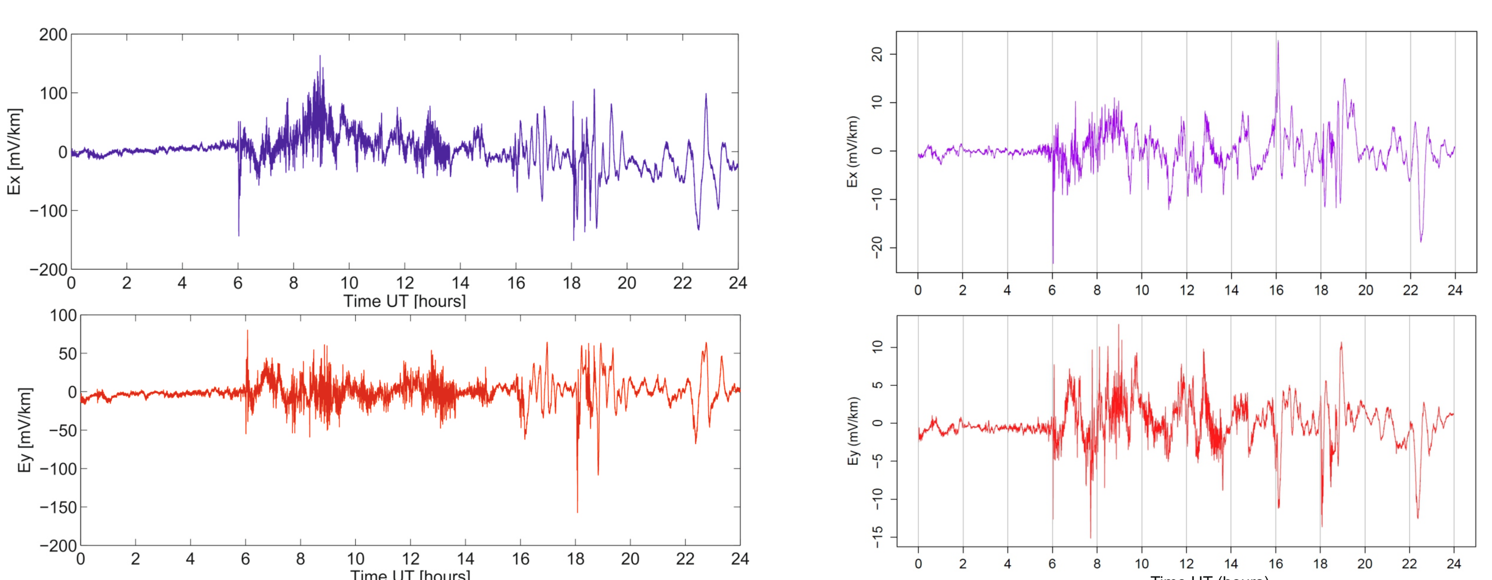
The following steps were used to process the data:

- Approximately 19 years (Feb 94 to Aug 13) of digital data were analysed
- These data were 10-second NCK geoelectric data in *Ex* and *Ey*, the 3-hour NCK *K* index (geomagnetic activity) and NCK *T* (telluric) index
- Data were pre-screened using local K index > 7, giving 106 days to analyse
- For each day a least squares fit was removed to leave just the variations
- Variations analysed using the eXtremes software toolkit (Gilleland and Katz, 2005) through the *R* statistical package (R Development Core Team, 2008)
- The maximum 10-second values per 3 hour time block were used as our basic data set, providing a manageable reduction in data size, whilst permitting a reasonably 'fine-grained' analysis. This reduces the data set from around 1 million to approximately 53,000 data points
- A second data set containing the maximum 10-second values per day were also analysed for comparison
- The maxima for both *Ex* and *Ey* were determined for the time-span of data. The projected GPD distribution for periods of 100 and 200 years were computed. The 95% confidence levels in the extremes were also determined



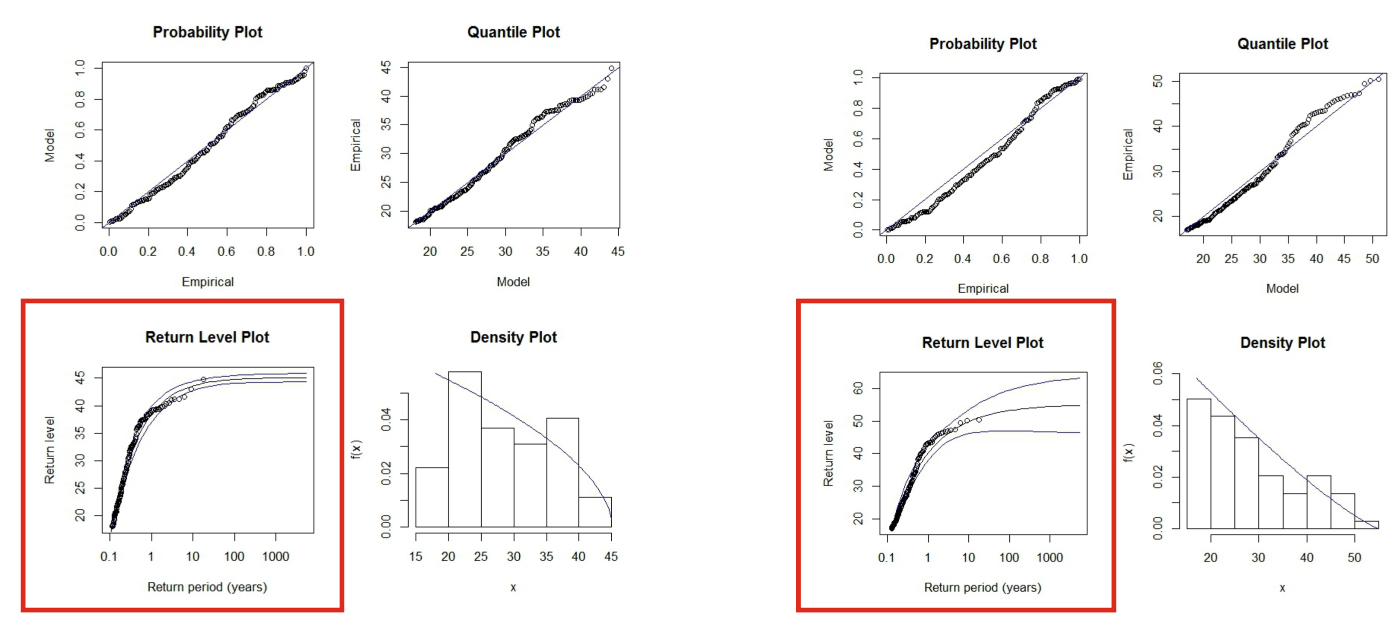
Two examples of geoelectric data for NCK are given in the Figures to the left. Plots of one month of data for a 'good' month (top) and a 'less good' month (bottom) are shown. Steps and spikes are clearly visible in the data and show the need for careful quality control. We note that the *E*-field can be relatively large even when the observatory *K* is only moderately disturbed.

In the Figure below left we show a daily plot of Eskdalemuir (Scotland) electric field data and to the right we show data from the same day at Nagycenk. The magnetic storm that occurred on this day is clearly visible in both data sets but the Eskdalemuir data are much larger. Applying EVS at both Eskdalemuir and Nagycenk should therefore yield quite different results.



4. EVS - Results

For each component we have extracted the Nagycenk peak variation predicted by the GPD to be exceeded for return periods of 100 and 200 years. To do this, appropriate geoelectric activity thresholds were determined separately for each component (threshold set at ~33% of peak observed *E*-fields for this data set).The results are summarised below (*Ex* below-left; *Ey* below-right).

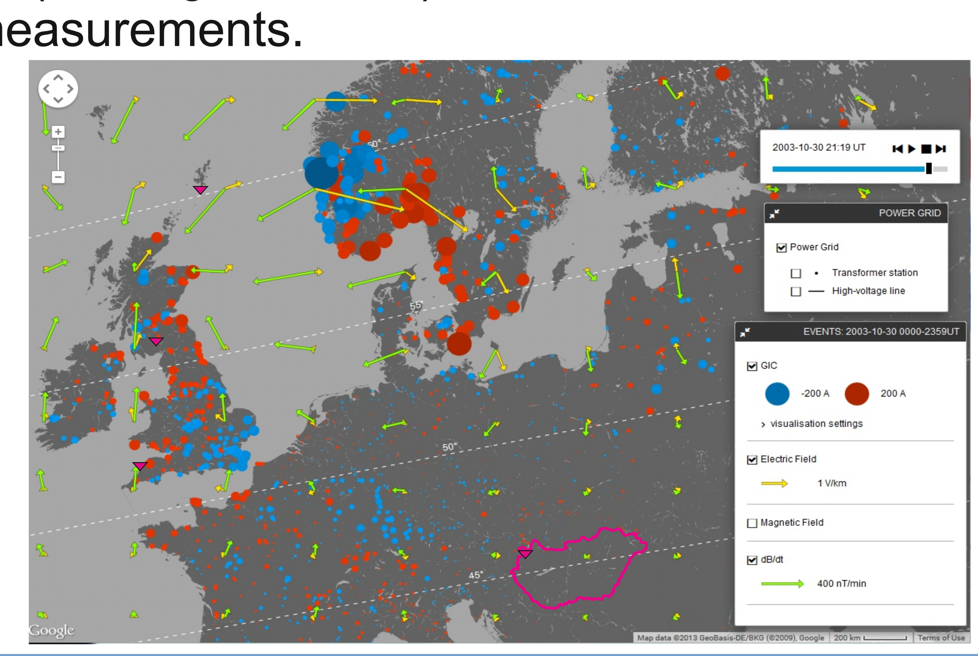


Output from extRremes analysis for NCK *Ex* (left) and *Ey* (right). The return level plot (red box) shows the return level, in mV/km, expected as a function of return period. Circles denote observed values.

Interestingly, the predicted extreme values for 100 and 200 year return periods are at levels already observed, something that suggests further study. The return level curves (particularly for *Ex*) shows signs of 'saturating', which is quite unlike geomagnetic data (e.g. Thomson *et al*, 2011). We also note a sensitivity of the results to threshold level, which may be a consequence of the pre-selection *K*-index method. This will be tested in future work.

Independent *E*-field model results for central Europe indicate that even during strong storms, such as October 2003, geoelectric field values are only fractions of a V/km (see Figure below). Such models are therefore validated by the NCK measurements.

Modelled *E*-fields, d*B*/dt and estimated geomagnetically induced current levels at 21:19 on 30th October 2003. Snapshot courtesy of EURISGIC project colleagues (www.eurisgic.eu).Triangles indicate locations of geoelectric monitoring sites.



5. Discussion & Future Work

What have we found?

- Both the 100 and 200 year return periods for NCK geoelectric data are about 45 mV/km and 54 mV/km for *Ex* and *Ey* respectively.
- The return level curves tend to 'saturate' meaning that such 'extremes' are already in the data set. This is unexpected and indicates that further work is required
- Using daily maxima of 10-second data produces similar results
- There is a sensitivity of the results to the choice of threshold

What will we do next?

- Repeat the analysis by including days when the local 3 hour geomagnetic *K* index is 6 (adding ~500 days more), or through better automatic correction of baselines, jumps and spikes
- Repeat the analysis with recently digitised (quality checked) telluric data for more of the 50 years on record
- Re-compute NCK *T* index (local telluric index) with unbounded upper limit and repeat the EVS analysis on this data set
- Compare results with similar analyses of Eskdalemuir and other UK stations

Acknowledgements:

The research leading to these results was part of the EURISGIC consortium project, having received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no 260330.

We would like to thank colleagues at GGI and Nagycenk Observatory, Hungary, for providing their data. 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. A04204, doi:10.1029/2006JA011900
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.

BGS Electric Field website:
http://www.geomag.bgs.ac.uk/research/electric_field.html

EURISGIC