Natural Resources Canada   Ressources naturelles Canada

GEOLOGICAL SURVEY OF CANADA
OPEN FILE 7905

Extreme Values Statistical Assessment for Geomagnetic and Geoelectic Field Variations for Alberta

L.Nikitina and L.Trichtchenko

2015

Canada

GEOLOGICAL SURVEY OF CANADA
OPEN FILE 7905

Extreme Values Statistical Assessment for Geomagnetic and Geoelectic Field Variations for Alberta

L.Nikitina and L.Trichtchenko

2015

doi:10.4095/296956

Recommended citation
Nikitina, L. and Trichtchenko, L., 2015. Extreme values statistical assessment for geomagnetic and geoelectric field variations for Alberta; Geological Survey of Canada, Open File 7905, 23 p. doi:10.4095/296956

Contents

1 Summary

Disturbances of the geomagnetic field produced by space weather events can have an impact on power systems and other critical infrastructure. This study aims to provide an understanding of possible geomagnetic effects on power systems in Alberta. Extreme value statistics has been applied to 40 years of magnetic data from three Canadian magnetic observatories to estimate maxima of the geomagnetic variations and horizontal geoelectric fields

- once per 50 years;
- once per 100 years.

This approach to extreme value estimation was tested by using of 16 years of data to estimate 40-years maximum of magnetic variations. Comparison with the actual maximum for the complete 40-year database showed an error of only 4%.

The estimated values can be used with a power network model to assess geomagnetically induced currents in the Alberta power network.

2 Introduction

Disturbances of the geomagnetic field produced by space weather events cause variable geoelectric field at Earth's surface which drive electric currents called geomagnetically induced currents GIC. These GIC produce effects in power systems, for example transformer saturation resulting in production of harmonics and increased reactive power demand that can cause misoperation of protective relays, voltage sags and damage to equipment. In extreme cases, as during the magnetic storm in March 13, 1989, this can result in burnt-out transformers and system collapse.

This study aims to provide understanding of possible extreme geomagnetic effects on power systems, pipelines and other ground critical infrastructure in Alberta Province. Analysis was started with a statistical evaluation of occurrence of geomagnetic and geoelectric activity in Alberta [Trichtchenko et al., 2015]. In the present study this analyses was extended to estimate extreme scenarios for geomagnetic and geoelectric disturbances, once in 50 years value and once in 100 years value.

Extreme events are of special interests in many areas of natural and human sciences. Extreme value statistics have been applied to estimate financial risks as well as risk of natural catastrophes like extreme flooding, wind speed, and precipitation levels. Historically, statistical analysis has been mainly concentrated on understanding the average behavior of physical, biological, and social systems, but in some cases it is more important to study extreme events. Extreme events happen rarely, but can have significant impact on life and property. Extreme value analysis is used to provide a hazard assessment which can be used to develop the system design which could mitigate extreme events.

Long recordings of geomagnetic data provided by Canadian geomagnetic observatories were used for the extreme value statistical analysis for Alberta. The geomagnetic activity is greatest in the auroral zone, where most of the province is located. This study used almost 40 years of geomagnetic data recorded in Meanook (MEA) observatory, the only observatory which is located in Alberta. To estimate the spatial variability of the geomagnetic disturbances, the data from the observatory available to the north of the provincial border (Yellowknife, YKC) were used. 40 years of recordings in Ottawa (OTT) were used for comparison.

In Trichtchenko et al. (2015), the 95%, 99% and maxima annual values for geomagnetic hourly ranges were estimated and it was shown that time variations of maxima do not repeat time variations of 95% and 99% values. To check if the maximum values are correlated with 95% or 99% percentiles, these values are plotted on Figure 1a and 1b. Correlation between 95% and 99% values for Yellowknife data is high, the correlation coefficient is r=0.968 (see Figure 1a). This means that these two values are statistically dependent. But maximum values are not well correlated neither with 95% nor with 99% values. This is demonstrated by Figure 1b where the annual maxima of geomagnetic hourly ranges at Yellowknife are plotted as a function of the annual 99% values.

Correlation is weak, the correlation coefficient is r=0.462 and it means that the 99% values cannot be used for extreme values estimation. This result shows that it is necessary to use the special theory of extreme value statistics to evaluate the size of events expected once in 50 years and once in 100 years.

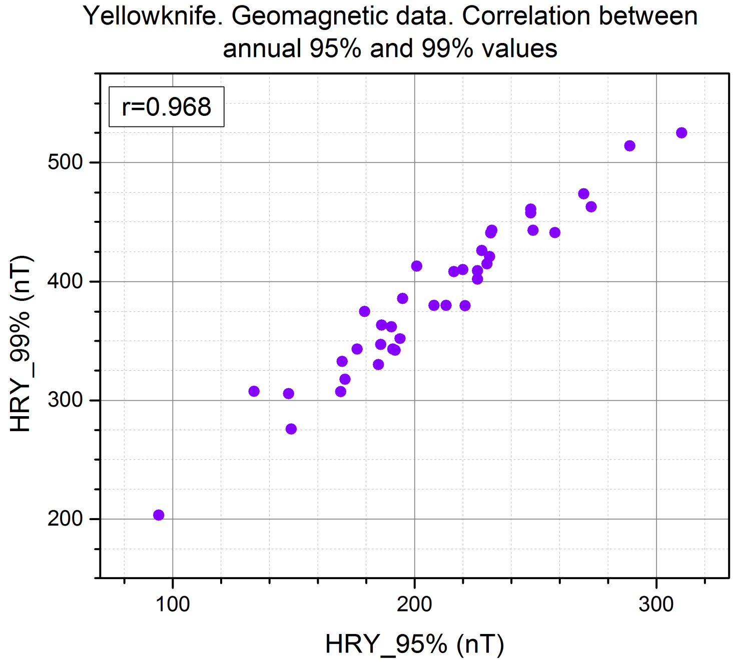


Figure1a. Correlation between 95% and 99% annual values for geomagnetic hourly ranges. Yellowknife data, 1975-2012. Correlation coefficient r=0.968.

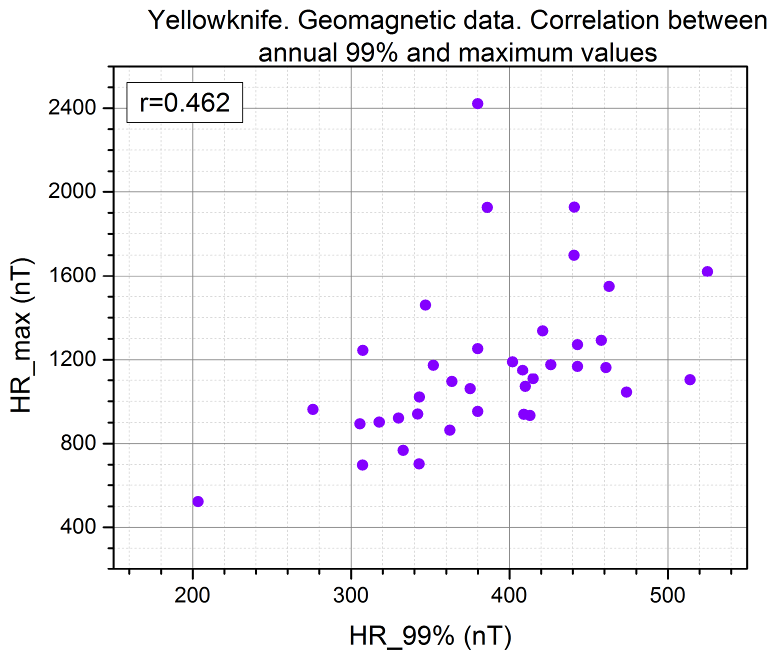


Figure 1b. Correlation between hourly maxima of hourly ranges and 99% annual values of magnetic hourly ranges. Yellowknife, 1975-2012. Correlation coefficient r=0.462.

2. Theoretical background

Extreme value theory is focused on the behavior of extreme values that can appear in series of data. In this study the extreme value theory is used for evaluation of extreme values in geomagnetic hourly ranges and geoelectric hourly maxima.

Extreme value theory (Coles, 2004) is concerned with the statistical behavior of maxima of big sets of data. This theory provides a statistical distributions of M_n where

$$M_n=\max\{X_1, X_2, X_3, \ldots, X_n\} \tag{1}$$

and X_1, .., X_n is a sequence of independent observations of the same physical parameter. In our case X_i are geomagnetic hourly ranges or geoelectric hourly maximum. A key part of extreme value theory is the theorem of Fisher-Tippett-Gnedenko which proves that the limit distribution for maxima does not depend on the initial distribution of X values. When n is enough large, the maximum M_n (under some conditions) has one of three possible distributions. These three classes of distributions are termed Gumbel, Fréchet and Weibull distributions. The cumulative probability p for these distributions can be written in the form

$$p = \exp\left(-\left(\frac{x}{s}\right)^{-\alpha}\right) \text{ for Fréchet distribution;} \tag{2}$$

$$p = 1 - \exp\left(-\left(\frac{x}{s}\right)^{\alpha}\right) \text{ for Weibull distribution;} \tag{3}$$

$$p = \exp\left(-\exp\left(-\frac{x}{s}\right)\right) \text{ for Gumbel distribution;} \tag{4}$$

where S and α are constants which can be obtained by a fitting procedure. Gumbel distribution is applicable for the data which are unbounded from both sides. Geomagnetic variations are larger than zero, and for our analyses we use Fréchet or Weibull distributions. Some examples of these distributions are plotted on Figure 2. The probability density functions are on the left panel, and the right panel provides the cumulative probability function. Fréchet distribution is known as a distribution with a ‘heavy tail’ while Weibull distribution has a ‘lighter tail’ which means that the extreme values have larger probability for Fréchet distributions than for Weibull distribution. For our analyses we use the one of these two distributions which fits better to our data.

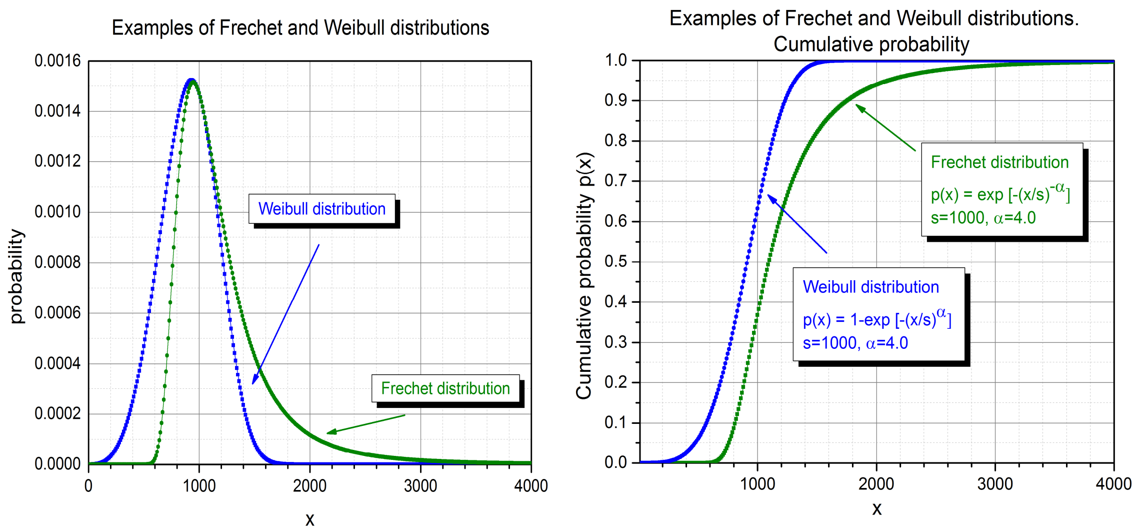


Figure 2. Examples of Weibull and Fréchet distributions. The probability density functions are on the left panel and the cumulative probability functions are on the right panel.

To fit data to one of these distributions, it is convenient to use special coordinates which transform these distributions to the straight lines (see Figure 3) and allow us to use the linear fitting procedure. We will use

$$H_F = -\ln(-\ln(p)) \quad \text{for Fréchet distribution;} \tag{5}$$

$$H_W = \ln(-\ln(1-p)) \text{ for Weibull distribution.} \tag{6}$$

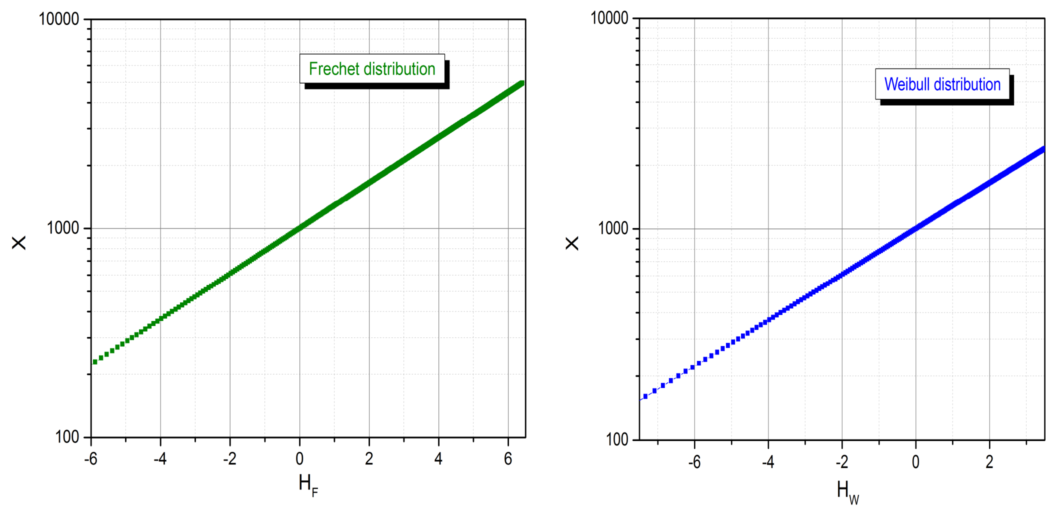


Figure 3. Transformation of Weibull and Fréchet distributions into straight lines with use of H_F and H_W coordinates

To find one event per 50 years or one event for 100 years the return period concept is used. The return period T is the time of recurrence of large events and can be calculated using a probability function

$$T = \frac{1}{1-p}. \quad (7)$$

The return period for an extreme value is an average time interval when this extreme value can occur or be exceeded. If, say, the return period for hourly range HR=1000 nT is 5 years, it means that during 50 years this value could be exceeded approximately 10 times. Examples of Fréchet and Weibull distributions in terms of the return period are plotted on Figure 4.

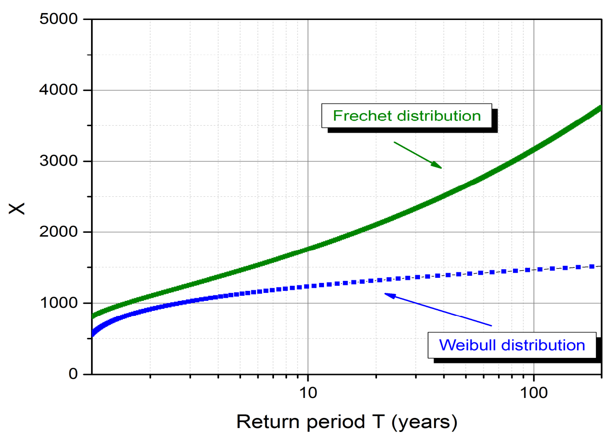


Figure 4. Fréchet and Weibull distributions plotted with use of the return period T.

3 Magnetic data preparation and fitting

Extreme value statistics is applicable for maxima of large sets of independent data. All the data for 40 years for three geomagnetic observatories (Meanook, Yelloknife and Ottawa) were separated into blocks each containing data for one month. The maximum values for each month were used for the extreme value analysis. Space weather event can last up to several days so by using monthly maxima we can be confident that the data are independent. And additionally the analysed maxima were checked to be sure that the time period between two consequent maxima is larger than at least 2-3 days and they are not related to the same event.

To fit maxima to one of the extreme value distributions, usually the values larger than some certain level were considered. As shown in [Thompson et al., 2011] this threshold approximately corresponds to 99.97% value of all the data for a given geomagnetic observatory. Thresholds for geomagnetic data at three geomagnetic observatories are given in Table 1.

Table 1. Thresholds for extreme values of the magnetic hourly ranges which are used for extreme value statistics

Geomagnetic observatory	Years	Threshold
MEA	1972-2012	1100 nT
YKC	1975-2012	1200 nT
OTT	1973-2012	650 nT

To demonstrate the fitting procedure we consider Meanook data (see Figure 5). Here we took maximum values above the threshold 1100 nT. To fit the data to an extreme value distribution we plotted the data in the coordinate system where the horizontal coordinate is the double logarithmic coordinate H_F which is used for Frechet distribution (equation 5) and the vertical coordinate is the logarithm of the hourly range of the magnetic variations. In these coordinates the data should fit to the straight line (see Fig. 5, left panel) and the linear fitting procedure is used to find the linear coefficients. These coefficients were used to convert the graph back to the cumulative probability (Fig. 5, right panel)

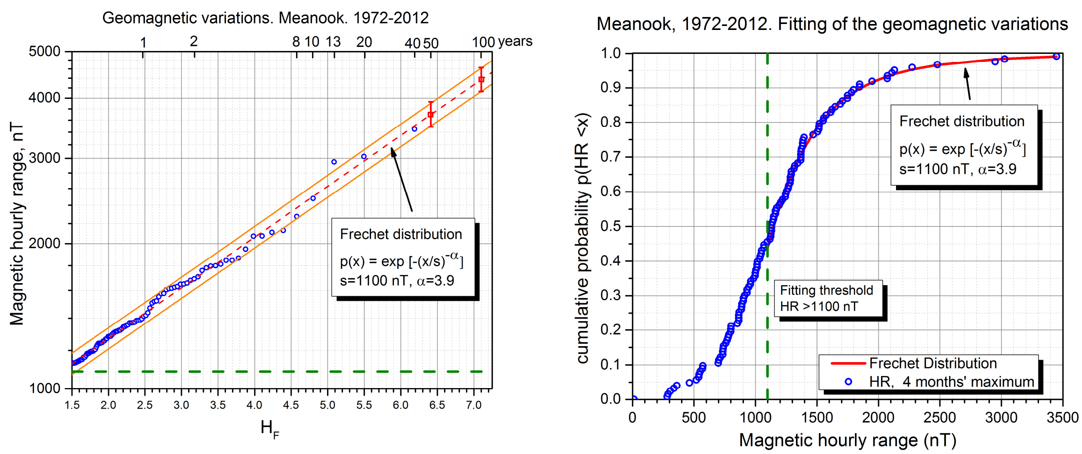


Figure 5. Fitting of Meanook data to Fréchet distribution. Blue circles are maxima of the Meanook data exceeding the threshold 1100 nT (green dashed line on both panels). On the left panel these maxima are fitted to the straight line (red line) in the Fréchet coordinate H_F (eq. 5). On the right panel the result of the fitting procedure is plotted. Blue circles lie on the Fréchet distribution line (red line).

4 Extreme values statistics for the geomagnetic variations. Results

The above mentioned approach has been applied to the geomagnetic activity range index for each observatory under investigation. In this study the geomagnetic activity index has been chosen as the largest per hour between HRX and HRY, therefore the directions of the geomagnetic variations are not considered. This has been done for two reasons, one is that the directional sensitivity of each specific station in the power grid might be different, i.e. the largest GIC might be due to X- component or Y-component or any other direction depending on the particular topology of the power grid. The second reason is that the Canadian Space Weather Forecast Centre is providing the forecasts of the local geomagnetic activity indices based on the same statistics, i.e. forecasted is only one value, which is the largest of the two horizontal components, without specific determination either it is in X- or Y-direction. Thus, in order to use these forecasts, only one component of the geomagnetic activity index is needed to be used.

The monthly maximum of the geomagnetic activity indices obtained from hourly values recordings at three stations, MEA, YKC and OTT, for period of time close to 40 years, were fitted to extreme value distributions. The better of two distributions was selected, and then the extrapolation has been done to estimate the extreme values with the return period of 50 and 100 years with 99% confidence interval. The results are presented in Figure 6.

On this figure, the left column of panels are the straight lines fitting the logarithm of hourly range (vertical coordinate) vs. the double logarithmic coordinates as described in the previous paragraph. The dashed line is the linear fit of the monthly maximum hourly range, the solid lines represent the 99% confidence interval and the red squares are the values corresponding to the once per 50 year and once per 100 year occurrences. The right set of panels is demonstrating the same results plotted in the different coordinate system. The vertical axis is the monthly maximum of the hourly range values, the horizontal axis is a return period in years. The red squares correspond to the return periods of 50 and 100 years. The obtained extrapolated numerical values corresponding to the once per 50 years and once per 100 years occurrences together with the 99% confidence intervals (CI) are listed in Table 2 and presented in Figure 7 for each observatory used in the study.

Table 2. Estimation for extreme values of the geomagnetic variations, once per 50 years, once per 100 years

Geomagnetic observatory	1 in 50 years value and 99% confidence interval, in nT	1 in 100 years value and 99% confidence interval, in nT
MEA	3695	4370
	[3500, 3900]	[4135, 4625]
YKC	2855	3085
	[2615, 3120]	[2820, 3375]
OTT	2625	2980
	[2310, 2990]	[2610, 3405]

The estimated extreme values with their 99% confident intervals, as shown in Figure 7 for each of geomagnetic observatory, demonstrate the following features. Maximum values recorded in the 40 years (brown boxes) are slightly less than the estimated values for 50 year repeat period (green boxes). Their confidence intervals practically touch the 40-year values and values for 100 year (red boxes) are the largest.

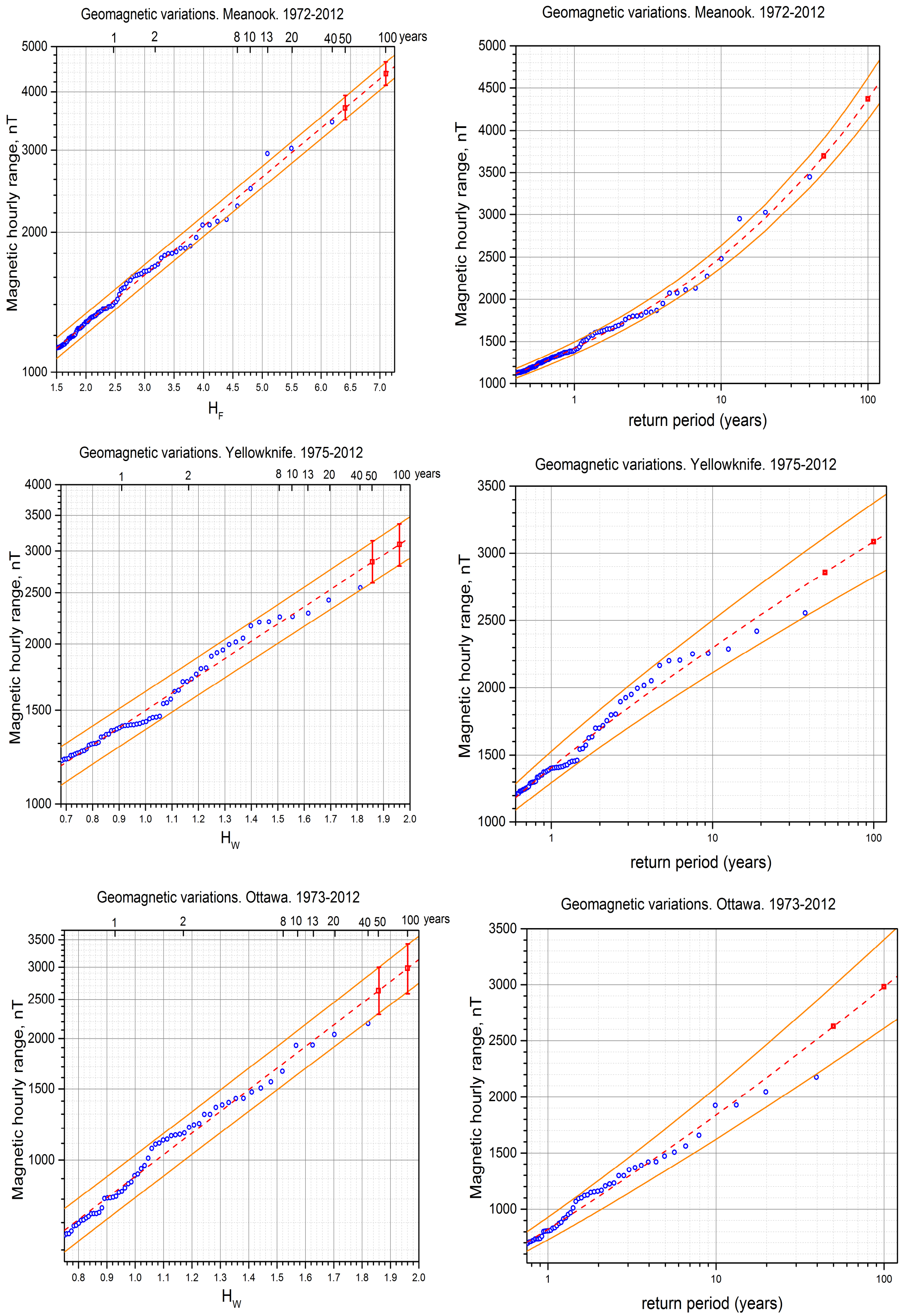


Figure 6. Monthly maximum of the hourly range of the magnetic variations are fitted to extreme value distributions. Left panels illustrate the linear fitting of the data. Right panels provides the return period for extreme magnetic hourly ranges.

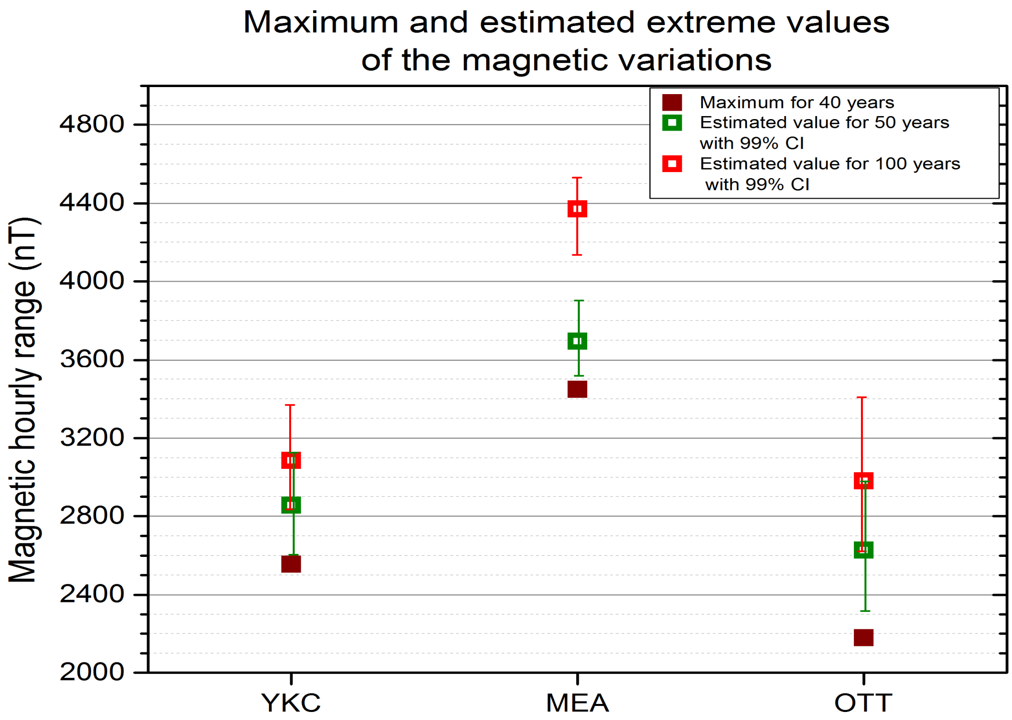


Figure 7. Estimated extreme values (50 years-green, 100 years-red) with 99% confidence intervals and recorded 40 years (brown) maxima of the geomagnetic hourly ranges at three stations.

The extremes in geomagnetic activity at YKC (auroral) observatory are more closely spaced and in general are less than values for MEA and are only slightly larger than the extreme values in Ottawa (sub-auroral location). This can be explained by the highly dynamical nature of geomagnetic disturbances, so that the maximum of the geomagnetic activity during the strongest geomagnetic storms moves to the location of MEA instead of the more statistically average location in the high auroral zone (YKC). This can be inferred also from Figure 1.8 of [Trichtchenko et al., 2015] where the annual median and mean values for MEA are smaller than for YKC, while 95% and 99% values are close or sometimes larger than in YKC. It means that in Meanook the regular climatological pattern is largely destroyed in the case of extreme values.

More detailed comparative analysis of the geomagnetic activity at these three locations can be done based on Figure 8, which shows the dependence of the extreme values on the return period. It is clearly seen, that for the smaller return periods (less than 8 years) the extreme geomagnetic activity is the same at YKC and MEA, while for return period of above 10 years the extremes in geomagnetic activity in MEA are higher and stay higher for estimated values with return periods of 50 and 100 years. At the same time, the extreme geomagnetic activity in OTT estimated for 100 year return period is becoming close in estimation to that of YKC.

It should be noted, that the above conclusions have to be regarded as preliminary, based on the statistical studies of 40 years of data and only for these three stations. To compare extreme magnetic activity in auroral and sub-auroral zones, data from other magnetic observatories should be included in this analysis.

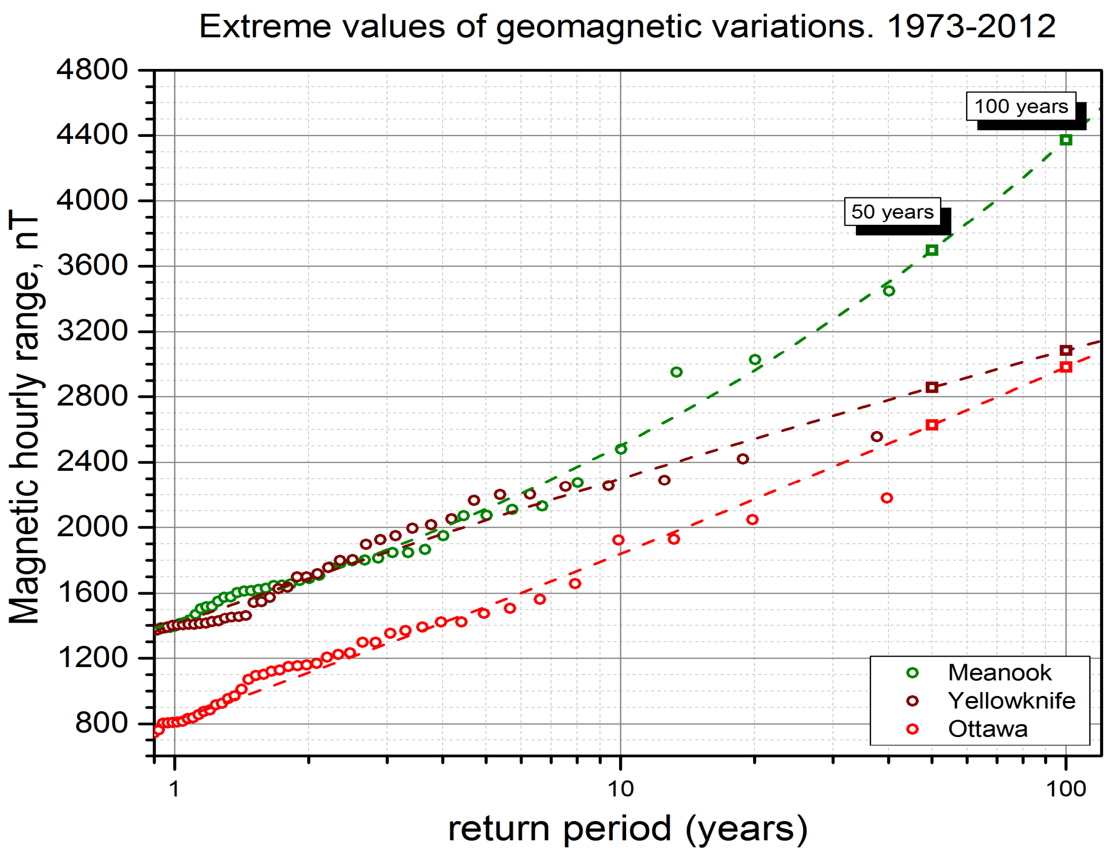


Figure 8. The recorded extreme values for geomagnetic hourly ranges at three observatories versus return period. Once in 50 years and once in 100 years estimated values are shown as squares.

5 Verification test

The extreme value approach was tested by comparing the estimated extreme values with actual recordings for one observatory, Yellowknife. Based on the data recordings in 1990-2005 years (16 years), the predicted extreme value for 40 years has been calculated within the 99% confidence interval and has been compared with the actual measured value. As it is shown in Figure 9, the predicted 40-years value is 2655 nT (green circle) with the 99% confidence interval [2415, 2920] (yellow lines). The observed maximum value of the geomagnetic variations in 1975-2012 was in 1985, i.e. in the year which was not included in the data for this estimate, and the measured maximum value is 2555 nT (red dot). This value is inside the 99% confidence interval, and the relative error of the forecast is 4%.

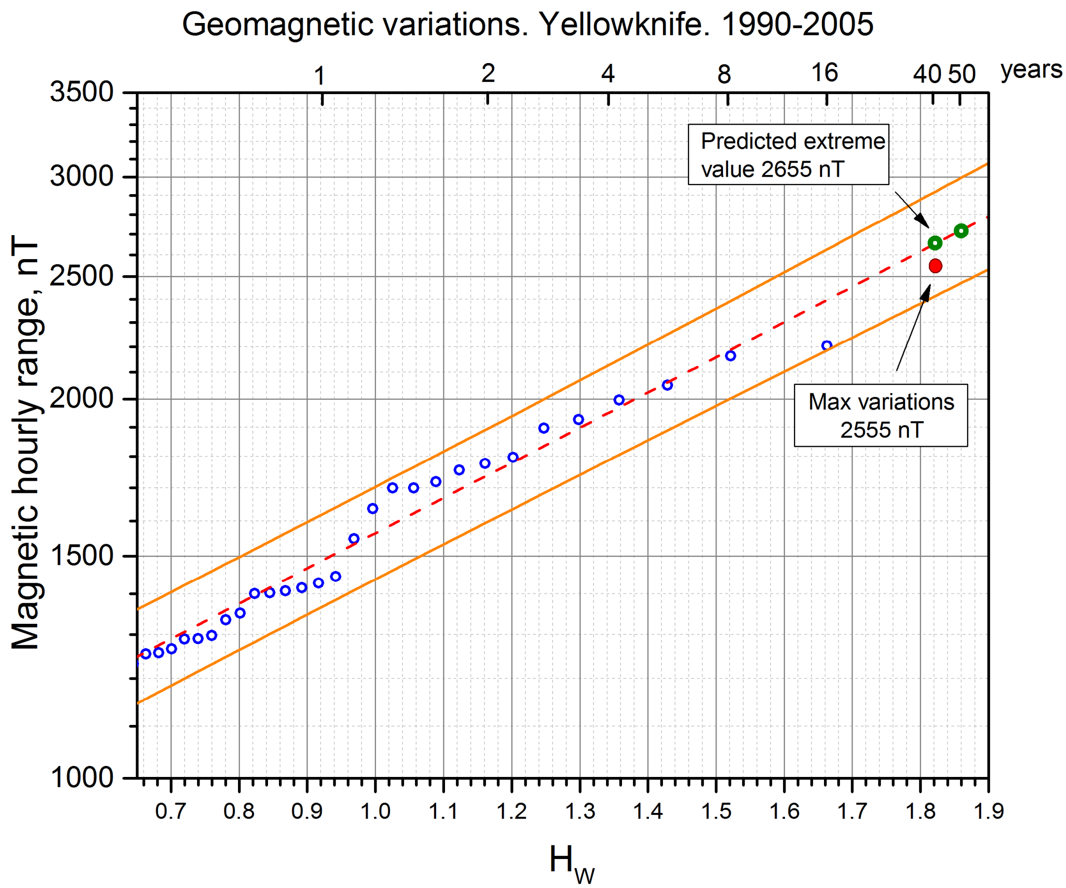


Figure 9. Verification test example. Data for 16 years (1990-2005) of the geomagnetic hourly ranges for Yellowknife are shown as blue circles, the best fit is plotted by red dashed line and 99% confidence interval are yellow lines. Green points are the predicted values for once per 40 and once per 50 years events. Red point is the maximum for all the data measured at Yellowknife for 40 years.

6 Extreme values of geoelectric activity

The same approach has been applied to the hourly maximum index of the modelled geoelectric field values. The geoelectric field has been calculated from MEA,YKC and OTT observatories data using the procedure described in Trichtchenko et al., (2015):

1. Conversion of the geomagnetic data from time into frequency domain using Fast Fourier Transform (FFT);
2. Multiplication by the surface impedance, obtained from one-dimensional resistivity profile of particular area;
3. Inverse transform of geo-electric spectrum into time domain by using inverse FFT.

The surface impedances for the Earth used in step 2 were calculated based on 10 layered earth conductivity models provided in Trichtchenko et al., (2015), Chapter 2. Sequence of steps is illustrated by scheme (Fig. 10).

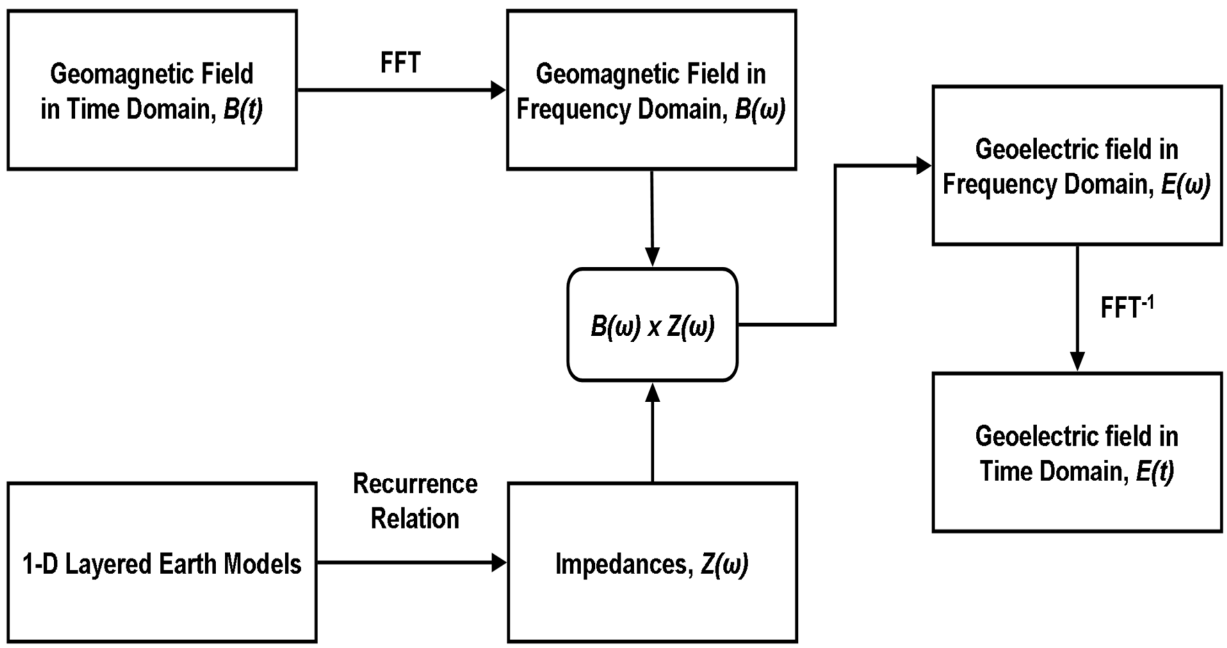


Figure 10. Calculation of the electric field from the geomagnetic field and layered earth model

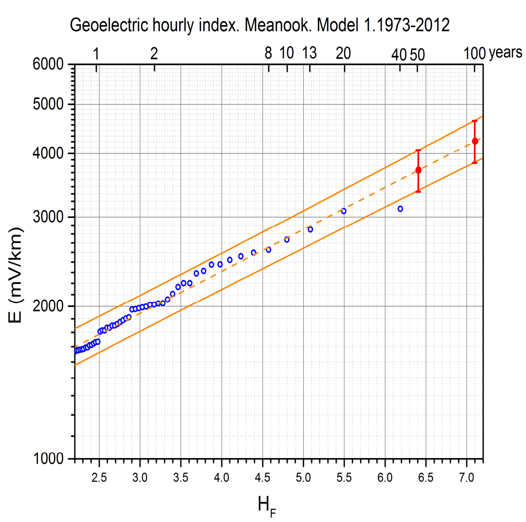
In the procedure in Figure 10 minute data for X and Y components of the magnetic field are used to model Y and X components of the geoelectric field correspondingly. The total horizontal component is derived from X and Y components. The hourly maximum of the geoelectric field has been chosen as a geoelectric activity index, therefore the direction of the geoelectric field is not considered. This has been done because the directional sensitivity of each specific station in the power grid might be different, i.e. the largest GIC might be due to X- component or Y-component or any other direction depending on the particular topology of the power grid.

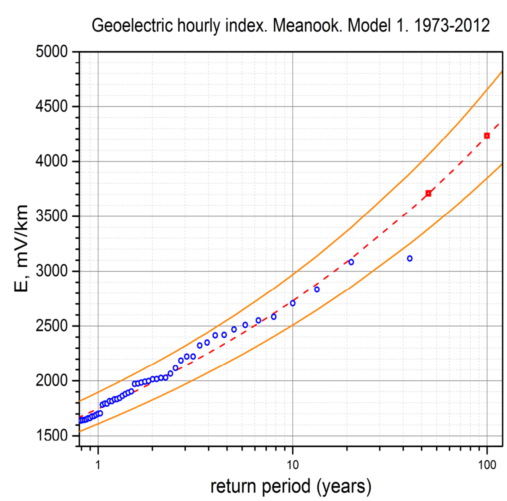
40 years of geomagnetic data at MEA, YKC and OTT were used for geoelectic field modeling. Monthly maxima of absolute value of geoelectric field were taken for statistical analyses and were fitted to one of the extreme value distributions. Then the extrapolation has been done to estimate the extreme values with the return period of 50 and 100 years with 99% confidence interval. Several examples are presented in Figure 11.

In Figure 11, the left column of panels are the straight line fitting of the monthly maximum of the geoelectric field index (mV/km) plotted in logarithmic scale (vertical coordinate) and double logarithmical coordinate corresponded to the cumulative probability, as described in paragraph 2. The dashed line is the linear fit of the monthly maximum, the solid lines represent the 99% confidence interval and the red squares are the values corresponding to the once per 50 year and once per 100 year occurrences. For example, the top left plot demonstrates the geoelectric field modeled with use of MEA geomagnetic data and surface impedance model 1. In the second and third plots the fit for extreme geoelectric fields modeled by using YKC magnetic data with surface impedances model 7 and model 9 are shown.

The bottom plot shows values for geoelectric field modeled using Ottawa magnetic data with Ottawa surface impedance. It is interesting to note, that the highest value for the 99% confidence interval of the estimated extreme value for YKC with model 7 and MEA with model 1 are close to each other, while OTT shows a reduction in values.

The plots located in the right hand side column of Figure 11 represent the same extreme values re-plotted in the coordinate system of maximum geolelectric field values (in mV/km) versus return period (in years). It is clearly seen, that while for MEA the estimated value for the 50 year return period together with its confidence interval is higher than the existing value for 40 years maximum, for the YKC and surface model 7 the lowest value of confidence interval for once in 50 years value is close to the existing maximum value of the electric field, and for the surface impedance model 9 it is even less than the existing maximum value for 40 years.





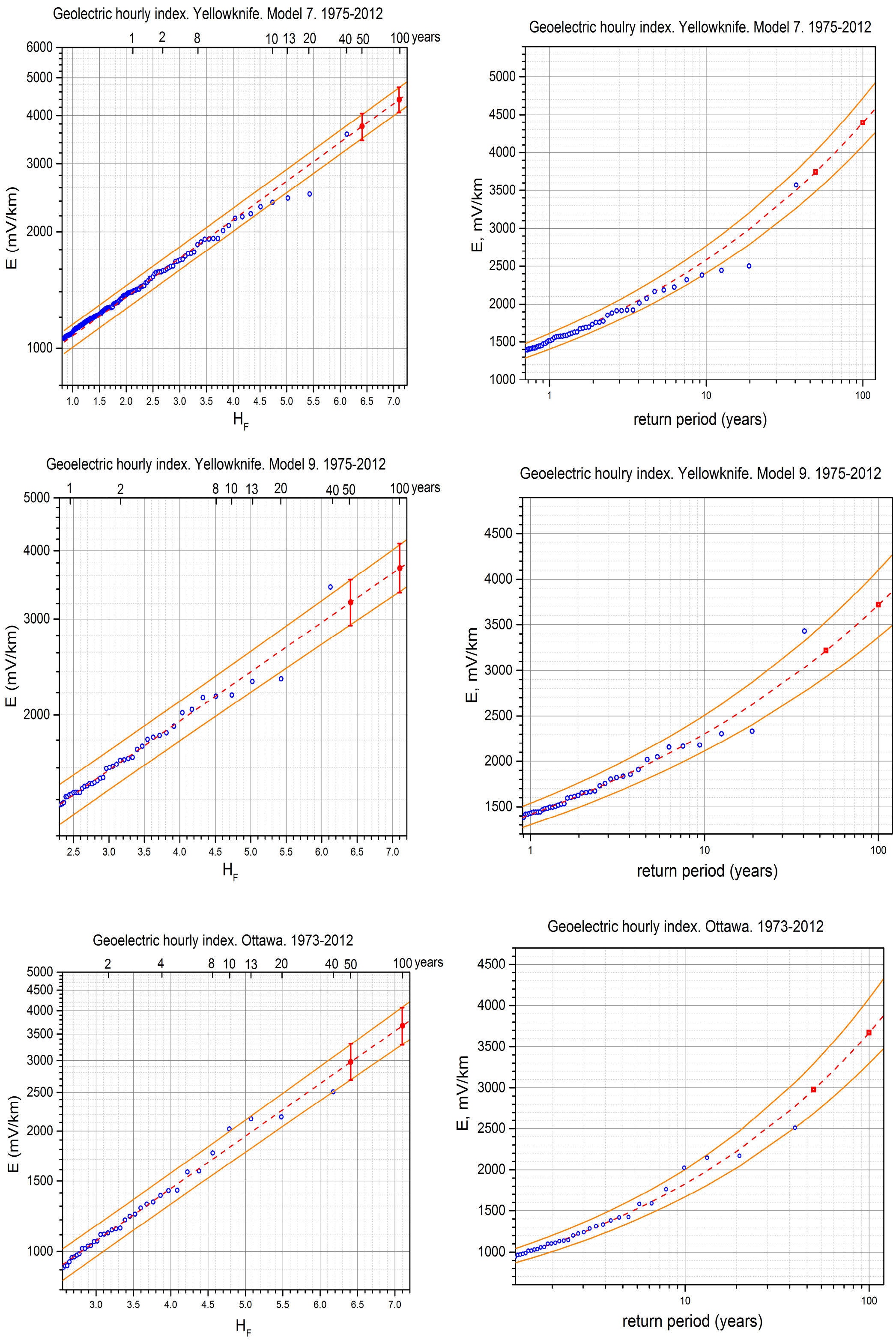


Figure 11. Example of fitting geoelectic data. Left panel – linear fitting to the extreme value distribution; right panel – return period for the extreme values in geoelectric field

In more details the impact of different earth conductivity models is presented in Figures 12-14. As can be seen from Figure 12a), the amplitude of the surface impedance is the largest for model 1, lowest for model 3 and we choose model 5 as the intermediate value. The corresponding modeled extreme electric field hourly indices (with MEA geomagnetic) and forecasted for 50 and 100 years return periods are plotted in Figure 12 b). It is clear, that the higher the amplitude of the surface impedance, the larger the extreme values of the geoelectric field.

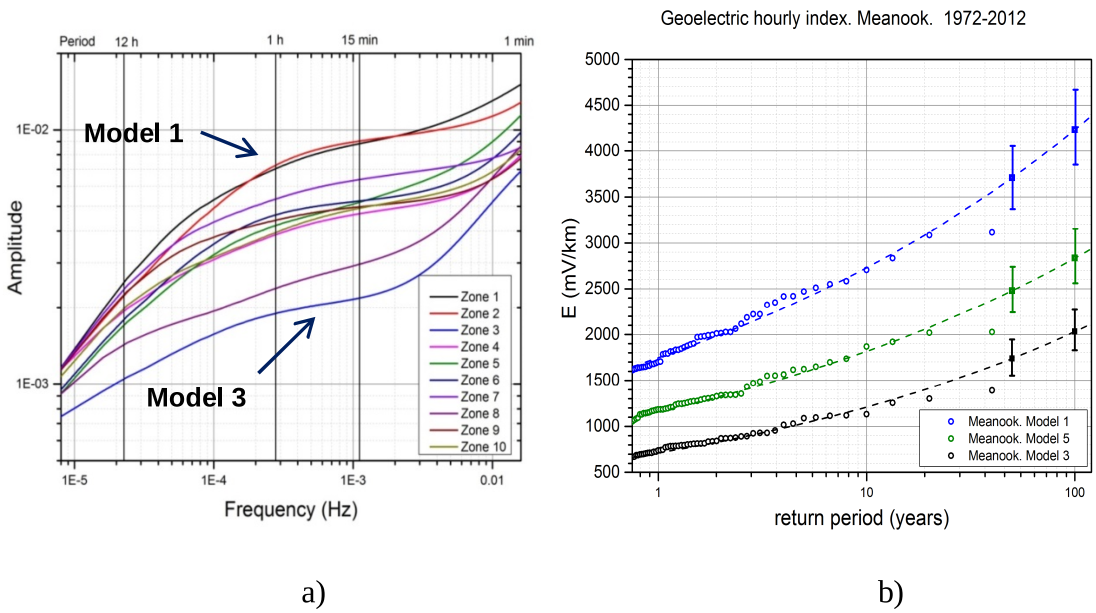
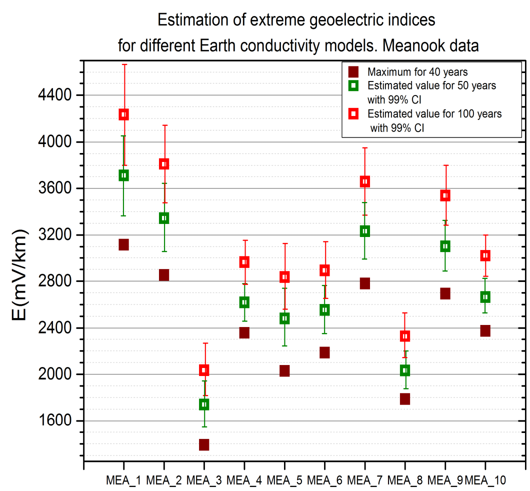


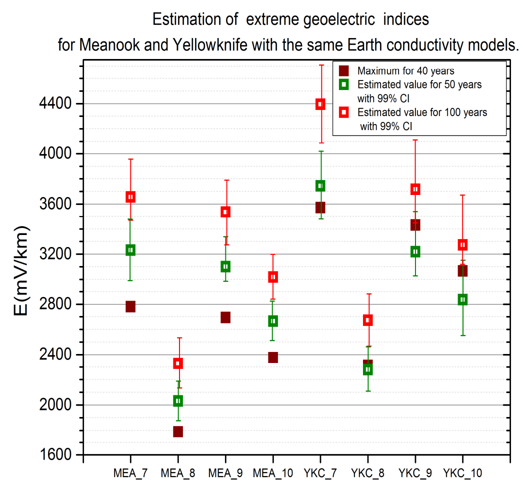
Figure 12. Impact of surface impedance models on the estimation of the extreme values. a) the amplitude of the surface impedance models (see Chapter 2 in Trichtchenko et al., (2015); b) extreme geoelectric values for 3 chosen surface impedances with MEA geomagnetic data.

The results of the calculated geoelectric field for the same magnetic and different surface impedance models are illustrated by Figure 13. Figure 13a) demonstrates the extreme values of the geoelectric field which were calculated for the same geomagnetic field data from Meanook observatory and 10 different impedance models. Figure 13b) is plotted for two sets of geomagnetic data (for Meanook and Yelloknife observatories) and for models 7-10 of surface impendance. On these graphs the 40 years extreme values of the geoelectric field as well as estimated for return periods of 50 and 100 years with their confidence intervals are presented.

The impact of surface impedance (Fig. 13a) is clearly corresponding to its relative amplitude, i.e. models 3 and 8 (with the smallest amplitude of surface impedance) provide the smallest values of geoelectric field, models 1, 2 and 7 with the largest impedance give the largest values of geoelectric field, as high as twice larger. As well, from plots presented in Figure 13 b) it follows, that for every given impedance model the extremes of geoelectric field in YKC are larger than in MEA.



a)



b)

Figure 13. Impact of earth conductivity and geomagnetic field on the extreme values of the geoelectric indices a) same geomagnetic, but different surface impedances, b) same 4 surface impedance models (7-10), but different geomagnetic data (MEA and YKC)

The values of the extreme geoelectric field modeled with use of the relevant to location geomagnetic observatory data and corresponding surface impedance with the confidence intervals are presented in Figure 14. These values are listed in Table 3.

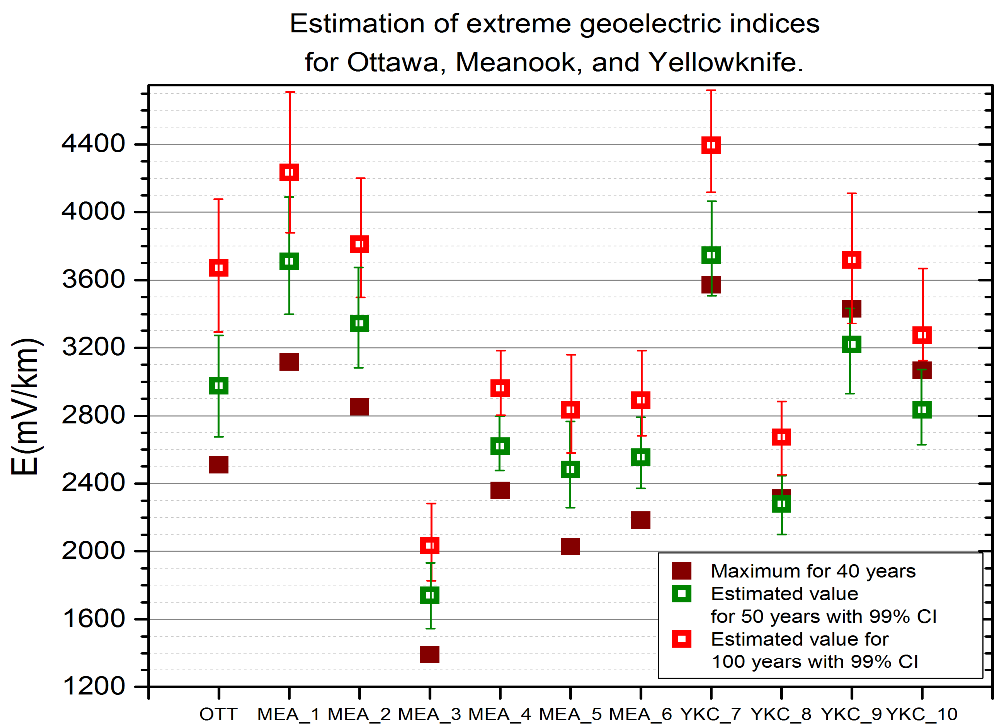


Figure 14. Extreme values of geoelectric indices

Table 3. Estimation for extreme values of geoelectric variations, once per 50 years, once per 100 years

Geomagneitc observatory/Earth resistivity model	Geoelectric field. 1 in 50 years forecast with 99% confidence interval, mV/km	Geoelectric field. 1 in 100 years forecast with 99% confidence interval, mV/km
MEA Earth Model 1	3709	4233
	[3386, 4063]	[3849, 4655]
MEA Earth Model 2	3343	3810
	[3072, 3639]	[3485, 4164]
MEA Earth Model 3	1739	2033
	[1553, 1947]	[1810, 2283]
MEA Earth Model 4	2618	2961
	[2466, 2779]	[2780, 3153]
MEA Earth Model 5	2480	2835
	[2247, 2736]	[2559, 3140]
MEA Earth Model 6	2554	2891
	[2355, 2770]	[2655, 3148]
MEA Earth Model 7	3231	3656
	[2997, 3483]	[3379, 3957]
MEA Earth Model 8	2031	2328
	[1880, 2196]	[2148, 2525]
MEA Earth Model 9	3101	3536
	[2895, 3322]	[3292, 3799]
MEA Earth Model 10	2664	3018
	[2519, 2816]	[2847, 3200]
YKC Earth Model 7	3744	4394
	[3488, 4021]	[4090, 4720]
YKC Earth Model 8	2278	2672
	[2105, 2467]	[2467, 2895]
YKC Earth Model 9	3218	3716
	[2929, 3536]	[3366, 4104]
YKC Earth Model 10	2835	3273
	[2538, 3167]	[2911, 3681]
OTT	2975	3670
	[2690, 3295]	[3300, 4090]

7 Conclusions

40 years of geomagnetic data recordings from three geomagnetic observatories were used to derive the extreme values of the geomagnetic hourly range index. While the results for the Meanook and Yellowknife locations are in close proximity to Alberta, the data for Ottawa were analyzed for setting some "background" geomagnetic activity values for comparison with some other, close to mid-latitudes, locations.

The statistical analyses including the extreme value statistics has been applied to estimate maxima of the geomagnetic variations

- once per 50 years;
- once per 100 years.

It has been estimated that for Meanook the extreme value for once in 100 year case with the 99% confidence interval is (4350±250) nT and for Yellowknife is (3100±300) nT. For comparison, in Ottawa the estimations give the extreme value of (2900±500) nT.

The geoelectric field has been modelled using the geomagnetic field data and earth resistivity models to derive the surface impedances at different locations. Then the monthly maximum values of the hourly peak amplitude of the geoelectric field have been used for extreme values statistics to estimate maxima of the horizontal geoelectric field

- once per 50 years;
- once per 100 years.

The extreme values of the geoelectric field in 100 years could reach 4250±400 mV/km at Meanook and 4400±350 mV/km at Yellowknife. For comparison, in Ottawa this value is 3670±420 mV/km.

It should be noted, that the geoelectric field results are very dependent on the surface impedance models, so that the difference in these models can give extreme values up to three times larger (smaller).

The method of extreme value statistics was tested by using 16 years of magnetic data to predict 40 years maximum in magnetic hourly range. Comparison with the actual maximum for the complete 40-year database showed an error of only 4%.

Acknowledgements

We acknowledge Dr. David Boteler for useful comments, Alberta power company AltaLink for support of this study and the staff of Geomagnetic Laboratory, NRCan, for provision of the geomagnetic data.

References:

Trichtchenko, L., Nikitina, L., Fernberg, P., Boteler, D., 2015. Space weather hazard assessment for Alberta. Geological Survey of Canada, Open file 7906, xxx p., doi: xxxx

Coles, S. 2004. An introduction to statistical Modelling of Extreme values, Springer, London.

Thomson, A., E. Dawson, and S. Reay, 2011. Quantifying extreme behaviour in geomagnetic activity, *Space Weather*, **9**, S10001, doi:10.1029/2011SW000696, 2011.