Estimating the Occurrence of Geomagnetic Activity Using the Hilbert-Huang Transform and Extreme Value Theory

Elvidge, Sean

DOI:
10.1029/2020SW002513

ADVANCING EARTH AND SPACE SCIENCE

Space Weather

RESEARCH ARTICLE
10.1029/2020SW002513

Key Points:
- A Hilbert-Huang transform is applied to the geomagnetic aa data to identify the solar cycle dependency in the data
- Extreme value theory is applied to the aa data separately in solar minimum and maximum conditions
- March 1989 (Quebec) event is shown to overall be a 1-in-25-year event, but a 1-in-130-year event during solar minimum

Correspondence to:
S. Elvidge,
s.elvidge@bham.ac.uk

Citation:
Elvidge, S. (2020). Estimating the occurrence of geomagnetic activity using the Hilbert-Huang transform and extreme value theory. *Space Weather*, *17*, e2020SW002513. https://doi.org/10.1029/2020SW002513

Received 10 APR 2019
Accepted 22 APR 2020
Accepted article online 28 APR 2020

Estimating the Occurrence of Geomagnetic Activity Using the Hilbert-Huang Transform and Extreme Value Theory

Sean Elvidge[1]

[1]Space Environment and Radio Engineering Group (SERENE), University of Birmingham, Birmingham, UK

Abstract In this paper extreme value theory (EVT) has been used to estimate the return levels for geomagnetic activity based on the aa index. The aa index is the longest, continuously recorded, geomagnetic data set (from 1868 to present). This long, 150-year data set is an ideal candidate for extreme value analysis. However, the data are not independent and identically distributed as required for EVT since they are impacted by the approximately 11-year solar cycle. The Hilbert-Huang transform has been used to identify the solar cycle component in the data, and the data have been split into solar maximum and minimum times. In these two regimes the generalized extreme value distribution has been fit to the data sets. These have also been combined for an estimate of the overall return times. The results suggest that the largest event in the database (March 1989) is a 1-in-25-year event. However, considering separate solar maximum and minimum times has a large impact on the return times. During solar minimum conditions the return time of the March 1989 event is 130 years. This suggests that the occurrence of extreme space weather events is conditionally dependent on where in the solar cycle we are.

1. Introduction

Geomagnetic storms are disturbances in the Earth's magnetosphere. They are caused by changes in the solar wind, which impact the magnetosphere. Two of the main causes for these changes in the solar wind are coronal mass ejections and high-speed solar wind streams (Schwenn, 2007). Coronal mass ejections usually have large speeds (approximately five times faster than the background solar wind), high energies, and large magnetic field strengths (Riley & Love, 2017). High-speed solar wind streams come from solar coronal holes, and the fast wind from these regions interacts with the slower upstream wind, which create corotating interaction regions (Garton et al., 2018). These regions have increased magnetic field strength and higher particle density (Schwenn, 2007). A lot of the largest space weather impacts are associated with geomagnetic storms: geomagnetically induced currents, radio scintillation, solar energetic particle events, and enhanced fluxes of relativistic electrons (Cannon, 2013).

Indices are used to quantify the relative strength of geomagnetic events. These include Dst (World Data Center for Geomagnetism Kyoto et al., 2015b), Kp and Ap (Bartels, 1957), AE, AO, AL, and AU (World Data Center for Geomagnetism Kyoto et al., 2015a), am, as, and an (Mayaud, 1980), and aa and Aa (Mayaud, 1972). The indices are calculated at different cadences from hourly to daily. Of these, the Dst index is used for identifying and quantifying the severity of geomagnetic storms (e.g., Loewe & Prölss, 1997). However, the various indices are, for the most part, closely related to each other.

Due to the impact of space weather events on human health and technology there is interest in estimating the return time for the most extreme events. A common reference point for extreme space weather events is the so called "Carrington event" (Carrington, 1859), which was one of the largest space weather events in the last 200 years (Cliver & Svalgaard, 2004). A key question is what is the return time (likelihood) of extreme space weather events. This is a difficult question to answer satisfactorily as it requires investigating the tails of probability distributions, where there are little data. However, this can be done rigorously using extreme value theory (EVT)/statistics.

EVT is mathematical rigorous and provides sensible measures of uncertainty, which can be very large when there are few data points. Therefore, one of the key difficulties associated with using EVT is the need for sets of large independent samples. For example, when looking at extreme temperatures (hot or cold) in meteorology the annual maxima or minima are suitable time scales since, on the whole, yearly temperatures are

Table 1
Weighting Factors for the Stations Used to Compile the aa Index (International Service of Geomagnetic Indices, 2013)

Northern Hemisphere station (UK)			Southern Hemisphere station (Australia)		
Time range	Station	Weighting factor	Time range	Station	Weighting factor
1868–1925	Greenwich	1.007	1868–1919	Melbourne	0.967
1926–1956	Abinger	0.934	1920–1979	Toolangi	1.033
1957 to present	Hartland	1.059	1980 to present	Canberra	1.084

independent, while daily temperatures are not. In the space weather domain it would be ideal to take solar cycle (11 years) minima or maxima time series. Unfortunately, there are not enough recorded data to have enough data points remaining for effective analysis.

A number of authors have applied EVT to the space weather domain, including Elvidge and Angling (2018), Koons (2001), Meredith et al. (2015), Silbergleit (1996, 1999), Siscoe (1976), Thomson et al. (2011), and Tsubouchi and Omura (2007). Extreme space weather events have also been estimated using techniques other than EVT such as fitting power law distributions, log normal distributions, and generalized Pareto distribution (Chapman et al., 2018, 2020; Riley, 2012; Riley & Love, 2017). The overall goal of each of the papers is to try to quantify the statistics of a particular measure associated with an extreme space weather event.

In terms of investigating geomagnetic activity Silbergleit (1996) and Tsubouchi and Omura (2007) used EVT to investigate extreme events in the Dst index using 23 and 44 years of data, respectively, Koons (2001) used the Ap index using 66 years of data, Siscoe (1976) and Silbergleit (1999) used variants of the aa index, using 91 and 124 years of data, and Thomson et al. (2011) used the rate of change of the magnetic field using 31 years. Riley and Love (2017) also estimated the probability of extreme Dst events using 60 years of data. Of those, the 91- and 124-year data sets of Siscoe (1976) and Silbergleit (1999) are very useful for EVT since they capture the longest time period (eight to 11 solar cycles). However, as well as the length of the data sets, the number of data used is also crucial in reducing the uncertainty in the analysis. Siscoe (1976) performed EVT only using the three largest events in each solar cycle during the test period, resulting in 27 data points, and Silbergleit (1999) used the maximum value from each solar cycle, resulting in 12 data points. While these approaches break the data into suitable scale sizes, few data points remain, which means that there are substantial uncertainties in the results.

In this paper the aa index (1868–2018; 150 years) is analyzed using the annual maximum values. This results in the largest temporal span of data for EVT in the space weather domain. Recent work by Chapman et al. (2020) has used a linear "mapping" between the top few percent of the aa index and the annual minimum Dst index value to estimate the probability of extreme Dst values. This work takes advantage of the extra data points from this aa-to-Dst mapped set.

While geomagnetic storms tend to last between 2 and 7 days, the annual maxima have been chosen to increase the likelihood of having independent, identically distributed data, a requirement for EVT. Using, for example, weekly rather than annual maxima can introduce further dependencies in the data set as individual events may, or may not, originate from the same solar active region. This could result in the analyzed data not being identically distributed. However, the disadvantage of using annual maxima is that the solar variability over the ~11-year cycle remains embedded in the data. This temporal dependence is accounted for by using the Hilbert-Huang transform (HHT) (Huang & Wu, 2008) to split the data into solar maximum and minimum times.

2. Extreme Value Theory

EVT provides a sophisticated approach for estimating probability distribution functions and specifically for looking at the tail of such distributions. The method avoids any starting assumption about the underlying distribution (Coles, 2001). The key result from EVT is the Fisher-Tippett-Gnedenko theorem, which states that the maximum of an independent and identically distributed (iid) random variable converges to one of only three possible distributions: the Gumbel distribution (Gumbel, 1935), the Fréchet distribution

10.1029/2020SW002513

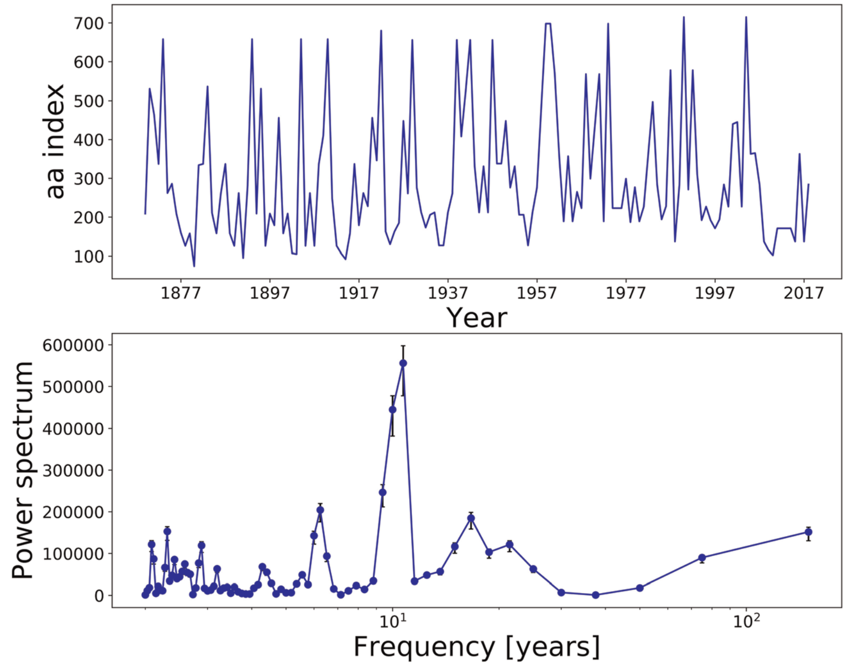

Figure 1. Top panel shows the time series of the annual maxima aa index. Bottom panel shows the periodogram of the series.

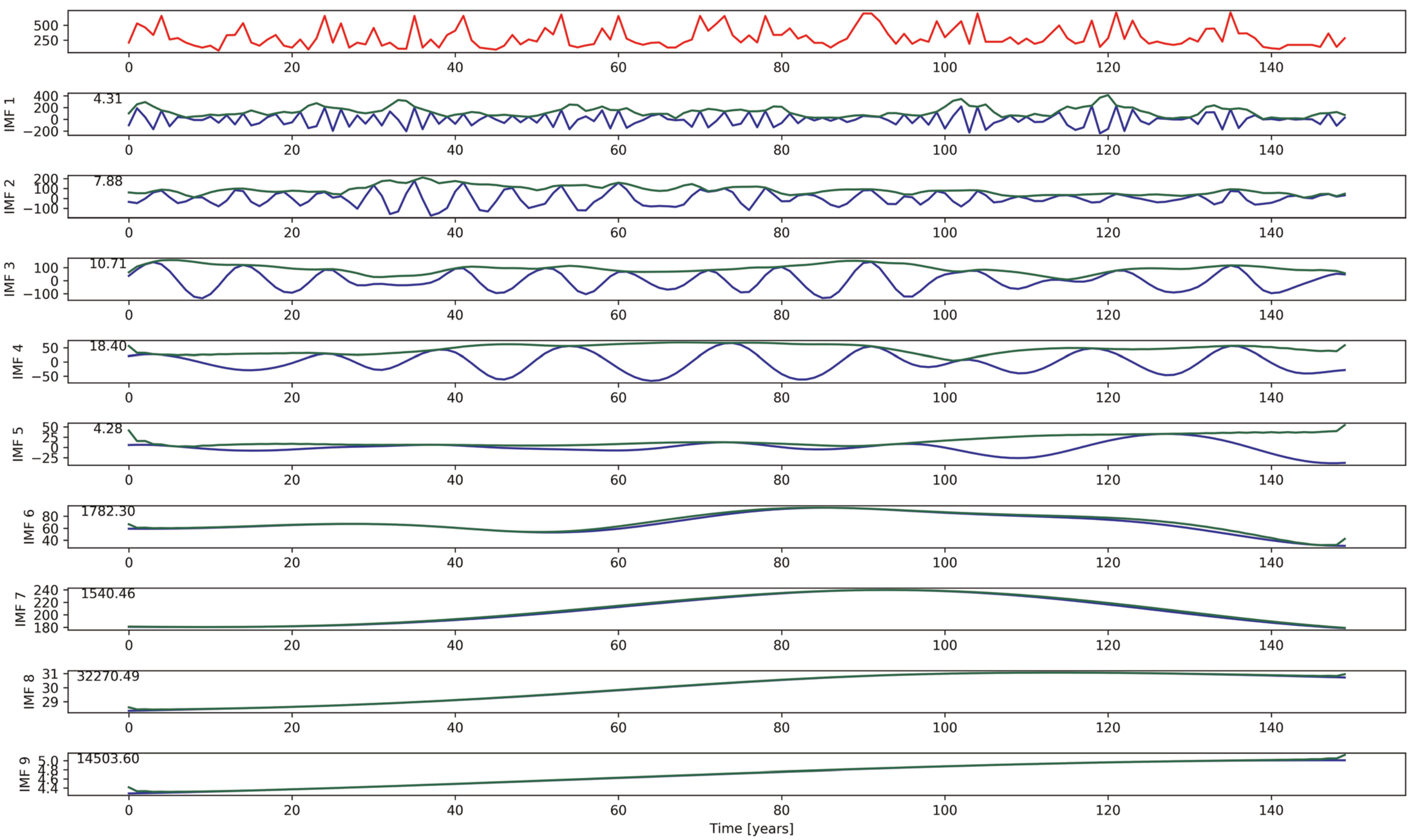

Figure 2. aa index decomposed into its nine IMFs. The original aa index is shown in the top panel (red), and each panel below shows an IMF (blue). For each IMF the envelope has been found using the Hilbert transform (green), and the dominant instantaneous frequency value is shown in the upper left of each IMF plot.

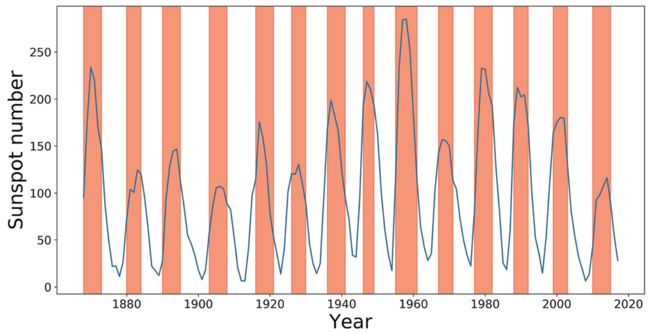

Figure 3. Annual maximum sunspot number (SILSO World Data Center, 2018). Solar maximum times (from the third IMF from the aa index) are shaded.

(Fréchet, 1927), or the Weibull distribution (Weibull, 1951), which can be grouped into the generalized extreme value (GEV) distribution.

Specifically, for a sequence of iid random variables $X_1, X_2, ..., X_n$ with common distribution function F let $M_n = \max\{X_1, \ldots, X_n\}$ and $w = \sup\{x : F(x) < 1\}$. Then,

$$\Pr(M_n \leq x) = \Pr(X_1 \leq x, ..., X_n \leq x) = F^n(x). \tag{1}$$

Then, as $n \to \infty$, $F^n(x) \to 0$ if $x < w$ and $F^n(x) \to 1$, otherwise $M_n \to w$. To avoid a degenerate distribution $F^n(x)$ is normalized. Assuming there is a nondegenerate distribution G such that, for normalizing constants $a_n > 0$ and b_n:

$$\lim_{n\to\infty} F^n(a_n x + b_n) = G(x), \tag{2}$$

where G is the GEV distribution defined by

$$G(x) = \exp\left\{-1\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}\right\}, \tag{3}$$

defined for $1 + \dfrac{\xi(x-\mu)}{\sigma} > 0$ and where μ is the location parameter, $\sigma > 0$ is the scale parameter, and ξ is the shape parameter (Coles, 2001). For $\xi < 0$ the GEV reduces to the Weibull distribution. For $\xi > 0$ the Fréchet distribution, and in the limit $\xi \to 0$, $G(x)$ reduces to

$$G(x) = \exp\left\{-\exp\left(-\frac{x-\mu}{\sigma}\right)\right\}, \tag{4}$$

the Gumbel distribution.

The parameters of the GEV are usually estimated using a maximum log likelihood method (Coles, 2001). However, the requirement that the variables must be iid is usually a barrier with using raw data directly, and some form of preprocessing is normally required.

3. Data

The aa index is a global geomagnetic index, which is based on the largest horizontal deviation of the magnetic field measured in nT. It is based on data from two nearly antipodal stations, one in the United Kingdom and another in Australia, and has been continuously recorded since 1868. Over the 150 years the stations where the data have been recorded have changed. In order to maintain a constant value for the index the weighting of the different stations has varied over time (Table 1).

The main advantage of using the aa index for EVT is that it is the longest running planetary index of geomagnetic activity. This long sample time helps in reducing the uncertainties in the EVT extrapolation. However, on this time scale the impact of the solar cycles becomes apparent, which has been shown to have an impact on the results of extreme value modeling (Riley & Love, 2017). Figure 1 shows the time series of the aa index in the top panel (each point is the annual maximum aa value), and the bottom panel of the figure shows the periodogram created with a Hamming windowing function on the data. The large peak in the periodogram corresponds to 10.7 years and is the solar cycle contribution to the data. To perform EVT on the data set this temporal dependence should be accounted for. In this work the HHT is used to identify solar maximum and minimum times, which are assumed to each be iid, and EVT can be performed on each.

The HHT decomposes a time series into intrinsic mode functions (IMFs) and then finds the instantaneous frequency of each IMF (Huang & Wu, 2008). The first step of the HHT is to use empirical

Table 2
The GEV Fit Parameters for the Solar Maximum and Minimum Conditions

	Number data points	μ	σ	ξ
Maximum	78	279 (17.8)	130 (12.6)	−0.03 (0.12)
Minimum	72	174 (9.24)	68.8 (7.16)	0.15 (0.10)

Note. The standard error is shown in parentheses.

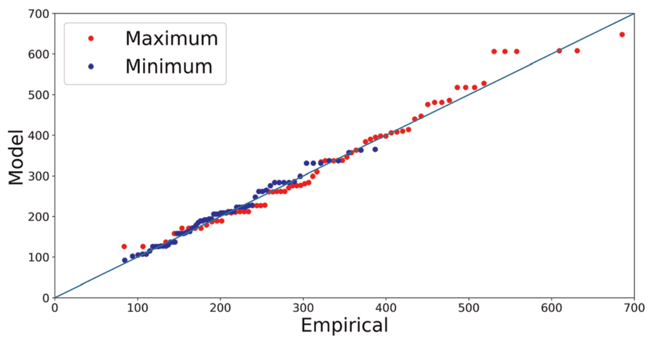

Figure 4. Quantile plot of the GEV fit for solar maximum and minimum times.

mode decomposition (Huang et al., 1998). Similar to the Fourier and wavelet transforms, empirical mode decomposition splits a signal into its components, called IMFs. An IMF is a function in which the number of extrema and zero crossings differs by at most 1, and at each point the mean value of the envelopes, defined by the local maxima and minima, is 0. The sum of the IMFs reconstitutes the original signal. Hilbert spectral analysis can then be used by applying the Hilbert transform to each IMF to find the instantaneous frequency (Huang et al., 1998). Unlike Fourier and wavelet transforms, HHT is an algorithmic approach rather than theoretical. However, the main advantage of the HHT over Fourier and wavelet is that it is suitable for nonlinear and nonstationary data.

Applying the HHT to the aa index data (top panel of Figure 1) results in nine IMFs shown in Figure 2. The top panel of the figure shows the original aa index values (in red). Then, each of the IMFs is shown in blue. Using Hilbert spectral analysis, the envelope of each IMF has been found (shown in green), and the value of the instantaneous time period (ITP) (one over the instantaneous frequency) is shown in the upper left of each IMF plot. A single value for the ITP is found by fitting a linear polynomial through the data. In each case the polynomial was of Order 0 (as expected), and the constant term is shown. From the figure it can be seen that the third IMF has an ITP, which corresponds to the peak time period from Figure 1. This provides confidence that this particular IMF corresponds to the sunspot cycle component in the aa index data.

This particular IMF could be used in the fitting of the GEV as part of a temporally varying location and scale parameters (Coles, 2001). However, this requires propagation of the IMF forward in time. With only simplistic ways of interpolating the IMF forward in time (it is hard to predict the next solar cycle), it makes finding the return times difficult (a key attribute for extreme value modeling).

Instead, the aa variables can be made to be iid by separating the solar maximum and minimum conditions. In each of the two cases it can be assumed that they are distributed according to the same function. Rather than using a separate data set to try and determine when these times are, the third (solar cycle dependent) IMF can be used. The IMF is centered at 0, and positive values can be used to describe solar maximum times, while negative values can be used as solar minimum times. Comparing the IMF estimated solar maximum/minimum times to the sunspot cycle (as independent verification of this approach) show that the estimated maximum overlaps with the peak of the sunspot number as would be expected (Figure 3).

4. Results

Fitting the GEV to the solar maximum and minimum time series using least log likelihood results in the estimated μ, σ, ξ (and standard errors) as shown in Table 2. One method for verifying the quality of the fit of the GEV distribution is by looking at the quantile plot, which shows the pairs

$$\left\{G^{-1}\left(\frac{i}{n+1}\right), x_i\right\}, \qquad i \in \{1, \ldots, n\}, \tag{5}$$

for the ordered annual aa values $\{x_1, x_2, \ldots, x_n\}$ and where $G^{-1}(x)$ is the inverse of equation 3, given by Coles (2001):

$$G^{-1}\left(\frac{i}{n+1}\right) = \mu - \frac{\sigma}{\xi}\left(1 - \left(-\log\left(\frac{i}{n+1}\right)\right)^{-\xi}\right). \tag{6}$$

The quantile plot for solar maximum and minimum times is shown in Figure 4. The plot being roughly linear is an indication of a good agreement between the model fit and the empirical data.

Using the fitted GEV distributions, the return time for any given event can be estimated for either solar maximum or minimum

Table 3
Return Levels for 10-, 50-, 100-, 500-, and 1,000-Year Return Periods for Solar Maximum and Minimum Conditions as Well as Combined Results

	10 years	50 years	100 years	500 years	1,000 years
Maximum	611 (50)	873 (146)	989 (208)	1,271 (408)	1,399 (519)
Minimum	365 (30)	565 (89)	670 (131)	969 (287)	1,127 (386)
Combined	631 (60)	898 (155)	1,019 (216)	1,321 (429)	1,463 (562)

Note. The standard error is shown in parentheses.

Table 4
Return Levels for This Work as Well as Previous EVT Work on Ap (Koons, 2001)

	Normalized 10 years	Normalized 50 years	Normalized 100 years	Normalized 500 years	Normalized 1,000 years
Aa (combined)	0.88 (0.08)	1.26 (0.21)	1.43 (0.30)	1.85 (0.60)	2.05 (0.79)
Ap (Koons, 2001)	0.80 (0.08)	1.09 (0.17)	1.21 (0.22)	1.43 (0.34)	1.54 (0.42)

Note. The values are normalized by dividing through by the index value of the event in March 1989.

conditions using the values in Table 2. However, another useful return time would combine both solar maximum and minimum. The combined return time can be found by solving

$$G_{\max}(z_p)G_{\min}(z_p) = 1 - \frac{1}{p}, \tag{7}$$

for z_p, the return level with return period $1/p$ (z_p is expected to be exceeded by the annual maximum aa value with probability p), and where $G_{\max}$, $G_{\min}$ are the GEV distributions defined by the parameters in Table 2 (Coles, 2001). The return levels for 10, 50, 100, 500, and 1,000 years are shown in Table 3, with the standard errors shown in parentheses. It is interesting to note the difference in return times between solar maximum and minimum times. This suggests that the probability of an extreme event is conditionally based on what part of the solar cycle we are in. This is in agreement with the findings of Riley and Love (2017) who determined that, assuming a power law distribution, the probability of geomagnetic storm exceeding the Carrington event (in terms of Dst) was 1.4% during solar minimum conditions and 28% for solar maximum conditions.

These results have been compared to previous EVT work on the Ap index (an index similar to the aa index) undertaken by Koons (2001). To compare the results, since the data are on different scales, they have been normalized by dividing through by the peak value of the March 1989 event (Feynman & Hundhausen, 1994). This is the largest value in the aa index database (tied with the "Halloween" event of 2003) and the second largest in Ap (the largest is an event is from November 1960). Koons (2001) provides the fit parameters for the Gumbel distribution, which was shown to have the best fit with the data. However, no standard error was reported in the results. So for ease of comparison in this work the same 66 years of data was analyzed and fit with the same Gumbel distribution as in Koons (2001) ($\mu = 99.1409$ and $\sigma = 42.9416$). The normalized results for both the aa and Ap index data are shown in Table 4.

Comparing the EVT results between the aa and Ap indices shows that they are very similar for "short" return times (10 years) with values of 0.88 and 0.80, respectively. These differences widen as the return periods get longer. This is to be expected as in this work 150 years of aa data have been used for the estimates compared to 66 years of Ap data from Koons (2001). The extra data points should provide better estimates for the longer return periods, especially the 100-year return level.

5. Conclusions

EVT has been used to estimate the return levels for geomagnetic activity based on the aa index. The aa index is the longest, continuously recorded, geomagnetic data set. This long, 150-year data set is an ideal candidate for extreme value analysis. While the aa index is not the most commonly used space weather index its close relationship with the more commonly used Dst index (Chapman et al., 2020) implies that similar geoeffective impacts of extreme Dst events would be felt during extreme aa events. However, the aa data are not independent and identically distributed (iid) as required for EVT as they are impacted by the approximately 11-year solar cycle. The HHT has been used to identify the solar cycle component in the data, and the data have been split into solar maximum and minimum times. In these two regimes the variables are assumed to be iid, and the GEV distribution has been fit to the two data sets. These have also been combined for an estimate of the return times.

The results suggest that the largest event in the database (March 1989/October 2003) is a 1-in-25-year event (but with a standard error of 86 years). While this may seem counterintuitive, since there are only two events of that size (aa of 715) in the 150-year database, there are in total eight events where the aa index exceeds 650.

Considering separate solar maximum and minimum times has a large impact on the return time. During solar minimum the return time of the March 1989 event is 130 years (with a standard error of 145 years). This value seems reasonable since there is one event with an aa >650 during solar minimum in the 150-year database (4 August 1972; Knipp et al., 2018). However, it is in contrast to the results of Riley and Love (2017) who report the probability as ~1 in 700 years during solar minimum (assuming a power law distribution). This demonstrates the uncertainty that can arise when extrapolating extreme events by using different underlying assumptions. Quantifying the impact of solar minimum is of particular importance in quantifying the likelihood associated with extreme space weather events if a period of extended solar minimum is entered. It has been estimated that there is a 15–20% chance of returning to Maunder minimum-like conditions within the next 40 years (Ineson et al., 2015; Lockwood, 2010).

Acknowledgments

The aa data can be downloaded from British Geological Survey (http://www.geomag.bgs.ac.uk/data_service/data/magnetic_indices/aaindex.html).

References

Bartels, J. (1957). The geomagnetic measures for the time-variations of solar corpuscular radiation, described for use in correlation studies in other geophysical fields. *Annals of the International Geophysical Year*, *4*, 227–236.

Cannon, P. S. (2013). Extreme space weather—A report published by the UK Royal Academy of Engineering. *Space Weather*, *11*, 138–139. https://doi.org/10.1002/swe.20032

Carrington, R. C. (1859). Description of a singular appearance seen in the Sun on September 1, 1859. *Monthly Notices of the Royal Astronomical Society*, *20*, 13–15.

Chapman, S. C., Horne, R. B., & Watkins, N. W. (2020). Using the aa index over the last 14 solar cycles to characterize extreme geomagnetic activity. *Geophysical Research Letters*, *47*, e2019GL086524. https://doi.org/10.1029/2019GL086524

Chapman, S. C., Watkins, N. W., & Tindale, E. (2018). Reproducible aspects of the climate of space weather over the last five solar cycles. *Space Weather*, *16*, 1128–1142. https://doi.org/10.1029/2018SW001884

Cliver, E. W., & Svalgaard, L. (2004). The 1859 solar-terrestrial disturbance and the current limits of extreme space weather activity. *Solar Physics*, *224*, 407–422.

Coles, S. (2001). *An introduction to statistical modeling of extreme values*. London: Springer-Verlag. https//doi.org/10.1007/978-1-4471-3675-0

Elvidge, S., & Angling, M. J. (2018). Using extreme value theory for determining the probability of Carrington-like solar flares. *Space Weather*, *16*, 417–421. https://doi.org/10.1002/2017SW001727

Feynman, J., & Hundhausen, A. J. (1994). Coronal mass ejections and major solar flares: The great active center of March 1989. *Journal of Geophysical Research*, *99*, 8451–8464 https://doi.org/10.1029/94JA00202

Fréchet, M. (1927). Sur la loi de probabilité de l'écart maximum. *Annales de la Societe Polonaise de Mathematique*, *6*.

Garton, T. M., Murray, S. A., & Gallagher, P. T. (2018). Expansion of high speed solar wind streams from coronal holes through the inner heliosphere. *The Astrophysical Journal Letters*, *869*, L12. https://doi.org/10.3847/2041-8213/aaf39a

Gumbel, E. J. (1935). Les valeurs extrêmes des distributions statistiques. *Annales de l'Institut Henri Poincaré*, *5*, 115–158.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., et al. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *454*, 903–995. https://doi.org/10.1098/rspa.1998.0193

Huang, N. E., & Wu, Z. (2008). A review on Hilbert-Huang transform: Method and its applications to geophysical studies. *Reviews of Geophysics*, *46*, RG2006. https://doi.org/10.1029/2007RG000228

Ineson, S., Maycock, A. C., Gray, L. J., Scaife, A. A., Dunstone, N. J., Harder, J. W., et al. (2015). Regional climate impacts of a possible future grand solar minimum. *Nature Communications*, *6*, 7535.

International Service of Geomagnetic Indices (2013), aa Index.

Knipp, D. J., Fraser, B. J., Shea, M. A., & Smart, D. F. (2018). On the little-known consequences of the 4 August 1972 ultra-fast coronal mass ejecta: Facts, commentary, and call to action. *Space Weather*, *16*, 1635–1643. https://doi.org/10.1029/2018SW002024

Koons, H. C. (2001). Statistical analysis of extreme values in space science. *Journal of Geophysical Research*, *106*, 10,915–10,921.

Lockwood, M. (2010). Solar change and climate: An update in the light of the current exceptional solar minimum. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *466*, 303–329. https://doi.org/10.1098/rspa.2009.0519

Loewe, C. A., & Prölss, G. W. (1997). Classification and mean behavior of magnetic storms. *Journal of Geophysical Research*, *102*(A7), 14,209–14,213. https://doi.org/10.1029/96JA04020

Mayaud, P. N. (1972). The aa indices: A 100-year series characterizing the magnetic activity. *Journal of Geophysical Research*, *77*, 6870–6874.

Mayaud, P. N. (1980). Derivation, meaning, and use of geomagnetic indices. *Geophysical Monograph Series*, *22*.

Meredith, N. P., Horne, R. B., Isles, J. D., & Rodriguez, J. V. (2015). Extreme relativistic electron fluxes at geosynchronous orbit: Analysis of GOES E >2 MeV electrons. *Space Weather*, *13*, 170–184. https://doi.org/10.1002/2014SW001143

Riley, P. (2012). On the probability of occurrence of extreme space weather events. *Space Weather*, *10*, S02012. https://doi.org/10.1029/2011SW000734

Riley, P., & Love, J. J. (2017). Extreme geomagnetic storms: Probabilistic forecasts and their uncertainties. *Space Weather*, *15*, 53–64. https://doi.org/10.1002/2016SW001470

Schwenn, R. (2007). Solar wind sources and their variations over the solar cycle. In D. N. Baker, B. Klecker, S. J. Schwartz, R. Schwenn, & R. von Steiger (Eds.), *Solar dynamics and its effects on the heliosphere and Earth* (pp. 51–76). New York, NY: Springer New York.

Silbergleit, V. M. (1996). On the occurrence of geomagnetic storms with sudden commencements. *Journal of Geomagnetism and Geoelectricity*, *48*, 1011–1016.

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

SILSO World Data Center (2018), The international sunspot number (1868–2018), International Sunspot Number Monthly Bulletin and online catalogue.

Siscoe, G. L. (1976). On the statistics of the largest geomagnetic storms per solar cycle. *Journal of Geophysical Research*, *81*.

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

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

Weibull, W. (1951). A statistical distribution function of wide applicability. *Journal of Applied Mechanics—ASME*, *3*, 293–297.

World Data Center for Geomagnetism Kyoto, Nose, M., Iyemori T., Sugiura M., & Kamei T. (2015a). Geomagnetic AE index, https://doi.org/10.17593/15031-54800

World Data Center for Geomagnetism Kyoto, Nose, M., Iyemori, T., Sugiura M., & Kamei T. (2015b). Geomagnetic Dst index, https://doi.org/10.17593/14515-74000