Solar Physics (2025) 300:50
https://doi.org/10.1007/s11207-025-02464-6

RESEARCH

^{14}C-Based Sunspot Numbers for the Last Millennium Encompass the Full Range of Variability: Extreme Value Theory

F.J. Acero[1,2] · V.M.S. Carrasco[1,2] · M.C. Gallego[1,2] · I.G. Usoskin[3] · J.M. Vaquero[1,2]

Received: 13 August 2024 / Accepted: 27 March 2025

Abstract
We examine the statistical properties of extreme solar activity levels through the application of the extreme value theory to the annual sunspot number series reconstructed from ^{14}C data spanning the last millennium. We have used the extreme value theory to study long-term solar variability by applying the peaks-over-threshold technique to an annual sunspot number series reconstructed from ^{14}C data for the last millennium. We have obtained a negative value of the shape parameter of the generalized Pareto distribution implying that an upper bound has been reached by the extreme sunspot number value distribution during the past millennium. The results obtained from the same analysis applied to two subperiods of the series, are consistent with that considering the whole series. We have also estimated return levels and periods for the extreme sunspot numbers. The maximum annual sunspot number (273.6) observed during the past millennium is slightly higher (lower) than that considering a 1000-year (10,000-year) return level, but they are within the 95% confidence interval in both cases. It approximately corresponds to a 3500-year return period. Our result implies that solar activity has reached its upper limit, and it would be unlikely to observe, in the near future, sunspot numbers significantly higher than those already observed during the past millennium.

Keywords Sunspots · Statistics · Solar cycle · Models · Cosmic rays · Galactic

1. Introduction

Sunspots have been observed with the aid of a telescope by hundreds of astronomers since the beginning of the 17th century (Muñoz-Jaramillo and Vaquero 2019; Arlt and Vaquero

✉ V.M.S. Carrasco
vmscarrasco@unex.es

1 Departamento de Física, Universidad de Extremadura, 06006 Badajoz, Spain

2 Instituto Universitario de Investigación del Agua, Cambio Climático y Sostenibilidad (IACYS), Universidad de Extremadura, 06006 Badajoz, Spain

3 Space Physics and Astronomy Research Unit and Sodankylä Geophysical Observatory, University of Oulu, 90014 Oulu, Finland

Published online: 07 April 2025

2020). The databases that include these records, available since 1610, represent the longest datasets using direct solar observations (Wolf 1860; Hoyt and Schatten 1998; Vaquero et al. 2016). The sunspot number series based on these databases are the most used indices to characterize long-term solar activity (Clette and Lefèvre 2016; Svalgaard and Schatten 2016; Usoskin et al. 2016; Clette et al. 2023). As an example, the sunspot number series is involved in the planning of future space missions (Khazanov 2016). Therefore, the understanding and prediction of solar activity in general, and of its extreme events in particular, are fundamental to avoid or at least mitigate the problems that solar activity can cause to our society (Pulkkinen 2007; Usoskin et al. 2023).

There are different types of prediction methods that make long-term forecasts of solar activity for recent solar cycles (Pesnell 2012; Petrovay 2020; Nandy 2021). For instance, when considering diverse predictions for the maximum amplitude of Solar Cycle 25, various approaches are evident. These include physics-based models as, e.g., that those proposed by Bhowmik and Nandy (2018), precursor-based predictions as seen in McIntosh et al. (2020), statistical forecasts such as Aparicio, Carrasco, and Vaquero (2023), and predictions utilizing Machine Learning techniques, as proposed by Rodríguez and Rodríguez (2020).

There is another branch to study (not to predict) longer-term solar variability. This is the extreme value theory (EVT), which is widely used in other scientific fields such as climatology (Beguería and Vicente-Serrano 2006; Acero, García, and Gallego 2011; Acero, Gallego, and García 2012; Acero et al. 2014) and engineering (Castillo et al. 2004). In the case of the space climate, the EVT has been applied by Siscoe (1976) to the aa index measured during large geomagnetic storms. Asensio (2007) used the EVT and the international sunspot number (version 1 of SILSO) to investigate the statistical properties of the extreme events of solar activity. Furthermore, we have previously applied the EVT to the international sunspot number (version 2) at different temporal scales (Acero et al. 2017), the hourly values of the Dst geomagnetic index for the period 1957 – 2014 (Acero et al. 2018a), and two decadal sunspot number series reconstructed from ^{14}C measured in tree trunks and ^{10}Be in polar ice (Acero et al. 2018b). Also, Zhang et al. (2023) used the extreme value theory applied to long-term sunspot areas to predict the trend of the 25th and 26th solar cycles.

Miyake et al. (2012) identified a significant increase in the ^{14}C levels measured in tree rings of two Japanese cedar trees between 774 and 775 AD. While various explanations were proposed to determine the origin of this event, subsequent research confirmed that the Sun was responsible for this increase (Usoskin et al. 2013). Following the groundbreaking work by Miyake et al. (2012), similar events were discovered (see, e.g., reviews by Cliver et al. 2022; Usoskin et al. 2023), such as the one between 993 and 994 AD (Miyake, Masuda, and Nakamura 2013), and others are yet to be confirmed, including those in 1052 AD and 1279 AD (Brehm et al. 2021). Such events indicate extremely strong bursts of solar activity, which might also be reflected in extreme sunspot numbers during those periods.

The objective of this work is to analyze the statistical properties of the extreme values of solar activity by applying the EVT to the annual sunspot number series (Usoskin et al. 2021) reconstructed from ^{14}C data for the last millennium (Brehm et al. 2021). We also examine whether the levels of solar activity in the year 994 AD, during a Miyake-type event documented and in 1052 AD and 1279 AD, two cases considered as potential Miyake-type events, meet the criteria for extreme events through the application of EVT. The outline of this work is as follows. The data used in this work and an explanation of the methodology followed is shown in Section 2. The results obtained together with its discussion are presented in Section 3. Section 4 includes the main conclusions of this work.

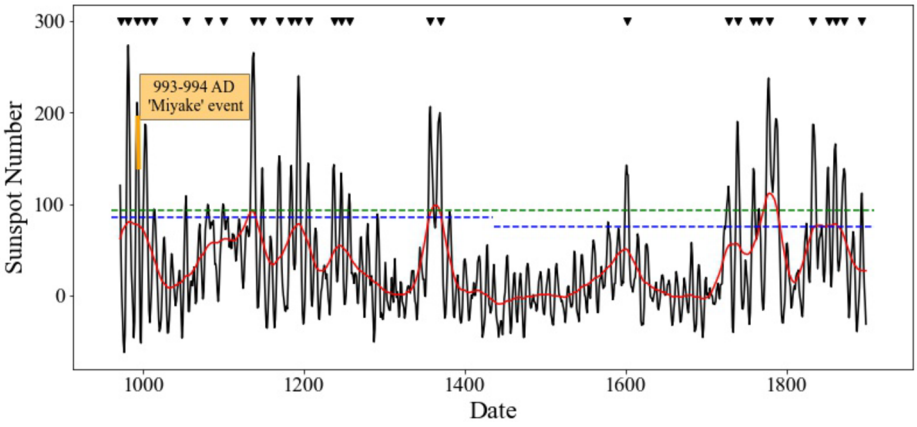

Figure 1 Sunspot number series by Usoskin et al. (2021) reconstructed from ^{14}C data. The black line represents the annual data, and the red line depicts the smoothed (22-year singular spectral analysis) sunspot number. The horizontal green and blue dashed lines show the threshold for the whole period and for the two subperiods chosen in this work, respectively. The solar activity level when a 'Miyake' event occurred in 993 – 994 AD is indicated in orange. Triangles on the top indicate independent clusters of activity defined in this work (see subsequent sections for more details).

2. Data and Methods

2.1. Data

Sunspot numbers reconstructed by Usoskin et al. (2021) from annual ^{14}C data are considered in this work. This series spans from 971 to 1899 AD. We note that this series includes various uncertainties associated to both the ^{14}C measurement errors and the reconstruction of the sunspot numbers from ^{14}C data (for more details, see Usoskin et al. 2021). The reconstructed sunspot numbers are available at the website: http://cdsarc.u-strasbg.fr/viz-bin/cat/J/A+A/649/A141. Figure 1 represents this as the black line along with a smoothed (22-year singular spectral analysis) sunspot number (red line). Furthermore, we have also taken into account the upper limit of this series. This upper limit was obtained by summing the annual sunspot numbers depicted in Figure 1 plus their corresponding 1-sigma uncertainties provided by Usoskin et al. (2021).

2.2. Methodology

The EVT can be applied using different approaches (Coles 2001). One of them is known as peaks-over-threshold (POT), which is chosen to be applied in this work. This technique considers values higher than a pre-selected threshold (u) that are called exceedances. The generalized Pareto distribution (GPD – see Equation 1) is used to model the probability distribution of the exceedances above the threshold. Thus, first, it is necessary to set a threshold.

We have employed two ways to find the best threshold. One is based on the mean residual life plot (Coles 2001), which is a technique that consists in plotting the sample mean of the exceedances against the value of u, searching for the value from which the graph is approximately linear (Figure 2). For sunspot number values slightly lower than 100, one can see flatness (where the GPD model becomes valid) in Figure 2. The second way is the assessment of the stability of the parameters' estimates (Figure 3). These are the scale and

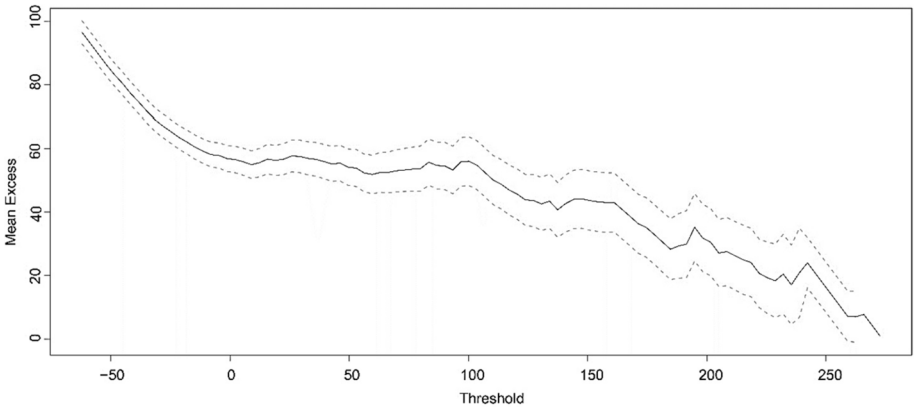

Figure 2 Mean residual life plot used in this work to choose the threshold of the studied sunspot number series. A 95 % confidence interval is represented by dashed lines.

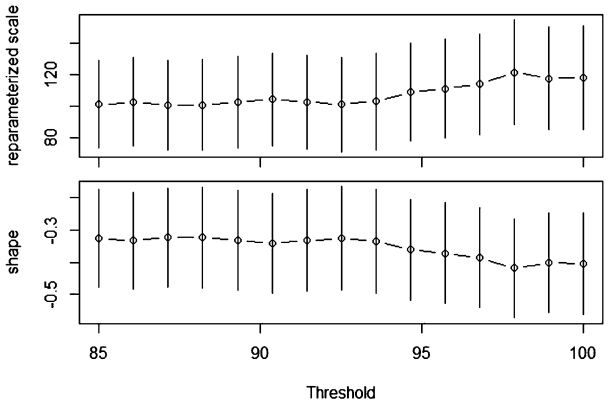

Figure 3 Parameter estimates against threshold using the studied annual sunspot number series. A 95 % confidence interval is represented by vertical lines.

shape parameters (see Coles 2001, for more details). The parameter estimates should be stable, i.e., approximately constant. Thus, the stability of these parameter estimates, fitting the GPD over a range of thresholds, can be seen roughly for the same sunspot number values as in the first case. Delimiting the threshold below 100 in Figure 2, the parameter estimates for thresholds between 85 and 100 are shown in Figure 3 in order to select an appropriate threshold. Both parameters (shape and scale) are nearly constant till $u = 93$, confirming this value as the best threshold. We note that the use of a threshold value between 90 and 95 does not alter our results.

Two subperiods halving the entire study period (971 to 1899 AD) are also considered to check the consistency of the result using the entire study period: 1) from 971 to 1435 AD and 2) from 1436 to 1899 AD. Applying the same methodology to get the best threshold in each subperiod, we obtained that the best thresholds are 85 for the first subperiod and 75 for the second one. Note that for a shorter period, the threshold is lower than that for the entire period. This ensures a sufficient number of exceedances in each subperiod to apply the EVT.

The POT approach requires the exceedances to be mutually independent to avoid short-range dependencies in the series. However, the sunspot number values studied in this work are grouped into clusters, since one can see that some consecutive years exceed the thresh-

old. Therefore, this requires a declustering procedure to identify independent exceedances of the threshold. For this purpose, we used the "runs declustering" (Leadbetter et al. 1989). This method assumes that exceedances belong to different clusters when they are separated by a certain number of data, called the run length (r), below u. In our case, we consider a run length $r = 1$ obtaining one data for each cluster exceeding the threshold. Then, we selected the year with the maximum sunspot number within each cluster, and a new time series with the date and the intensity was calculated.

This new time series was subjected to a GPD analysis. In the asymptotic limit for sufficiently large thresholds, the distribution of independent overruns $X = SN(t) - u$ with $SN(t) > u$, being SN the sunspot number, follows a GPD such as:

$$P(X < x) = 1 - (1 + \frac{\xi x}{\sigma})^{\frac{-1}{\xi}}, \tag{1}$$

with $x > 0$ and $1 + \frac{\xi x}{\sigma} > 0$, where σ is the scale parameter, and ξ the shape parameter ($\xi \neq 0$). Shape parameter values equal to or above zero mean that there is no upper limit in the distribution, whereas values below zero indicate that the distribution has an upper limit (Coles 2001). The scale and the shape parameters were estimated by maximum likelihood, which is a standard general-purpose statistical technique for fitting a given parametric distribution to a set of data. Generally, this technique is commonly used for the estimation of the parameters of the extreme value distribution in the extreme value theory (Coles 2001) due to its robustness. For this purpose, the in2extRemes statistical R software package for extreme values (Gilleland and Katz 2016) has been used. Once both parameters were estimated, their confidence interval were evaluated by a bootstrap procedure which is a statistical method that resamples a single data set to create many simulated subsamples widely used to estimate confidence intervals (see more details in Gilleland and Katz 2016).

The concept of the return level is used to evaluate the occurrence probability of future extreme events. It means the expected level to be exceeded once every certain number of years on average. We have estimated the return level for 1000 and 10,000 years. The same procedure as in the estimation of the parameters of the GPD was used to estimate the return levels and their confidence intervals with the bootstrap procedure. Here we assume that the sunspot number series is statistically uniform over the multi-millennial time scale, although it may exhibit longer-term trends, such as the about 2400-year Hallstatt cycle (e.g., Usoskin et al. 2016). However, this trend is small, within ±5 in sunspot number (see Figure 3c in Usoskin et al. 2016) and does not affect our main conclusions.

3. Results and Discussion

We have used the POT approach with a threshold of $u = 93$ to analyze the millennium-long sunspot number reconstructed from ^{14}C data. Thus, the number of overruns obtained considering the whole series was 127 with the highest sunspot number value of 273.6 in 981 AD. After applying the declustering process, we find 30 independent clusters (see down triangles in Figure 1). We note that some solar cycles may include several extreme sunspot numbers and other cycles do not have extreme values. The time series including those 30 clusters was fitted to a GPD (Equation 1). Table 1 shows the estimates for the scale and shape parameters (95% confidence intervals) obtained by bootstrapping. We obtained that the shape parameter is negative, which means that there is an upper bound of the extreme sunspot number distribution. This result is highly statistically significant, indicating that a

Table 1 Estimates of the GPD parameters and their 95 % confidence intervals (CI) obtained by bootstrapping.

Threshold u (Period)	Scale (σ) [95% CI]	Shape (ξ) [95% CI]
93 (971 – 1899 AD)	71.1 [57.4, 90.5]	−0.33 [−0.52, −0.20]
93 (971 – 2023 AD)	75.9 [64.4, 91.5]	−0.36 [−0.50, −0.27]
85 (971 – 1435 AD)	74.2 [56.7, 101.6]	−0.28 [−0.56, −0.13]
75 (1436 – 1899 AD)	77.0 [58.2, 110.7]	−0.43 [−0.75, −0.25]

positive value of the shape parameter is highly unlikely. We note that, for negative values of the shape parameter, the EVT establishes that the upper bound can be estimated as $u - \sigma/\xi$ (Coles 2001). In this case, the upper bound obtained for the sunspot number by Usoskin et al. (2021) is 311.51 (with values between 131.21 and 449.54 considering a 95% confidence interval)

A similar analysis was also made by appending the sunspot number values (version 2) to the ^{14}C-based sunspot numbers after 1900. For the best threshold found ($u = 93$), we obtained similar results as previously: i) shape parameter $= -0.36$ with a 95% confidence interval of $-0.50 - -0.27$; ii) scale parameter $= 75.9$ with a 95% confidence interval of 64.4 – 91.5. In addition, we have applied the same methodology to the upper limit of the ^{14}C-based sunspot numbers obtaining similar results, that is, a negative value of the shape parameter: i) shape parameter $= -0.15$ with a 95% confidence interval of $-0.30 - -0.05$; ii) scale parameter $= 75.1$ with a 95% confidence interval of 63.8 – 90.7. Note that in this last case, the best threshold corresponds to a sunspot number value of 115.

It is necessary to check the accuracy of the GPD fit to the threshold exceedances for the studied sunspot number series. For this purpose, different diagnostic plots were used. The diagnostic plots for the GPD fit the maximum sunspot number values are shown in Figure 4. In the top panel of Figure 4, a quantile–quantile (QQ) plot compares empirical data quantiles with GPD fit quantiles, showing similar distributions with points aligning along the line y = x (solid line). The middle panel of Figure 4 displays a QQ-plot contrasting randomly generated data from the fitted GPD, using the in2extRemes statistical R software package for extreme values (Gilleland and Katz 2016), with empirical data quantiles, exhibiting also nearly linear trends along with a 95% confidence interval. In addition, Figure 4 (bottom panel) demonstrates the consistency between the empirical density of observed sunspot number maxima and the modeled GPD fit density through corresponding density estimates. It is worth mentioning that the sunspot number values provided by Usoskin et al. (2021) can be formally negative, but are consistent with zero within the uncertainties. Thus, the distribution in Figure 4 (bottom panel) is generally consistent with the theoretical one. For all these reasons, the results shown in Figure 4 confirm the validity of the fitted model.

In the case of the analysis for the two subperiods, the maximum sunspot number is 273.6 in the first subperiod and 237.6 in the second one. We find 22 and 15 clusters after the declustering process for the first and second subperiods, respectively. Note that these values are lower than the number of independent clusters found for the whole period, but sufficient to apply the EVT as the length of both subperiods is lower. Table 1 also includes the values of the scale and shape parameters (95% confidence intervals) applied to these subperiods. As in the case of the analysis in the entire period, we obtained a negative value for the shape parameter. Therefore, this second analysis for the two subperiods also confirms that there is an upper bound of the extreme sunspot number distribution.

Acero et al. (2017) used the block maxima approach with annual data of the sunspot number (version 2) and obtained a scale parameter of 53.48 with a 95% confidence interval

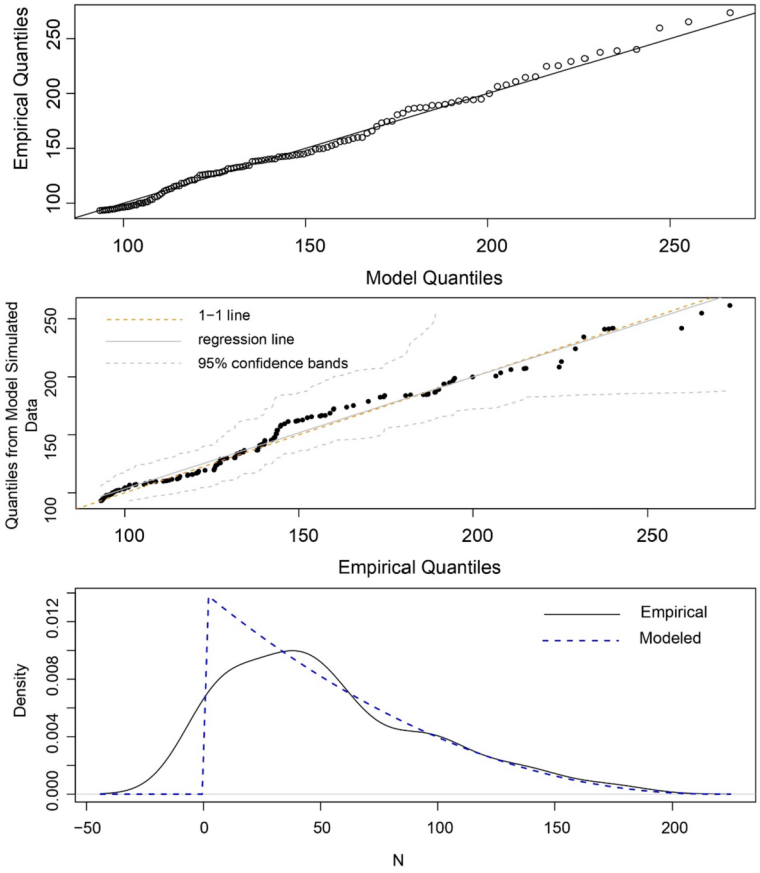

Figure 4 Diagnostic plots from fitting a GPD to the maximum sunspot number values analyzed here. (Top panel) QQ-plot of empirical data quantiles against GPD fit quantiles, (middle panel) QQ-plot of randomly generated data from the fitted GPD against the empirical data quantiles with 95 % confidence bands, and (bottom panel) empirical density of the observed maxima of sunspot number (solid black line) with GPD fit density (dark blue dashed line).

of 37.49 – 69.77. For the shape parameter, they also found a negative value equal to −0.34 and a confidence interval of −0.78 – −0.08. Thus, we have obtained a scale parameter value slightly larger than that by Acero et al. (2017) founding an upper bound in both studies.

We obtained that the solar activity level for the period 991 – 994 AD, when a "Miyake" event occurred in 993 – 994 AD, is one of the clusters identified in this work with extreme values since the solar activity level in those years exceed the defined threshold (Figure 1). This is found both in the analyses for the whole period and for the subperiods. Regarding the candidates for the "Miyake" event, we found that 1052 and 1053 AD are considered extreme events, but the sunspot number in 1279 AD (and any year around it) does not exceed the threshold to be considered an extreme event. This same result is found by the analysis of the subperiods defined above.

We have also estimated the return levels using the overruns of the 30 independent extreme value clusters previously defined for the entire period (Table 2). The maximum sunspot number observed (273.6) is higher than the value obtained for the 1000-year return level (263.5),

Table 2 Estimates of the GPD parameters and their 95 % confidence intervals (CI) obtained by bootstrapping.

Period	1000-year return level	10,000-year return level
971 – 1899 AD	263.5 [240.7, 284.4]	285.2 [253.2, 318.5]
971 – 1435 AD	277.2 [237.1, 317.0]	–
1436 – 1899 AD	226.7 [204.8, 245.4]	–

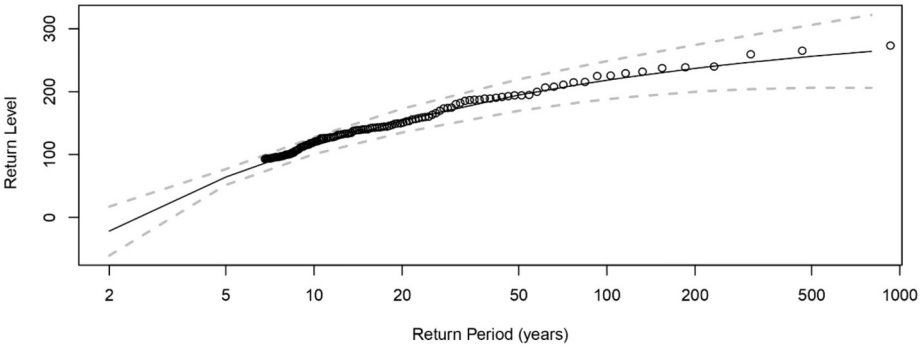

Figure 5 Return level plot for the maximum sunspot number values with 95 % confidence intervals (dashed lines).

although it is within the 95% confidence interval defined for that return level (240.7 – 284.4). Instead, the maximum sunspot number observed is lower than the estimation considering a 10,000-year return level (285.2), but it is also within the 95% confidence interval for that return level (253.2 – 318.5). This estimation is based on a long extrapolation of the low-statistic tail of the distribution. Thus, this result must be taken with caution, since the 10,000-year return level is estimated from a 1000-year-long dataset and it might not capture the full range of variability and extreme events that could occur over a 10,000-year period. We also note that our conclusions are similar to those reached by Acero et al. (2018b), who applied the extreme value theory to decadal sunspot numbers estimated from cosmogenic radioisotopes using a $\approx$ 10,000 year dataset.

In addition, we have estimated the 1000-year return level considering the uncertainties in the sunspot number provided by Usoskin et al. (2021). Including the upper limits of the sunspot numbers in the analysis, we obtain that the 1000-year return level is 383.8 with a 95% confidence interval of 338.7 – 428.8. The maximum sunspot number value considering the upper limit (373.6) is lower than that value, but it is within the interval defined by the 95% confidence interval, such as it occurred in the previous analysis.

The maximum sunspot number observed corresponds to a return period of 3500 years, approximately, which is longer than the study period. These results can be seen in Figure 5, where the return level is shown for each return period with a 95% confidence interval. The return level increases with greater return periods, but leads to a plateau for high return periods. Therefore, values of the sunspot number higher than the observed ones are not expected.

Considering the analysis for the two subperiods, the maximum sunspot number values observed in each subperiod are similar to the estimates found for the 1000-year return level. Moreover, the maximum sunspot number in the first subperiod (273.6) is not within the 95% confidence interval of the 1000-year return level obtained using only data from the second subperiod (204.8 – 245.4), but the maximum in the second subperiod (237.6) is within the

95% confidence interval of the fist subperiod (237.1 – 317.0). We also note that the 1000-year return level found in each subperiod is not within the 95% confidence interval obtained in the other subperiod. That is, 277.2 and 226.7, which are 1000-year return levels for the first and second subperiods, are outside the 95% confidence interval estimated in the second (204.8 – 245.4) and first (237.1 – 317.0) subperiods, respectively.

4. Conclusions

The extreme value theory has been applied to the annual sunspot number series reconstructed from ^{14}C data for the last millennium (Usoskin et al. 2021) to study long-term solar variability. We have used the peaks-over-threshold technique to identify the sunspot number values that exceed a threshold of 93, which is identified as an optimum value. We have also studied solar activity by dividing the entire period into two subperiods with thresholds of 85 and 75 for the periods of 971 – 1435 AD and 1436 – 1899 AD, respectively. The thresholds were selected considering both the mean residual life plot and the parameter estimates for different thresholds. We have made a declustering of these overruns to get independent extreme values, which were modeled using the generalized Pareto distribution.

The shape parameter obtained in this work is significantly negative, both in the analysis of the whole period and by subperiods, in agreement with the result by Acero et al. (2017) for the direct sunspot numbers. This means that there is an upper bound in the extreme values of this series. Furthermore, we have estimated the return levels for 1000 and 10,000 years. Regarding the entire study period, we found that the maximum sunspot number observed in this dataset (273.6 in 981 AD) is within the range considering 95% confidence intervals: 240.7 – 284.4 in the case of 1000-year return level and 253.2 – 318.5 in the case of 10,000-year return level. This implies that the analyzed 1000-year-long sunspot series is representative of the long-term solar variability and covers the full range of the values over at least ten millennia. We must be cautious with this result, because the 10,000-year return level is derived from an extrapolation of the significantly low-statistic tail of the distribution, which is defined using a dataset spanning 1000 years.

The return period for the highest sunspot number observed corresponds to approximately 3500 years. In the case of the analysis considering the two subperiods, we found that the 1000-year return level obtained for each subperiod is not within the range defined by the return level obtained in the other subperiod considering a 95% confidence interval. This result highlights the importance of the long-term series analysis on centennial and millennial scales to better understand the solar activity of the present and future.

We have identified that the Miyake event (extreme solar particle event, ESPE) of 993 – 994 AD corresponds to an extreme solar activity episode defined in this work. Regarding the two other Miyake event candidates in the study period, the one in 1052 – 1053 AD also corresponds to a sunspot number extreme, while that was not the case for the event of 1279 AD. We also note that the solar activity level in sunspot number terms estimated from ^{14}C measurements for years when Miyake events occurred should be taken with caution, since Miyake events are caused by extreme flares, which do not necessarily mean extreme values of sunspot number. This is similar to strong geomagnetic storms, which do not necessarily occur at the solar cycle maximum, as illustrated by the curious geomagnetic storm of 1903, very close to the solar cycle minimum, studied by Ribeiro et al. (2016) and Hayakawa et al. (2020). Moreover, the sunspot numbers around the Miyake events are less accurate since the events can distort the ^{14}C signal and consequently the reconstructed sunspot number.

It is unlikely that the maximum sunspot number values in the near future exceed (at least significantly) the sunspot number observed. This result agrees with the conclusions drawn in similar works such as Acero et al. (2017, 2018b). The methodology and results of this work are limited by the sunspot number values for the last millennium, including four grand minimum periods: Oort, Wolf, Spörer, and Maunder (Usoskin 2023).

Acknowledgements This research has been supported by the Junta de Extremadura (grant no. GR24049). I.U. acknowledges the Research Council of Finland (Project No.354280) and ISSI visiting fellowship.

Author contributions J.M. Vaquero had the idea of the manuscript. F.J. Acero made the calculations included in the manuscript. V.M.S. Carrasco mostly wrote the manuscript. All the authors discussed the results and reviewed the manuscript.

Funding Information Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature.

Data Availability The reconstructed sunspot numbers used in this study are available at the website: https://cdsarc.u-strasbg.fr/viz-bin/cat/J/A+A/649/A141.

Declarations

Competing interests The authors declare no competing interests.

References

Acero, F.J., Gallego, M.C., García, J.A.: 2012, Multi-day rainfall trends over the Iberian Peninsula. *Theor. Appl. Climatol.* **108**, 411. DOI.

Acero, F.J., García, J.A., Gallego, M.C.: 2011, Peaks-over-threshold study of trends in extreme rainfall over the Iberian Peninsula. *J. Climate* **24**, 1089. DOI.

Acero, F.J., García, J.A., Gallego, M.C., Parey, S., Dacunha-Castelle, D.: 2014, Trends in summer extreme temperatures over the Iberian Peninsula using nonurban station data. *J. Geophys. Res., Atmos.* **119**, 39. DOI.

Acero, F.J., Carrasco, V.M.S., Gallego, M.C., García, J.A., Vaquero, J.M.: 2017, Extreme value theory and the new sunspot number series. *APJ* **839**, 98. DOI.

Acero, F.J., Gallego, M.C., García, J.A., Usoskin, I.G., Vaquero, J.M.: 2018b, Extreme value theory applied to the millennial sunspot number series. *APJ* **853**, 80. DOI.

Acero, F.J., Vaquero, J.M., Gallego, M.C., García, J.A.: 2018a, A limit for the values of the Dst geomagnetic index. *Geophys. Res. Lett.* **45**, 9435. DOI.

Aparicio, A.J.P., Carrasco, V.M.S., Vaquero, J.M.: 2023, Prediction of the maximum amplitude of Solar Cycle 25 using the ascending inflection point. *Solar Phys.* **298**, 100. DOI.

Arlt, R., Vaquero, J.M.: 2020, Historical sunspot records. *Living Rev. Solar Phys.* **17**, 1. DOI.

Asensio, A.: 2007, Extreme value theory and the solar cycle. *A&A* **472**, 293. DOI.

Beguería, S., Vicente-Serrano, S.M.: 2006, Mapping the hazard of extreme rainfall by peaks over threshold extreme value analysis and spatial regresion techniques. *J. Appl. Meteorol. Climatol.* **45**, 108. DOI.

Bhowmik, P., Nandy, D.: 2018, Prediction of the strength and timing of sunspot cycle 25 reveal decadal-scale space environmental conditions. *Nat. Commun.* **9**, 5209. DOI.

Brehm, N., Bayliss, A., Christl, M., Synal, H.-A., Adolphi, F., Beer, J., Kromer, B., Muscheler, R., Solanki, S.K., Usoskin, I.G., Bleicher, N., Bollhalder, S., Tyers, C., Wacker, L.: 2021, Eleven-year solar cycles over the last millennium revealed by radiocarbon in tree rings. *Nat. Geosci.* **14**, 10. DOI.

Castillo, E., Hadi, A.S., Balakrishnan, N., Sarabia, J.M.: 2004, *Extreme Value and Related Models in Engineering and Science Applications*, Wiley-Interscience, New York.

Clette, F., Lefèvre, L.: 2016, The new sunspot number: assembling all corrections. *Solar Phys.* **291**, 2629. DOI.

Clette, F., Lefèvre, L., Chatzistergos, T., Hayakawa, H., Carrasco, V.M.S., Arlt, R., Cliver, E.W., Dudok de Wit, T., Friedli, T.K., Karachik, N., Kopp, G., Lockwood, M., Mathieu, S., Muñoz-Jaramillo, A., Owens, M., Pesnell, P.A., Svalgaard, L., Usoskin, I.G., van Driel-Gesztelyi, L., Vaquero, J.M.: 2023, Recalibration of the sunspot-number: status report. *Solar Phys.* **298**, 44. DOI.

Cliver, E.W., Schrijver, C.J., Shibata, K., Usoskin, I.G.: 2022, Extreme solar events. *Living Rev. Solar Phys.* **19**, 2. DOI.

Coles, S.: 2001, *An Introduction to Statistical Modeling of Extreme Values*, Springer, London.

Gilleland, E., Katz, R.W.: 2016, in2extRemes: into the R Package extRemes – Extreme Value Analysis for Weather and Climate Applications. Technical report, National Center for Atmospheric Research. DOI.

Hayakawa, H., Ribeiro, P., Vaquero, J.M., Gallego, M.C., Knipp, D.J., Mekhaldi, F., Bhskar, A., Oliveira, D.M., Notsu, Y., Carrasco, V.M.S., Caccavari, A., Veenadhari, B., Mukherjee, S., Ebihara, Y.: 2020, The extreme space weather event in 1903 October/November: an outburst of quiet sun. *Astrophys. J. Lett.* **897**, L10. DOI.

Hoyt, D.V., Schatten, K.H.: 1998, Group sunspot numbers: a new solar activity reconstruction. *Solar Phys.* **179**, 189. DOI. ADS

Khazanov, G.V.: 2016, *Space Weather Fundamentals*, CRC Press, New York.

Leadbetter, M.R., Weissman, I., Haan, L.D., Rootzen, H.: 1989, On clustering of high values in statistically stationary series. Technical report, Centre for Stochastic Processes, University of North Carolina.

McIntosh, S.W., Chapman, S., Leamon, R.J., Egeland, T., Watkins, N.W.: 2020, Overlapping magnetic activity cycles and the sunspot number: forecasting sunspot cycle 25 amplitude. *Solar Phys.* **295**, 163. DOI.

Miyake, F., Masuda, K., Nakamura, T.: 2013, A signature of cosmic-ray increase in ad 774 – 775 from tree rings in Japan. *Nat. Commun.* **4**, 1748. DOI.

Miyake, F., Nagaya, K., Masuda, K., Nakamura, T.: 2012, A signature of cosmic-ray increase in ad 774 – 775 from tree rings in Japan. *Nature* **486**, 240. DOI.

Muñoz-Jaramillo, A., Vaquero, J.M.: 2019, Visualization of the challenges and limitations of the long-term sunspot number record. *Nat. Astron.* **3**, 205. DOI.

Nandy, D.: 2021, Progress in solar cycle predictions: sunspot cycles 24 – 25 in perspective. *Solar Phys.* **296**, 54. DOI.

Pesnell, W.D.: 2012, Solar cycle predictions. *Solar Phys.* **281**, 507. DOI.

Petrovay, K.: 2020, Solar cycle prediction. *Living Rev. Solar Phys.* **17**, 2. DOI.

Pulkkinen, T.: 2007, Space weather: terrestrial perspective. *Living Rev. Solar Phys.* **4**, 1. DOI.

Ribeiro, P., Vaquero, J.M., Gallego, M.C., Trigo, R.M.: 2016, The first documented space weather event that perturbed the communication networks in Iberia. *Space Weather* **14**, 464. DOI.

Rodríguez, J.V., Rodríguez, I.: 2020, Machine learning-based prediction of sunspots using Fourier transform analysis of the time series. *Publ. Astron. Soc. Pac.* **134**, 124201. DOI.

Siscoe, G.L.: 1976, On the statistics of the largest geomagnetic storms per solar cycle. *J. Geophys. Res.* **81**, 4782. DOI.

Svalgaard, L., Schatten, K.H.: 2016, Reconstruction of the sunspot group number: the backbone method. *Solar Phys.* **291**, 2653. DOI.

Usoskin, I.G.: 2023, A history of solar activity over millennia. *Living Rev. Solar Phys.* **20**, 2. DOI.

Usoskin, I.G., Kromer, B., Ludlow, F., Beer, J., Friedich, M., Kovaltsov, G.A., Solanki, S.K., Wacker, L.: 2013, The AD775 cosmic event revisited: the Sun is to blame. *A&A* **552**, L3. DOI.

Usoskin, I.G., Kovaltsov, G.A., Lockwood, M., Mursula, K., Owens, M., Solanki, S.K.: 2016, A new calibrated sunspot group series since 1749: statistics of active day fractions. *Solar Phys.* **291**, 2685. DOI.

Usoskin, I.G., Solanki, S.K., Krivova, N.A., Hofer, B., Kovaltsov, G.A., Wacker, L., Brehm, N., Kromer, B.: 2021, Solar cyclic activity over the last millennium reconstructed from annual 14C data. *A&A* **649**, A141. DOI.

Usoskin, I.G., Miyake, F., Baroni, M., Brehm, D.S.N., Hayakawa, H., Hudson, H., Timothy Jull, A.J., Knipp, D., Koldobskiy, S., Maehara, H., Mekhaldi, F., Notsu, Y., Poluianov, R.E.S., Shapiro, A., Spiegl, T., Sukhodolov, T., Uusitalo, J., Wacker, L.: 2023, Extreme solar events: setting up a paradigm. *Space Sci. Rev.* **219**, 73. DOI.

Vaquero, J.M., Svalgaard, L., Carrasco, V.M.S., Clette, F., Lefèvre, L., Gallego, M.C., Arlt, R., Aparicio, A.J.P., Richard, J.-G., Howe, R.: 2016, A revised collection of sunspot group numbers. *Solar Phys.* **291**, 3061. DOI.

Wolf, R.: 1860, Mittheilungen über die Sonnenflecken. *Vierteljahrsschr. Nat.forsch. Ges. Zür.* **1**, 1.
Zhang, R., Chen, Y.Q., Zheng, S.G., Zheng, S., Xiao, Y.S., Deng, L.H., Zeng, X.Y., Huang, Y.: 2023, Extreme value theory applied to long-term sunspot areas. *J. Astrophys. Astron.* **45**, 14. DOI.