



**CENTRE DE RECERCA MATEMÀTICA**

Title:	<i>Probability estimation of a Carrington-like geomagnetic storm</i>
Journal Information:	<i>Scientific Reports,</i>
Author(s):	Moriña D., Serra I., Puig P., Corral Á..
Volume, pages:	9 1, DOI:[10.1038/s41598-019-38918-8]

# SCIENTIFIC REPORTS

OPEN

## Probability estimation of a Carrington-like geomagnetic storm

David Morina<sup>1,2,3</sup>, Isabel Serra<sup>1,2,4,5</sup>, Pedro Puig<sup>1,2</sup> & Álvaro Corral<sup>1,2,4,6</sup>

**Intense geomagnetic storms can cause severe damage to electrical systems and communications. This work proposes a counting process with Weibull inter-occurrence times in order to estimate the probability of extreme geomagnetic events. It is found that the scale parameter of the inter-occurrence time distribution grows exponentially with the absolute value of the intensity threshold defining the storm, whereas the shape parameter keeps rather constant. The model is able to forecast the probability of occurrence of an event for a given intensity threshold; in particular, the probability of occurrence on the next decade of an extreme event of a magnitude comparable or larger than the well-known Carrington event of 1859 is explored, and estimated to be between 0.46% and 1.88% (with a 95% confidence), a much lower value than those reported in the existing literature.**

Solar activity, through solar flares, coronal mass ejections, and solar wind, strongly influences the Earth's magnetosphere, triggering there a range of complex phenomena such as geomagnetic storms and aurorae. These phenomena are characterized by highly intermittent and turbulent dynamics, and have also been related to a self-organized-critical state<sup>1</sup>. A geomagnetic storm is a disturbance in the magnetosphere quantified by changes in the Dst (disturbance-storm time) index<sup>2,3</sup>. This index measures the globally averaged change of the horizontal component of the Earth's magnetic field at the magnetic equator and it is recorded once per hour<sup>4</sup>. During quiet times, the Dst index varies between  $-20$  and  $+20$  nT (nanotesla).

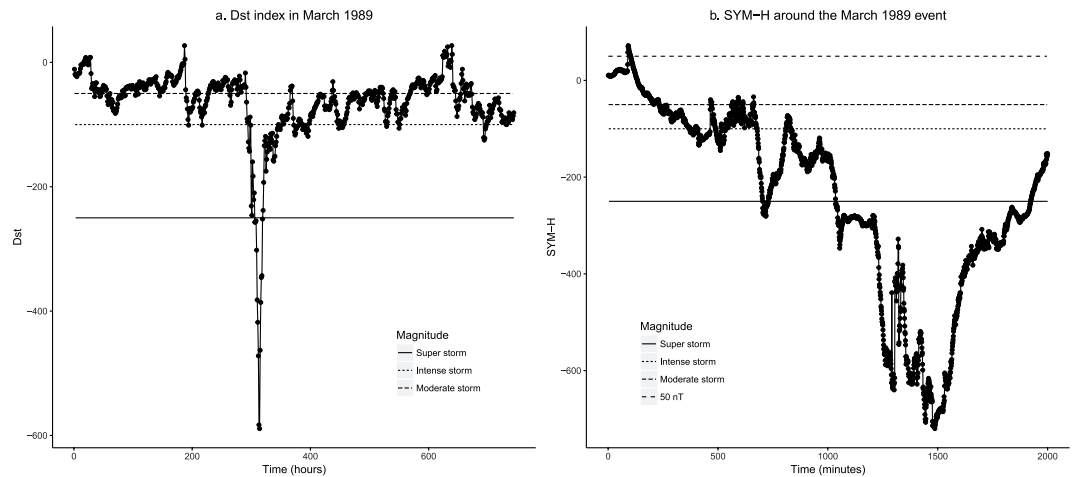
Three phases can usually be identified in a geomagnetic storm<sup>2</sup>. The initial phase, which is not present in all storms and is also known as storm sudden commencement, is characterized by Dst (or its one-minute component, SYM-H) increasing by 20 to 50 nT in a few minutes (but notice that not all sudden increases lead to a storm). In the main phase, whose duration is typically several hours, the Dst decreases to less than  $-50$  nT (the more it decreases, the more intense the storm). Finally, the Dst returns to its quiescent-time range in the recovery phase, a process that can be quite erratic and may last from hours to days. In general, two consecutive storms with less than 48 hours of difference between them are considered to be the same<sup>5</sup>. Figure 1(a) shows the time evolution of the Dst index in March 1989, including the most intense event on record; the storm sudden commencement at the beginning of the event can be noticed on Fig. 1(b), showing the SYM-H evolution.

The minimum Dst value reached during the main phase (always negative) will give the intensity of the storm (also referred to as magnitude). This allows one to distinguish three categories of geomagnetic storms: super-storms (under  $-250$  nT), intense (under  $-100$  nT) and moderate (under  $-50$  nT). The introduction of an intensity threshold in the original Dst time series leads to the fact that the storms defined in this way constitute a point process where the occurrence times of the events are the times of threshold crossing from above. The potential effects of geomagnetic storms on human life are discussed in the next section.

The Dst index has been traditionally modelled by means of its physical properties<sup>3,6</sup>, although some work has also focused on exploring its statistical properties<sup>7</sup>. As far as we know, all efforts in statistical modelling have been based on the assumption that the occurrence of a geomagnetic storm follows an homogeneous Poisson counting process (see for instance ref.<sup>3</sup>). In general, a counting process  $N(t)$  arises when dealing with the number of events occurring over time, and its rate  $\lambda(t)$  is a stochastic process defined by

<sup>1</sup>Barcelona Graduate School of Mathematics (BGSMath), Edifici C, Campus Bellaterra, E-08193, Barcelona, Spain.

<sup>2</sup>Departament de Matemàtiques, Universitat Autònoma de Barcelona (UAB), E-08193, Barcelona, Spain. <sup>3</sup>Unit of Infections and Cancer - Information and Interventions (UNIC - I&I), Cancer Epidemiology Research Program (CERP), Catalan Institute of Oncology (ICO)-IDIBELL, L'Hospitalet de Llobregat, Barcelona, Spain. <sup>4</sup>Centre de Recerca Matemàtica, Edifici C, Campus Bellaterra, E-08193, Barcelona, Spain. <sup>5</sup>Barcelona Supercomputing Center, E-08034, Barcelona, Spain. <sup>6</sup>Complexity Science Hub Vienna, Josefstadt Straße 39, 1080, Vienna, Austria. Correspondence and requests for materials should be addressed to D.M. (email: [david.morina@uab.cat](mailto:david.morina@uab.cat))



**Figure 1.** Evolution of the Dst index in March 1989 (a) and values of the SYM-H index for the extreme geomagnetic storm of that month (b). Horizontal lines mark different thresholds. Time origins are 1989-03-01 (a) and 1989-03-13 (b).

$$\lambda(t) = \frac{1}{dt} P(N(t + dt) - N(t) = 1 | \text{past}), \quad (1)$$

where  $P$  denotes probability (with the vertical bar for conditional probability) and  $N(t)$  counts the number of events from the time origin up to time  $t$ . The simplest case arises for a homogeneous Poisson (counting) process, in which the rate  $\lambda(t) = \lambda$  is constant, and this leads to the fact that the times between consecutive events, known as inter-occurrence times, are independent and exponentially distributed<sup>8</sup>. For a non-homogeneous Poisson process,  $\lambda(t)$  is still independent on the past events, but evolves in time in a predetermined way. This independence constitutes the well-known lack-of-memory property of the Poisson process.

As shown in this paper, the homogeneous-Poisson assumption for the occurrence of geomagnetic storms is not supported by statistical testing. In consequence, we propose a model based on a Weibull counting process, a generalization that contains the homogeneous Poisson process as a particular case (an analogous model was introduced and applied to demographic data in ref.<sup>9</sup>). By means of the Weibull inter-occurrence times, the model allows us to compute the probability of threshold surpassing of future values of the Dst index over different time windows, and can be extrapolated to the most extreme geomagnetic storms thanks to the stability of the shape parameter and the exponential dependence of the Weibull scale parameter with the intensity threshold.

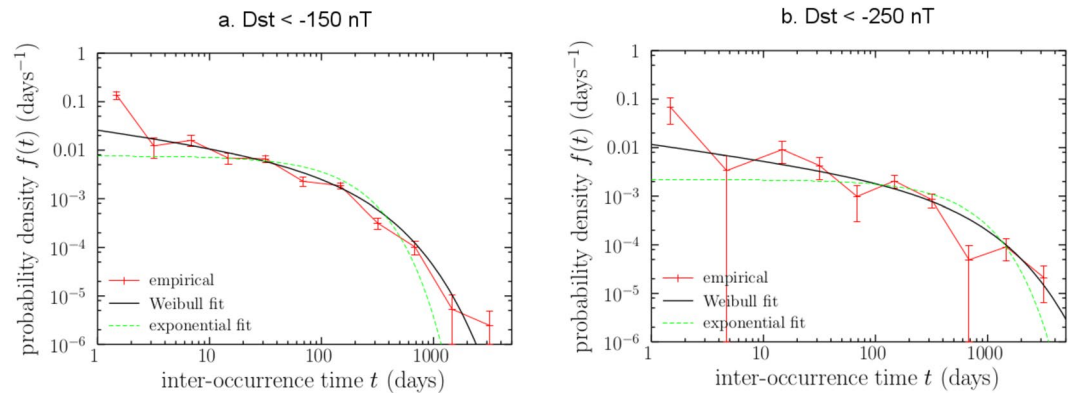
## Effects of Geomagnetic Storms

Geomagnetic storms of remarkable intensity may have a potential effect on many aspects of human life, including health effects due to radiation hazards, disruption of electrical systems like communications appliances, navigation systems, satellites... The biological mechanisms altered by geomagnetic storms are explored in ref.<sup>10</sup>, and geomagnetic activity has been recently related to several human conditions as hypertension<sup>11</sup>, multiple sclerosis<sup>12–14</sup> or cancer<sup>15</sup>. However, it is still unclear to what extend and under what threshold these health effects appear or the risk increases. Regarding communications, in the period under study the most intense event happened on March 1989, reaching a peak Dst under  $-600$  nT, Fig. 1, which caused the collapse of the Hydro-Québec power grid, and resulted in the loss of electrical power to six million people<sup>16</sup>. Another event on July 15–16, 2000 with a peak Dst under  $-300$  nT did not cause any terrestrial damage. Therefore, it seems reasonable to infer that the threshold to expect immediate observable consequences might be between these two values. In fact, some authors have speculated about the potential effects of an extreme geomagnetic event on today's society<sup>17,18</sup>.

The Carrington event is the largest known example of geomagnetic storm, occurred by the end of August and early September 1859 and is associated to a minimum Dst under  $-850$  nT<sup>19</sup>. Richard C. Carrington was observing sunspots on the solar disk and saw a large solar flare with optical brightness equaling that of the background sun, lasting several minutes and due to the destabilization of a large region of the sun causing an extremely fast coronal mass ejection towards Earth. These observations allowed to notice a delay of about 18 hours between the flare and the storm, explained in the recent literature in terms of propagation of the solar wind<sup>20</sup>. It is known that this geomagnetic storm had an effect on the telegraphic network, as several stations reported issues during these days, in different locations around the globe. Some stations were unusable for some hours, but there is no evidence that this interruption on the telegraph service had any impact on the economic activity<sup>18</sup>.

## Data and Storm Definition

To analyse the process of temporal occurrence of geomagnetic storms we use the Dst index, recorded hourly from 1957-01-01 to 2017-12-31 and available from the World Data Center for Geomagnetism in Kyoto<sup>21</sup>. When the Dst signal crosses a fixed negative threshold from above this defines the occurrence time or starting time of a geomagnetic storm with an intensity limited by the threshold; the storm ends when the signal gets over the



**Figure 2.** Empirical and fitted Weibull distributions for the time between consecutive geomagnetic storms, defined using two different thresholds,  $T = -150$  nT (**a**) and  $T = -250$  nT (**b**). The distributions are represented in terms of their probability densities,  $f(t)$ . The empirical ones have been estimated with the so-called “log-binning” method<sup>56</sup>. Exponential fits, shown for comparison, underestimate the probability of both short and long times. Times below two days (disregarded in our approach and also in these fits) are grossly underestimated by any of the models. Weibull fits use the parameters coming from Eq. (2). Plotted using Gnuplot (version 3.7), see ref.<sup>57</sup>.

threshold again. In this paper we consider different thresholds, ranging between  $-50$  nT to  $-400$  nT in steps of  $-10$  nT, although as discussed in the next section only storms under  $-150$  nT are used to build the model that allows forecasting.

The time between two consecutive storms (for a given threshold in Dst) is just the difference of their occurrence times, which defines the inter-occurrence time (measured in days, unless otherwise stated, and called waiting time in refs<sup>22–25</sup>). As the minimum inter-occurrence time to consider two storms as different (and not part of the same storm) is usually taken to be two days<sup>5</sup>, inter-occurrence times below this value are discarded in our procedure. This prevents overcounting arising from small spurious events that appear when the Dst signal remains around the threshold accidentally. As a sensitivity measure, we have explored alternative storms definitions, changing the inter-occurrence time threshold from two days to other (fixed) values, or defining a storm only if its duration is larger than a given fixed period (i.e., an event is considered as such only if the index remains below the threshold for at least 6 hours, for example). These different approaches lead to very similar final results.

## The Model

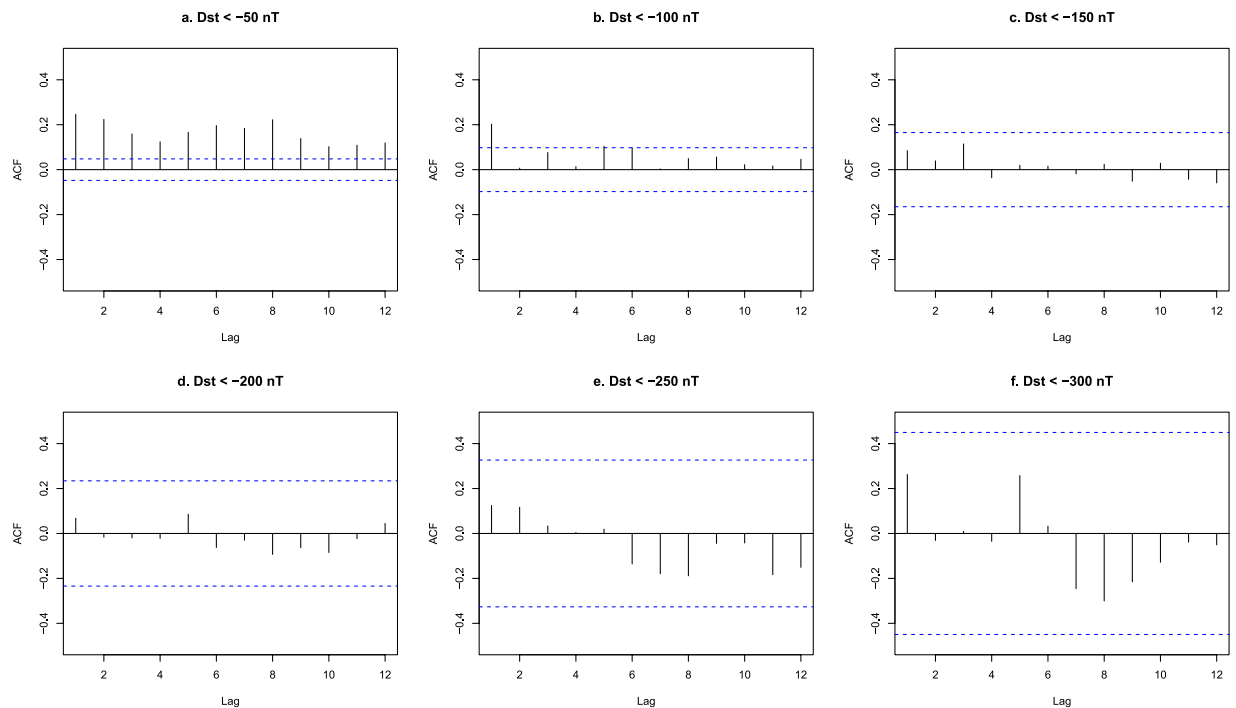
As already mentioned, the time of occurrence of geomagnetic storms has been traditionally analysed in the literature as a homogeneous Poisson process<sup>3</sup>. However, the homogeneous-Poisson-process hypothesis can be tested by checking if inter-occurrence times are exponential, and according to Fig. 2, this distribution seems to be far from doing the fitting optimally. Instead, inter-occurrence times seem to be well fitted by Weibull distributions (see again Fig. 2). In terms of the complementary cumulative distribution function,  $S(t) = P(X > t)$ , the Weibull distribution takes the form

$$S(t) = e^{-(t/\tau)^\gamma},$$

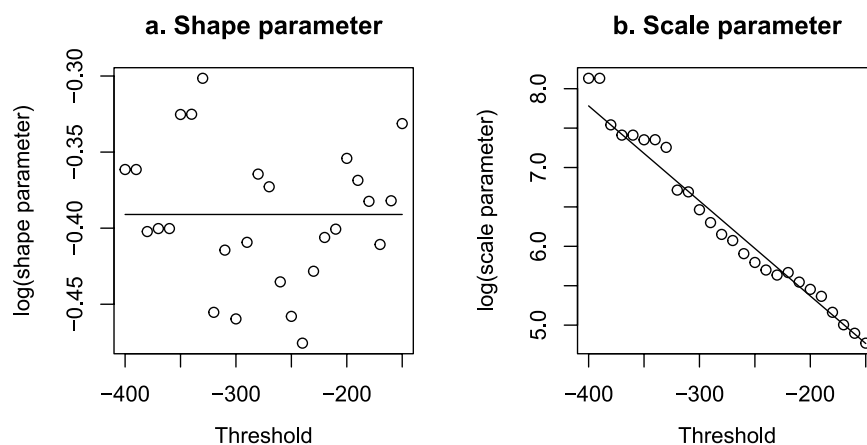
where  $X$  is the random variable representing inter-occurrence time,  $t$  a concrete value,  $\gamma$  the shape parameter and  $\tau$  the scale parameter. The Weibull probability density is obtained as  $f(t) = -dS(t)/dt$  and the rate, or hazard rate, as  $\lambda(t) = f(t)/S(t)$ . Our choice of the Weibull distribution is based on purely empirical grounds, as a common generalization of the homogeneous Poisson process, which is recovered as a particular case just setting  $\gamma = 1$ . Nevertheless, in the Discussion Section we speculate about the origin of the Weibull distribution (a derivation of it in the context of catastrophic failure of materials is given in ref.<sup>26</sup>).

For instance, for super-storms ( $\text{Dst} < -250$  nT), a Weibull distribution with shape parameter  $\gamma = 0.63$  and scale parameter  $\tau = 329$  days fits well the inter-occurrence times; this can be verified by means of a Lilliefors-corrected Kolmogorov-Smirnov test (p-value = 0.11) conducted using the R package *KScorrect*<sup>27</sup>. By means of the same technique, it can also be seen that an exponential distribution does not fit the data properly (p-value < 0.0001, which rules out the homogeneous Poisson process). As the shape parameter  $\gamma$  is below one, this Weibull distribution has a decreasing failure rate (DFR) or decreasing hazard rate, and therefore the associated count distribution (number of storms within this threshold in a period of time) should be overdispersed<sup>28</sup>, confirmed by a dispersion index  $d = \frac{\sigma^2}{\mu} = 2.10$  (i.e., above one), where  $\sigma^2$  and  $\mu$  are, respectively, the sample variance and the sample mean of the number of super-storms in one year during the study period.

Regarding correlations in the inter-occurrence times, geomagnetic storms defined by intensity thresholds equal or below  $-150$  nT show no significant temporal autocorrelations, as shown in Fig. 3 (the confidence intervals showed in Fig. 3 are computed taking into account that the autocorrelation coefficient at lag  $k$ ,  $\hat{\rho}_k$ , under the null hypothesis that the observations are independent and identically distributed, is asymptotically normal with variance  $\text{VAR}(\hat{\rho}_k) \sim \frac{1}{N}$ , where  $N$  is the number of observations, see ref.<sup>29</sup>); nevertheless, significant correlations



**Figure 3.** Autocorrelation function of the inter-occurrence times for moderate (a), intense (b), Dst < −150 nT (c), Dst < −200 nT (d), super-storms (e) and Dst < −300 nT (f). The horizontal axis (lag) counts storm separation in terms of number of storms. Horizontal lines denote the limits of 95% confidence intervals.



**Figure 4.** Relationship between Dst threshold (in nT) and Weibull shape (a) and scale (b) parameters, in log-scale, with scale parameter in days. Intensity thresholds range from −400 nT to −150 nT. The points correspond to maximum-likelihood estimates of the shape and scale parameters for fixed threshold values. The lines are the result of a Weibull regression model which is fitted directly to the whole data set of inter-occurrence times and thresholds (not to the shape and scale parameters).

are found for weaker events. This prevents the application of some statistical techniques to these weaker storms; therefore, to avoid temporal correlations, in what follows we focus on geomagnetic storms defined by a Dst threshold equal to or below −150 nT. These storms are, in fact, the ones that may have a deeper impact on human activity, as discussed in a previous section. However, it is important to remark that it is not that there are no correlations between the inter-occurrence times of the more intense storms, but that these correlations cannot be established as statistically significant.

**Forecasting.** The proposed model can be used to estimate the probability of an event below a certain threshold in Dst to occur in a period of time, based on studying the relationship between the threshold and the Weibull shape and scale parameters. This relationship can be even used to forecast extreme events with a threshold outside the observed range. It can be seen (Fig. 4) that the shape parameter does not depend on the threshold while a linear relationship seems to appear for the logarithm of the scale parameter and the threshold.

Indeed, several models were fitted to the scale parameter using the threshold in Dst as explanatory variable, and the one that showed a better fit was the linear regression model

$$\log(\tau) = \beta_0 + \beta_1 \cdot T. \quad (2)$$

In consequence, the inter-occurrence times were finally fitted using a Weibull regression model (see ref.<sup>30</sup>) where the scale parameter  $\tau$  changes with the threshold according to Eq. 2 and the shape parameter  $\gamma$  is constant (independent of the threshold); in other words, we fit the threshold-dependent distribution

$$S(t; T) = \exp[-(t/e^{\beta_0 + \beta_1 T})^\gamma], \quad (3)$$

using R *survival* package (ref.<sup>31</sup>). The inter-occurrence times analyzed were those arising from the thresholds  $T = -150, -160, -170, \dots$  up to  $-400$  nT. Both scale and shape parameters can be recovered from the Weibull regression model ( $\tau = e^{\beta_0 + \beta_1 T}$ , in units of days). The estimates are  $\log(\hat{\gamma}) = -0.39$  ( $SE = 0.023$ ),  $\hat{\beta}_0 = 2.96$  ( $SE = 0.17$ ) and  $\hat{\beta}_1 = -0.0121$  ( $SE = 0.0008$ , units of  $\text{nT}^{-1}$ ).

Regression models based on other distributions might be appropriate according to the behavior described in Fig. 2, such as a gamma regression model, which is implicit in previous works on earthquakes and solar flares (see refs.<sup>32–34</sup> and in concrete Eq. 5 of ref.<sup>35</sup> and Fig. 13 of ref.<sup>36</sup>), although not recognized there as regression models and not fitted as such. In order to justify our final choice of the Weibull regression model, the goodness of fit of the Weibull and the gamma distributions for fixed thresholds have been compared by means of the Akaike Information Criterion (AIC), showing that the fit of the Weibull model is better regardless of the threshold. The details can be found in the supplementary material (Table S1), comparing also with the standard exponential distribution, which showed the worst fit.

**Forecasting another Carrington event.** Some efforts have been recently devoted to estimate the probability of another Carrington event in the next 10 years<sup>3</sup>. According to our results, the estimated shape parameter of the corresponding Weibull distribution is  $\hat{\gamma} = e^{-0.39} = 0.68$ , and Eq. 2 with threshold  $T = -850$  nT leads to the scale parameter,  $\hat{\tau} = 563882$  days (1544 years). Knowing that the original Carrington event happened in 1859, about 58000 days ago, one can compute the probability of having a Carrington or more intense event during the next decade (2018–2027) conditioned to the fact that no event like this has happened since 1859,

$$P(X \leq t_C + t_d | X \geq t_C) = \frac{S(t_C) - S(t_C + t_d)}{S(t_C)} = 1 - \exp\left[\left(\frac{t_C}{\tau}\right)^\gamma - \left(\frac{t_C + t_d}{\tau}\right)^\gamma\right] = 0.0092, \quad (4)$$

with  $t_C = 58000$  days and  $t_d = 3652$  days (10 years), and so, the probability is 0.92%. Note that the value reported in ref.<sup>3</sup> was about 12%, in sharp contrast with our result. This important difference can be explained in part by the more flexibility of the Weibull model and in part by the fact that the tail of the Dst distribution seems to be better fitted by an exponential distribution than by a power law (p-value = 0.1 and p-value < 0.0001 respectively, but other distributions have been proposed<sup>37</sup>), according to the methodology described in ref.<sup>38</sup> and implemented in the R package *powerLaw*<sup>39</sup>. Note that the exponential behavior of the tail of the distribution is directly related to Eq. 2.

In addition to the probability point estimate, an approximation of its standard error can be obtained by bootstrapping inter-occurrence times and thresholds and fitting the described Weibull regression model for each bootstrapped data set. Using this method the estimated standard error of the resulting probability is  $SE = 0.0037$ , and a 95% confidence interval can be built by computing percentiles 2.5% and 97.5%, leading to the interval (0.46%, 1.87%). Taking into account that the probability in Eq. 4 is a function of  $\gamma$  and  $\tau$ , with the latter depending on  $\beta_0$  and  $\beta_1$ , its standard error could also have been estimated by using the *delta method*, leading to a very similar result of  $SE = 0.0031$ . The R package *car*<sup>40</sup> was used with this purpose, and the R code for both methods is available as supplementary material. In the latter case, however, for the construction of the corresponding confidence interval asymptotic normality needs to be assumed but the distribution of the probabilities given by Eq. 4 obtained by bootstrapping is not symmetrical, so the bootstrap confidence interval is preferred.

The Carrington event was first estimated to have a Dst of about  $-1760$  nT<sup>41</sup>. If we assume this threshold, the corresponding Weibull distribution would have shape and scale parameters 0.68 and  $3.18 \times 10^{10}$  days ( $8.7 \times 10^7$  years) respectively, so the probability of having such an event within the next decade would be 0.0005%, with a 95% confidence interval ( $9.11 \times 10^{-5}\%$ ; 0.0033%). A probability of 1.5% was reported in ref.<sup>3</sup>, again much higher than the values in our interval.

Using the expectation of the underlying Weibull counting process  $N(t)$ , the expected number of storms with intensity below a specific threshold can be calculated for any time period  $t$ , and is given by the expression<sup>9</sup>

$$E(N(t)) = \sum_{n=1}^{\infty} \sum_{j=n}^{\infty} \frac{n(-1)^{j+n} (t/\tau)^{\gamma j} \alpha_j^n}{\Gamma(\gamma j + 1)}, \quad (5)$$

where  $\alpha_j^0 = \frac{\Gamma(\gamma j + 1)}{\Gamma(j + 1)}$ ,  $j = 0, 1, \dots$  and  $\alpha_j^{n+1} = \sum_{m=n}^{j-1} \alpha_m^n \frac{\Gamma(\gamma j - \gamma m + 1)}{\Gamma(j - m + 1)}$ ,  $n = 0, 1, \dots$  and  $j = n + 1, n + 2, \dots$ . The formula is exact when a below-threshold event (which is not counted) occurred at the beginning of the time period.

For instance, refs.<sup>17,42</sup> reported a table including the estimated frequencies for geomagnetic storms with different thresholds. Our method yields substantially different results, as shown in Table 1. To obtain an estimation within the same order of magnitude, one could also use the asymptotic approximation  $E(N(t)) = \frac{t}{m}$ , where  $m$  is the average inter-occurrence time which can be calculated using the corresponding Weibull expectation, given by  $\tau \Gamma\left(1 + \frac{1}{\gamma}\right)$ .



Threshold (nT)	Frequency (Eq. 5)
−100	4.95 per 1 year
−200	1.78 per 1 year
−400	1.63 per 10 years
−800	1.37 per 1,000 years
−1600	0.19 per 1,000,000 years

**Table 1.** Estimated frequencies for geomagnetic storms of different thresholds, based on the exact expression (Eq. 5), calculated for the indicated time periods.

## Discussion

Several authors have proposed models in order to forecast the occurrence of geomagnetic storms of certain magnitude<sup>3,7,43,44</sup>, most of them based on physical properties of the Dst index, like in refs<sup>6,45</sup>. To date, the probabilistic approaches to forecast these events have been based, to our knowledge, on the assumption that the number of storms in a given period of time follows an homogeneous Poisson process<sup>3,5</sup>. This assumption might be violated, especially for intense and super-storms, as we have verified following the Lilliefors-corrected Kolmogorov-Smirnov test<sup>27</sup>, as also displayed visually in Fig. 2.

We have introduced a Weibull counting process which includes the homogeneous Poisson process as a particular case. The model is defined, for any intensity threshold, by independent Weibull inter-occurrence times, with a constant (independent on the threshold) and smaller-than-one shape parameter, and a scale parameter depending exponentially on the intensity threshold, Eq. 2. The estimated probability of occurrence of a Carrington or more extreme event ( $Dst < -850$  nT) in the next decade using the proposed model (~0.92%) is much lower than the values estimated in previous works using a different probabilistic approach (ref.<sup>3</sup> reports a probability around 12%) and the estimated probabilities are even more different for higher thresholds. The reason of this strong discrepancy is probably and in part due to the fact that the extrapolation to severe geomagnetic storms of a power-law<sup>46</sup> probability distribution of intensities is problematic<sup>47</sup>. This deviation from the power-law distribution was pointed out in ref.<sup>3</sup> and explored in refs<sup>37,42</sup> by fitting a gamma and a lognormal distribution to the storm maximum Dst, respectively.

In addition, our Weibull regression model is flexible enough to include other covariates, in addition to the Dst thresholds. According to ref.<sup>22</sup>, the geomagnetic storm inter-occurrence-time distribution is strongly related to solar cycle, so a similar model based on Eq. 2 including an extra term to account for the sunspot number was fitted, leading again to a better performance of the Weibull regression model compared to gamma and exponential alternatives and a very similar shape parameter estimate of  $\hat{\gamma} = 0.72$  ( $SE = 0.023$ ). This will be analyzed in depth in future research.

An interesting property of the Weibull distribution is that the hazard rate is just a power-law function,  $\lambda(t) \propto t^{\gamma-1}$ ; this is invariant under change of scale, which is physically suggestive, and the Weibull distribution is the only one having this property<sup>48</sup> (note that power-law distributions should also verify this property with  $\gamma = 0$ , but they are not defined for all  $t > 0$ ). In our study, the estimated Weibull model yields a shape parameter  $\gamma$  below one, leading therefore to a decreasing hazard rate. This means that the longer since the last super-storm happened, the less the chance to happen another one in the next unit of time. Thus, the estimated probability of occurrence of a Carrington event in the first decade after 1859 is 3.33%, which is 3.6 times greater than the estimated probability for the current decade. This property in which hazard decreases over time is not only counter-intuitive but fundamental for hazard assessment, and seems to be present in the occurrence of many geological and astronomical extreme events<sup>23,24,35,49–51</sup>. The origin can be attributed to long-range correlations in the Dst time series, as several authors<sup>50,52</sup> have argued that in such a case the inter-occurrence time statistics is of the decreasing-hazard-rate type. Note that, although the few number of storms available for too extreme Dst thresholds precludes that the correlations between the inter-occurrence time are significant, this does not rule out their existence.

Our findings also indicate some other sort of self-similarity in the temporal occurrence of geomagnetic storms (in addition to the hazard rate function becoming a power law, which is the hallmark of scale invariance for univariate functions). The fact that the shape parameter remains stable under changes in the threshold setting the severity of the storms implies that the temporal occurrence of storms follows the same dynamics for any of the analyzed threshold intensities (except for a trivial change of scale set by the scale parameter  $\tau$ ). It is noteworthy that many natural hazards share this property of self-similarity<sup>53</sup>, which is not linked to the particular model of the inter-occurrence times (i.e., not linked to the Weibull assumption). In the case of point events this has been observed for earthquakes<sup>35,36</sup> and forest fires<sup>54</sup>, whereas for time series (the case under consideration here) this behavior is present in rainfall events<sup>49</sup>, annual climate records (temperature, river floods, etc.<sup>50</sup>), and (of particular interest for its relation with geomagnetic storms) solar flares<sup>23,24</sup>. Further, the fact that the scale parameter  $\tau$  depends exponentially on the threshold  $T$  implies that the waiting time can be rescaled as  $t/e^{\beta_1 T}$ , see Eq. 3. Defining a new variable as  $E \propto e^{\alpha|T|}$ , the same rescaling can be written as  $t/E^{|\beta_1|/\alpha}$ , which implies that Eq. 3 can be rewritten in the form of a time-energy scaling law, as found previously for earthquakes and other natural hazards<sup>36</sup>.

Although there are physical limitations to the possible scale-invariance of geomagnetic storms (due to theoretical upper limits to their size based on the saturation properties of the magnetospheric ring current and polar-cap conductivity, these limitations should affect the distribution of storm size, but not necessarily the temporal properties of the storms, given by Eq. (3)). In other words, this equation gives scale-invariant properties no matter the size distribution of the storms. Note that time-size scale invariance without power-law distributed sizes has been claimed in other systems in geoscience<sup>54</sup>.

Regarding the effects of an extreme geomagnetic storm like the Carrington event, it seems reasonable to think that these would have a deeper impact nowadays than in 1859. For instance, there were no radio communications in 1859 and a global blackout of these and any communications involving ionospheric effects can be expected due to a severe geomagnetic storm. Power distribution networks would probably be the most sensitive equipment as major communication infrastructure is based on optical fibres links, which are much less sensitive to perturbations than the 1859 network. In early November 2003 an important perturbation happened in space (known as the Halloween event), although associated with a higher Dst than Carrington event's<sup>55</sup>. These storms interfered with satellite communications (the Japanese ADEOS-2 was partly damaged) and produced a brief power outage in Sweden. This event and its consequences are described in ref.<sup>17</sup>.

## Data Availability

All data generated or analysed during this study are included in this published article (and its Supplementary Information files).

## References

- Aschwanden, M. *Self-Organized Criticality in Astrophysics*. (Springer, 2011).
- Gonzalez, W. D. *et al.* What is a geomagnetic storm? *Journal of Geophysical Research* **99**, 5771–5792 (1994).
- Riley, P. On the probability of occurrence of extreme space weather events. *Space Weather* **10**, S02012, <https://doi.org/10.1029/2011SW000734> (2012).
- Sugiura, M. & Kamei, T. Equatorial Dst Index 1957–1986. *International Association of Geomagnetism and Aeronomy (IAGA)* **40** (1991).
- Tsubouchi, K. & Omura, Y. Long-term occurrence probabilities of intense geomagnetic storm events. *Space Weather* **5**, 1–12, <https://doi.org/10.1029/2007SW000329> (2007).
- Kataoka, R. Probability of occurrence of extreme magnetic storms. *Space Weather* **11**, 214–218, <https://doi.org/10.1002/swe.20044> (2013).
- Yokoyama, N. & Kamide, Y. Statistical nature of geomagnetic storms. *Journal of Geophysical Research: Space Physics* **102**, 14215–14222, <https://doi.org/10.1029/97JA00903> (1997).
- Daley, D. J. & Vere-Jones, D. *An Introduction to the Theory of Point Processes*. (Springer-Verlag, New York, 1988).
- McShane, B., Adrian, M., Bradlow, E. T. & Fader, P. S. Count models based on Weibull interarrival times. *Journal of Business and Economic Statistics* **26**, 369–378, [http://www.blakemcshane.com/Papers/jbes\\_weibull.pdf](http://www.blakemcshane.com/Papers/jbes_weibull.pdf), <https://doi.org/10.1198/073500107000000278> (2008).
- Krylov, V. V. Biological effects related to geomagnetic activity and possible mechanisms. *Bioelectromagnetics* **38**, 497–510, <http://www.ncbi.nlm.nih.gov/pubmed/28636777>, <https://doi.org/10.1002/bem.22062> (2017).
- Azcárate, T. & Mendoza, B. Influence of geomagnetic activity and atmospheric pressure in hypertensive adults. *International Journal of Biometeorology* **61**, 1585–1592, <http://www.ncbi.nlm.nih.gov/pubmed/28357508>, <https://doi.org/10.1007/s00484-017-1337-x> (2017).
- Sajedi, S. A. & Abdollahi, F. Geomagnetic disturbances may be environmental risk factor for multiple sclerosis: an ecological study of 111 locations in 24 countries. *BMC Neurology* **12**, 100, <http://www.ncbi.nlm.nih.gov/pubmed/22998435>, <https://doi.org/10.1186/1471-2377-12-100> (2012).
- Sajedi, S. A. & Abdollahi, F. Geomagnetic disturbance: A new field in multiple sclerosis research. *Clinical Neurology and Neurosurgery* **151**, 142, <http://www.ncbi.nlm.nih.gov/pubmed/27771023>, <https://doi.org/10.1016/j.clineuro.2016.10.003> (2016).
- Papathanasopoulos, P. *et al.* The possible effects of the solar and geomagnetic activity on multiple sclerosis. *Clinical Neurology and Neurosurgery* **146**, 82–89, <http://www.ncbi.nlm.nih.gov/pubmed/27161905>, <https://doi.org/10.1016/j.clineuro.2016.04.023> (2016).
- Swanson, J. & Kheifets, L. Could the geomagnetic field be an effect modifier for studies of power-frequency magnetic fields and childhood leukaemia? *Journal of Radiological Protection* **32**, 413–418, <http://www.ncbi.nlm.nih.gov/pubmed/23006769>, <https://doi.org/10.1088/0952-4746/32/4/413> (2012).
- Bolduc, L. GIC observations and studies in the Hydro-Québec power system. *Journal of Atmospheric and Solar-Terrestrial Physics* **64**, 1793–1802, [https://doi.org/10.1016/S1364-6826\(2002\)](https://doi.org/10.1016/S1364-6826(2002)).
- Schieb, P. A. & Gibson, A. Geomagnetic Storms, Office of Risk Management and Analysis, United States Department of Homeland Security, report IFP/WKP/FGS, <http://www.oecd.org/gov/risk/46891645.pdf> (2011).
- Muller, C. The Carrington solar flares of 1859: consequences on life. *Origins of life and evolution of the biosphere: the journal of the International Society for the Study of the Origin of Life* **44**, 185–195, <http://www.ncbi.nlm.nih.gov/pubmed/25351684>, <https://doi.org/10.1007/s11084-014-9368-3> (2014).
- Siscoe, G., Crooker, N. U. & Clauer, C. R. Dst of the Carrington storm of 1859. *Advances in Space Research* **38**, 173–179 <https://www.sciencedirect.com/science/article/S0273117705012081>, <https://doi.org/10.1016/j.asr.2005.02.102> (2006).
- Tsurutani, B. T., Gonzalez, W. D., Lakhina, G. S. & Alex, S. The extreme magnetic storm of 1–2 September 1859. *Journal of Geophysical Research* **108**, 1268, <https://doi.org/10.1029/2002JA009504> (2003).
- World Data Center for Geomagnetism. Data analysis center for geomagnetism and space magnetism. <http://wdc.kugi.kyoto-u.ac.jp> (2018).
- Wheatland, M. S. & Litvinenko, Y. E. Understanding Solar Flare Waiting-Time Distributions. *Solar Physics* **211**, 255–274, <https://doi.org/10.1023/A:1022430308641> (2002).
- Paczuski, M., Boettcher, S. & Baiesi, M. Interoccurrence times in the Bak-Tang-Wiesenfeld sandpile model: A comparison with the observed statistics of solar flares. *Phys. Rev. Lett.* **95**, 181102, <https://doi.org/10.1103/PhysRevLett.95.181102> (2005).
- Baiesi, M., Paczuski, M. & Stella, A. L. Intensity thresholds and the statistics of the temporal occurrence of solar flares. *Phys. Rev. Lett.* **96**, 051103, <https://doi.org/10.1103/PhysRevLett.96.051103> (2006).
- Wanliss, J. A. & Weygand, J. M. Power law burst lifetime distribution of the SYM-H index. *Geophys. Res. Lett.* **34**, L04107, <https://doi.org/10.1029/2006GL028235> (2007).
- Yakovlev, G., Turcotte, D. L., Rundle, J. B. & Rundle, P. B. Simulation-based distributions of earthquake recurrence times on the San Andreas Fault System. *Bulletin of the Seismological Society of America* **96**, 1995, <https://doi.org/10.1785/0120050183> (2006).
- Novack-Gottshall, P. & Wang, S. C. KScorrect: Lilliefors-Corrected Kolmogorov-Smirnov Goodness-of-Fit Tests, <https://CRAN.R-project.org/package=KScorrect>. R package version 1.2.3 (2018).
- Barlow, R. E. & Proschan, F. *Statistical Theory of Reliability and Life Testing. Probability Models* (To Begin With, Silver Spring, MD 20904, 1981).
- Chatfield, C. *The Analysis of Time Series: An Introduction (Sixth Edition)*, <https://www.crcpress.com/The-Analysis-of-Time-Series-An-Introduction-Sixth-Edition/Chatfield/p/book/9781584883173> (CRC Press, 2003).
- Hosmer, D. W., Lemeshow, S. & May, S. *Applied survival analysis: regression modeling of time-to-event data*, <http://www.wiley.com/en-us/Applied+Survival+Analysis+%3A+Regression+Modeling+of+Time+to+Event+Data+%2C+2nd+Edition-p-9780471754992> (Wiley-Interscience, 2008).
- Therneau, T. M. & Grambsch, P. *Grambsch. Modeling Survival Data: Extending the Cox Model*. (Springer, New York, 2000).



32. Corral, A. Local distributions and rate fluctuations in a unified scaling law for earthquakes. *Phys. Rev. E* **68**, 035102, <https://doi.org/10.1103/PhysRevE.68.035102> (2003).
33. Lippiello, E., de Arcangelis, L. & Godano, C. Different triggering mechanisms for solar flares and coronal mass ejections. *Astronomy & Astrophysics* **488**, L29–L32, <https://doi.org/10.1051/0004-6361:200810164> (2008).
34. Bottiglieri, M., De Arcangelis, L., Godano, C. & Lippiello, E. Multiple-time scaling and universal behavior of the earthquake interevent time distribution. *Physical Review Letters* **104**, 5–8, <https://doi.org/10.1103/PhysRevLett.104.158501> (2010).
35. Corral, A. Dependence of earthquake recurrence times and independence of magnitudes on seismicity history. *Tectonophysics* **424**, 177–193, <https://doi.org/10.1016/j.tecto.2006.03.035> (2006).
36. de Arcangelis, L., Godano, C., Grasso, J. R. & Lippiello, E. Statistical physics approach to earthquake occurrence and forecasting. *Physics Reports* **628**, 1–91, <https://doi.org/10.1016/j.physrep.2016.03.002> (2016).
37. Love, J. J., Rigler, E. J., Pulkkinen, A. & Riley, P. On the lognormality of historical magnetic storm intensity statistics: Implications for extreme-event probabilities. *Geophysical Research Letters* **42**, 6544–6553, <https://doi.org/10.1002/2015GL064842> (2015).
38. Clauset, A., Shalizi, C. R. & Newman, M. E. J. Power-Law Distributions in Empirical Data. *SIAM Review* **51**, 661–703, <https://doi.org/10.1137/070710111> (2009).
39. Gillespie, C. S. Fitting heavy tailed distributions: The powerLaw package. *Journal of Statistical Software*, **64**, 1–16, <http://www.jstatsoft.org/v64/i02/> (2015).
40. Fox, J. & Weisberg, S. *An R Companion to Applied Regression*, second edn. <http://socserv.socsci.mcmaster.ca/jfox/Books/Companion> (Sage, Thousand Oaks CA, 2011).
41. Lakhina, G. S., Alex, S., Tsurutani, B. T. & Gonzalez, W. D. Research on historical records of geomagnetic storms. Coronal and Stellar Mass Ejections. *IAU Symp. Colloq. Proc. Ser* **226**, 3–15, <https://doi.org/10.1017/S1743921305000074> (2005).
42. Love, J. J. & Gannon, J. L. Revised Dst and the epicycles of magnetic disturbance: 1958–2007. *Annales Geophysicae* **27**, 3101–3131, <https://www.ann-geophys.net/27/3101/2009/>, <https://doi.org/10.5194/angeo-27-3101-2009> (2009).
43. Baker, D. N. *et al.* A major solar eruptive event in July 2012: Defining extreme space weather scenarios. *Space Weather* **11**, 585–591, <https://doi.org/10.1002/swe.20097> (2013).
44. Temerin, M. & Li, X. A new model for the prediction of Dst on the basis of the solar wind. *Journal of Geophysical Research: Space Physics* **107**, 1–13, <https://doi.org/10.1029/2001JA007532> (2002).
45. Kovaltsov, G. A. & Usoskin, I. G. Occurrence Probability of Large Solar Energetic Particle Events: Assessment from Data on Cosmogenic Radionuclides in Lunar Rocks. *Solar Physics* **289**, 211–220, <https://doi.org/10.1007/s11207-013-0333-5> (2014).
46. Norman L. J., Adrienne W. & Kemp, S. K. *Univariate Discrete Distributions - 3rd Edition*, vol. 17, 3 edition edn. <http://linkinghub.elsevier.com/retrieve/S0167947394901155>. arXiv:1011.1669v3 (Wiley, 2005).
47. Serra, I. & Corral, A. Deviation from power law of the global seismic moment distribution. *Sci. Rep.* **7**, 40045, <https://doi.org/10.1038/srep40045> (2017).
48. Dasgupta, R. Characterization theorems for Weibull distribution with applications. *Journal of Environmental Statistics* **6** (2014).
49. Deluca, A., Moloney, N. R. & Corral, A. Data-driven prediction of thresholded time series of rainfall and selforganized criticality models. *Phys. Rev. E* **91**, 052808, <https://doi.org/10.1103/PhysRevE.91.052808> (2015).
50. Bunde, A., Eichner, J. F., Kantelhardt, J. W. & Havlin, S. Long-term memory: a natural mechanism for the clustering of extreme events and anomalous residual times in climate records. *Phys. Rev. Lett.* **94**, 048701, <https://doi.org/10.1103/PhysRevLett.94.048701> (2005).
51. de Arcangelis, L., Godano, C., Lippiello, E. & Nicodemi, M. Universality in solar flare and earthquake occurrence. *Phys. Rev. Lett.* **96**, 051102, <https://doi.org/10.1103/PhysRevLett.96.051102> (2006).
52. Altmann, E. G. & Kantz, H. Recurrence time analysis, long-term correlations, and extreme events. *Phys. Rev. E* **71**, 056106, <https://doi.org/10.1103/PhysRevE.71.056106> (2005).
53. Corral, A. *Scaling in the timing of extreme events*. *Chaos. Solit. Fract.* **74**, 99–112 (2015).
54. Corral, A., Telesca, L. & Lasaponara, R. Scaling and correlations in the dynamics of forest-fire occurrence. *Phys. Rev. E* **77**, 016101, <https://doi.org/10.1103/PhysRevE.77.016101> (2008).
55. Viljanen, A., Myllys, M. & Nevanlinna, H. Russian geomagnetic recordings in 1850–1862 compared to modern observations. *Journal of Space Weather and Space Climate* **4**, A11, <https://doi.org/10.1051/swsc/2014008> (2014).
56. Deluca, A. & Corral, A. Fitting and goodness-of-fit test of non-truncated and truncated power-law distributions. *Acta Geophys.* **61**, 1351–1394, <https://doi.org/10.2478/s11600-013-0154-9> (2013).
57. Williams, T., Kelley, C. & many others. Gnuplot 3.7: an interactive plotting program. <http://gnuplot.sourceforge.net> (1999).

## Acknowledgements

We appreciate the criticisms of Álvaro González and the comments from two anonymous reviewers and editor, that helped improve and clarify some aspects of this manuscript. Our work was partially supported by grants from the Instituto de Salud Carlos III-ISCIII (Spanish Government) cofunded by FEDER funds/European Regional Development Fund (ERDF) - a way to build Europe (References: RD12/0036/0056, PI11/02090), from the Agència de Gestió d'Ajuts Universitaris i de Recerca (2014SGR 756, 2014SGR 1307), from RecerCaixa 2015 (2015ACUP00129) and from the Spanish MINECO (FIS2015-71851-P, MTM2015-69493-R). David Moríña acknowledges financial support from the Spanish Ministry of Economy and Competitiveness, through the María de Maeztu Programme for Units of Excellence in R&D (MDM-2014-0445) and from Fundación Santander Universidades.

## Author Contributions

David Moríña, Isabel Serra, Pedro Puig and Álvaro Corral participated in the analysis of the data, interpretation and discussion of the results, reviewed the manuscript and approved its final form.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-019-38918-8>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019