

Analysis of Japan and World Records in the 100 m Dash Using Extreme Value Theory

Fumio Maruyama

Department of Sports and Health Science, Matsumoto University, Matsumoto, Japan

Email: fmaruya@nagoya-u.jp

How to cite this paper: Maruyama, F. (2021) Analysis of Japan and World Records in the 100 m Dash Using Extreme Value Theory. *Journal of Applied Mathematics and Physics*, 9, 1442-1451.

<https://doi.org/10.4236/jamp.2021.97097>

Received: May 22, 2021

Accepted: July 5, 2021

Published: July 8, 2021

Copyright © 2021 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Extreme value theory provides methods to analyze the most extreme parts of data. We predicted the ultimate 100 m dash records for men and women for specific periods using the generalized extreme value (GEV) distribution. The various diagnostic plots, which assessed the accuracy of the GEV model, were well fitted to the 100 m records in the world and Japan, validating the model. The men's world record had a shape parameter of -0.250 with a 95% confidence interval of $[-0.391, -0.109]$. The 100 m record had a finite limit and a calculated upper limit was 9.46 s. The return level estimates for the men's world record were 9.74, 9.62, and 9.58 s with a 95% confidence interval of $[9.69, 9.79]$, $[9.54, 9.69]$, and $[9.48, 9.67]$ for 10-, 100- and 350-year return periods, respectively. In one year, the probability of occurrence for a new world record of men, 9.58 s (Usain Bolt), was $1/350$, while that for women, 10.49 s (Florence Griffith-Joyner), was about $1/100$, confirming it was more difficult for men to break records than women.

Keywords

Athletics, 100 m Running, Extreme Value Theory, GEV Model

1. Introduction

Extreme value theory (EVT) has emerged as an important statistical discipline in applied science. EVT deals with statistical problems concerning the far tail of the probability distribution and is unique as a statistical tool since it develops models and techniques to describe the unusual event rather than the usual. Using EVT, the theoretical distribution and its population parameter that the maximum value follows are estimated from long-term observation data. Additionally, the maximum value or a large value that occurs once every century can be pre-

dicted based on the estimated results.

Extreme value techniques are also becoming widely used for portfolio adjustment in the insurance industry, risk assessment on financial markets, and traffic prediction in telecommunications [1]. Statistical approaches focused on extreme values have shown promising results in forecasting unusual events in earth sciences, genetics, and finance. For instance, EVT was developed in the 1920s and has been used to predict the occurrence of events such as droughts and flooding [2] or financial risk [3] [4]. The application of extreme value modeling has been performed in the fields of ocean wave modeling [5], biomedical data processing [6], thermodynamics of earthquakes [7] [8], climatology [9], food science [10], and public health [11].

In this paper, we focus on one of the most popular events in athletics: the 100 m dash for both men and women. The 100 m world record has had evolution. Usain Bolt (Jamaica) is the 1st of the men: 9.58 s (August 16, 2009). Florence Griffith-Joyner (United States) is the 1st of the women: 10.49 s (July 16, 1988). The ultimate records in athletics were calculated using extreme value theory [12] [13]. This study predicts 100 m records in the world and Japan using the extreme value theory.

2. Data and Method of Analysis

2.1. Data

We used 100 m records for men and women in the world and Japan [14].

2.2. Method

EVT concerns with phenomena of extreme data. The method of block maxima was used. A method for modeling the extremes of a stationary time series is block maxima, in which consecutive observations are grouped into non-overlapping blocks of length n , generating a series of m block maxima, $M_{n,1}, \dots, M_{n,m}$, to which the Generalized Extreme Value (GEV) distribution can be fitted for some large value of n . The usual approach is to consider blocks of a given time length, thus yielding maxima at regular intervals [1]. Here a block was considered as a year, *i.e.*, annual maxima values were used. Although the method of block maxima is suitable for the analysis of maximum value data, it has the disadvantage of being easily affected by one realization value and having a large variance of the estimator.

When data are taken to be the maxima (or minima) over certain blocks of time (such as annual maximum precipitation), it is appropriate to use the GEV distribution:

$$G(z) = \begin{cases} \exp \left\{ - \left[1 + \xi \left(\frac{z - \mu}{\sigma} \right) \right]^{-1/\xi} \right\}, & \xi \neq 0, \\ \exp \left\{ - \exp \left[- \left(\frac{z - \mu}{\sigma} \right) \right] \right\}, & \xi = 0, \end{cases} \quad (1)$$

where z is extreme values from blocks, μ a location parameter, σ a scale parameter, and ξ a shape parameter. $G(z)$ is defined for all z such that $(1 + \xi(z - \mu)/\sigma) > 0$ for $\xi \neq 0$, and all z for $\xi = 0$. Three families of GEV distributions are defined depending on the value of ξ . For $\xi > 0$, we get the Fréchet distribution with a heavy tail, for $\xi = 0$, the Gumbel distribution with a lighter tail, and for $\xi < 0$, the Weibull distribution with a finite tail.

We want to know how small the value is as the fastest speed; hence it is necessary to multiply the data by -1 to put it in the framework of extreme value statistics that considers the maximum.

If a GEV distribution is fitted to observations, it becomes possible to estimate the probability of an event that has not yet been observed. Estimates of extreme quantiles of the annual maximum distribution are obtained by inverting Equation (1):

$$z_p = \begin{cases} \mu - \frac{\sigma}{\xi} \left[1 - \{-\log(1-p)\}^{-\xi} \right], & \xi \neq 0, \\ \mu - \sigma \log[-\log(1-p)], & \xi = 0, \end{cases} \tag{2}$$

where $G(z_p) = 1 - p$. z_p is the return level associated with the return period $1/p$, since to a reasonable degree of accuracy, z_p is expected to be exceeded on average once every $1/p$ year. More accurately, z_p is exceeded by the annual maximum in any particular year with probability p [1].

Modeling was performed using the evd package in R for the GEV calculations. We also tried a non-stationary model in the GEV, but it did not work.

3. Results

The 100 m records for men in the world and Japan for 1970-2009 are shown in **Figure 1**. In the world the changes were small for 1975-2000 and records decreased for 2005-2009. In Japan, the records decreased for 1980-1998. The two data were getting closer and closer. **Figure 1** shows the 100 m records of women in the world and Japan for 1970-2003. The change in the world was small and that in Japan was large. The closeness of the two data was small and the difference was still large.

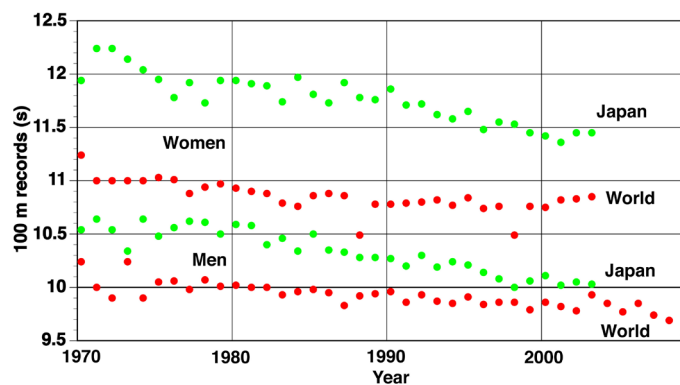


Figure 1. Plot of the 100 m records for men and women.

3.1. Men

Table 1 shows the GEV parameter estimates, which were the results of the GEV modeling on the 100 m records for men in the world using the block maxima method. The GEV parameters were estimated using the maximum likelihood estimation (MLE).

The model has three parameters: location parameter, μ ; scale parameter, σ ; and shape parameter, ξ . Because ξ was negative, the 100 m records in the world had a finite upper limit.

Table 2 shows the predicted maximum return levels for the return periods of 10, 20, 50, 100, and 350 years along with their respective 95% confidence intervals. For the 10-year return period, we estimated the return level to be 9.74 s, with a 95% confidence interval of [9.69, 9.79]. For the 100-year return period, we estimated the return level to be 9.62 s, with a 95% confidence interval of [9.54, 9.69]. Another way to interpret the plot is to say that there is approximately a 1% chance (1/100) each year that the 100 m record will not exceed 9.62 s. There is approximately a 10% chance (1/10) each year that the 100 m records will not exceed 9.74 s.

Various diagnostic plots for assessing the accuracy of the GEV model fitted to the 100 m records of men in the world are shown in **Figure 2**. Straight lines and curves are estimated functions. Each point plot is a realization value. The lines on both sides represent the 95% confidence interval. Neither the probability plot nor the quantile plot gives cause to doubt the validity of the fitted model: each set of plotted points is near-linear. The corresponding density estimate seems consistent with the data. Following the negative value of ξ , the tails are finite, and the return level curve is nonlinear. Various diagnostic plots gave little reason to doubt the validity of the GEV model.

Table 3 shows the GEV parameter estimates, which were the results of the GEV modeling on the 100 m records of men in Japan using the block maxima method. Because ξ was negative, the 100 m records in the world had a finite upper limit. **Table 4** shows the predicted maximum return levels. Various diagnostic plots for assessing the accuracy of the GEV model fitted to 100 m records for

Table 1. GEV parameter estimates of 100 m records for men in the world.

	μ	σ	ξ
Parameter estimate	9.95	0.120	-0.250
Standard errors	0.0208	0.0137	0.0718
95% CI	[9.90, 9.99]	[0.0934, 0.147]	[-0.391, -0.109]

Table 2. GEV return level estimates for men in the world.

Return period (year)	10	20	50	100	350
Return level (s)	9.74	9.69	9.65	9.62	9.58
Standard errors	0.0255	0.0286	0.00342	0.0393	0.0496
95% CI	[9.69, 9.79]	[9.64, 9.75]	[9.58, 9.71]	[9.54, 9.69]	[9.48, 9.67]

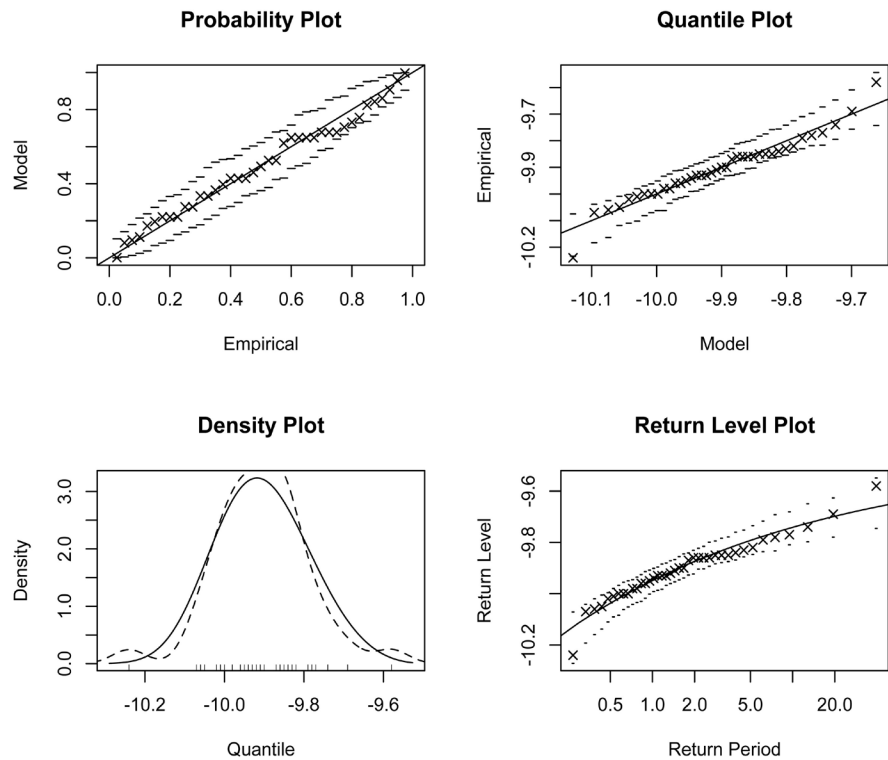


Figure 2. Diagnostic plots for GEV fit to 100 m records for men in the world.

Table 3. GEV parameter estimates of 100 m records for men in Japan.

	μ	σ	ξ
Parameter estimate	10.4	0.207	-0.393
Standard errors	0.0438	0.0367	0.225
95% CI	[10.3, 10.5]	[0.135, 0.279]	[-0.833, 0.0479]

Table 4. GEV return level estimates for men in Japan.

Return period (year)	10	20	50	100	500
Return level (s)	10.09	10.03	9.98	9.96	9.92
Standard errors	0.0361	0.0458	0.0698	0.0895	0.130
95% CI	[10.02, 10.16]	[9.94, 10.12]	[9.85, 10.12]	[9.78, 10.13]	[9.66, 10.17]

men in Japan are shown in **Figure 3**. Various diagnostic plots gave the validity of the GEV model.

3.2. Women

Table 5 shows the GEV parameter estimates, which were the results of the GEV modeling on the 100 m records of women in the world using the method of block maxima. The value of ξ was close to zero (-0.0497) and included zero as a

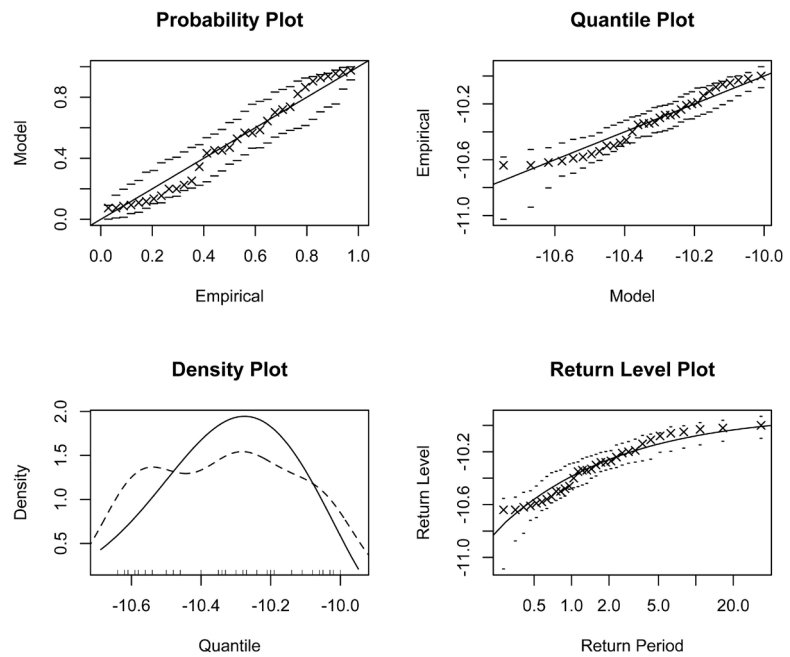


Figure 3. Diagnostic plots for GEV fit to 100 m records for men in Japan.

Table 5. GEV parameter estimates of 100 m records for women in the world.

	μ	σ	ξ
Parameter estimate	10.9	0.104	-0.0497
Standard errors	0.0202	0.0144	0.121
95% CI	[10.86, 10.94]	[0.0753, 0.132]	[-0.286, 0.187]

confidence interval. Therefore, ξ can be regarded as zero. When ξ is zero, there is no upper limit, but the probability of taking a large value is small. The 100 m records in the world did not have a finite upper limit, but the probability of taking a small 100 m record was small. **Table 6** shows the predicted maximum return levels for the return periods of 10, 20, 50, 100, and 500 years along with their respective 95% confidence intervals. Various diagnostic plots for assessing the accuracy of the GEV model fitted to the 100 m records for women in the world are shown in **Figure 4**. Part of the fluctuation in the upper data was not captured. In the return level curve, the estimated curve was close to linear because ξ was close to zero. Various diagnostic plots gave the validity of the GEV model.

Table 7 shows the GEV parameter estimates, which were the results of the GEV modeling on the 100 m records of women in Japan using the method of block maxima. Because ξ was negative, the 100 m records in the world had a finite upper limit. **Table 8** shows the predicted maximum return levels. The various diagnostic plots for assessing the accuracy of the GEV model fitted to the 100 m records for women in Japan are shown in **Figure 5**. Various diagnostic plots gave the validity of the GEV model.

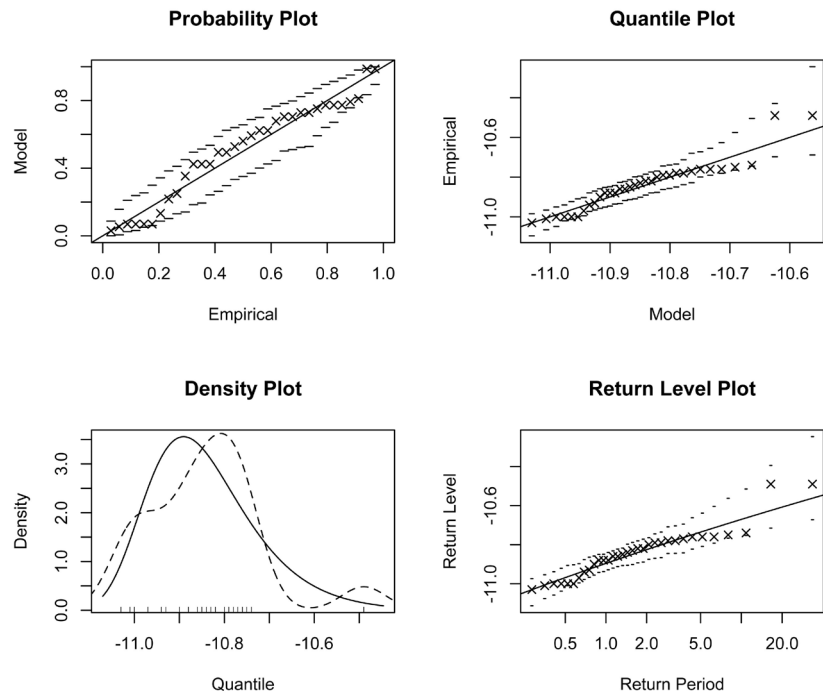


Figure 4. Diagnostic plots for GEV fit to 100 m records for women in the world.

Table 6. GEV return level estimates for women in the world.

Return period (year)	10	20	50	100	500
Return level (s)	10.68	10.61	10.53	10.47	10.34
Standard errors	0.0396	0.0540	0.0806	0.106	0.182
95% CI	[10.60, 10.75]	[10.50, 10.71]	[10.37, 10.69]	[10.26, 10.68]	[9.98, 10.70]

Table 7. GEV parameter estimates of 100 m records for women in Japan.

	μ	σ	ξ
Parameter estimate	11.8	0.243	-0.442
Standard errors	0.0476	0.371	0.150
95% CI	[11.7, 11.9]	[0.171, 0.316]	[-0.735, -0.148]

Table 8. GEV return level estimates for women in Japan.

Return period (year)	10	20	50	100	500
Return level (s)	11.48	11.43	11.38	11.35	11.31
Standard errors	0.0375	0.0398	0.0503	0.0603	0.0814
95% CI	[11.41, 11.55]	[11.35, 11.50]	[11.28, 11.47]	[11.23, 11.47]	[11.23, 11.47]

4. Discussion

The return level plot for the 100 m records for men is shown in **Figure 6**. In the

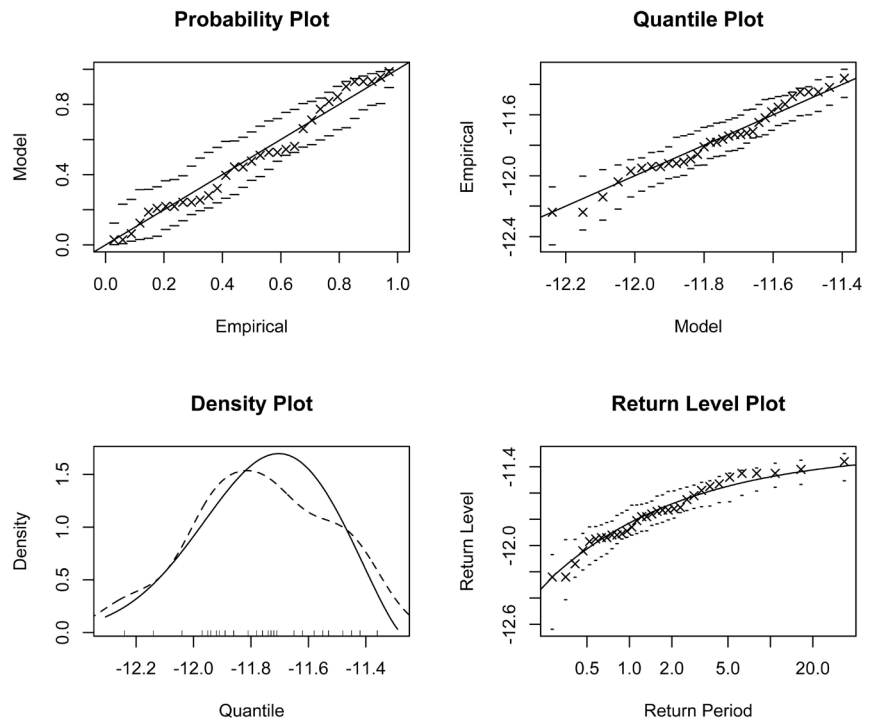


Figure 5. Diagnostic plots for GEV fit to 100 m records for women in Japan.

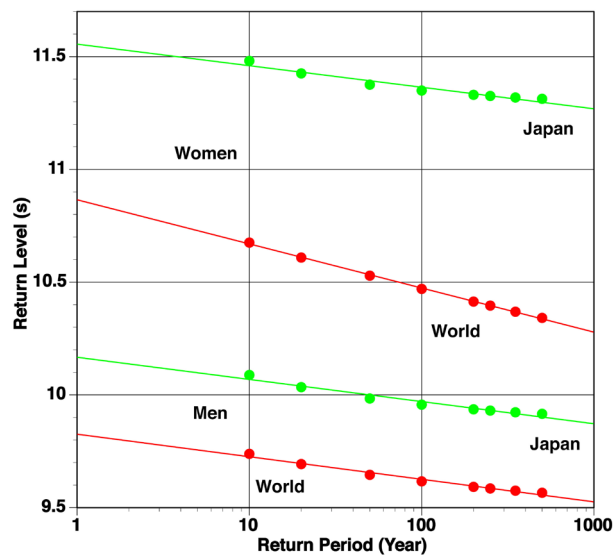


Figure 6. Return level plot for the 100 m records of men and women in the world and Japan.

$\xi < 0$ case, for men in the world and Japan, the plots deviated from the straight line and were convex downward. The calculated upper limit was 9.46 and 9.91 s in the world and Japan, respectively. For the 350-year return period, the return level for men in the world was obtained as 9.58 s, with a 95% confidence interval of [9.48, 9.67]. Hence, the probability of occurrence in one year for the record of Usain Bolt, 9.58 s, which was the 1st record in the world was 1/350. Einmahl estimated the ultimate world record and found 9.51 s for men [13].

The return level plot for the 100 m records for women in the world and Japan is shown in **Figure 6**. The slope of the approximate straight line for women in the world was the largest. That is, the fluctuation range by year was large. The $\xi = 0$ case had a heavy-tailed distribution. In the $\xi < 0$ case, the plots for the Japanese women's records deviated from the straight line with a downward convex shape. The calculated upper limit was 11.3 s in Japan. Compared with the Japanese women's record, the women's world record showed a larger change and improved with the return period. For the 100-year return period, the return level for women in the world was obtained as 10.47 s, with a 95% confidence interval of [10.26, 10.68]. The probability of occurrence in one year for the record of Florence Griffith-Joyner, 10.49 s, which was the 1st record for women in the world was about 1/100 from **Table 6**. Einmahl estimated the ultimate world record and found 10.33 s for women [13].

The probability of occurrence in one year for a new world record of men, 9.58 s (Usain Bolt), was 1/350, while that for women, 10.49 s (Florence Griffith-Joyner), was about 1/100, confirming it was more difficult for men to break records than women.

To break a record, a runner should accelerate quickly and maintain the maximum speed. Mechanically, acceleration, a , is determined by the angle θ at which the body is tilted with respect to the ground and can be expressed as

$$a = g/\tan \theta,$$

where g is the gravitational acceleration. To increase the acceleration, the body tilt should be increased and θ should be reduced.

5. Conclusions

Extreme value theory provides methods to analyze the most extreme parts of data. Here, we used the GEV distribution to predict the ultimate 100 m dash records for men and women for specific periods. The results are summarized as follows:

1) The various diagnostic plots, which assessed the accuracy of the GEV model, were well fitted to the 100 m records in the world and Japan, validating the model.

2) The men's world record had a shape parameter of -0.250 with a 95% confidence interval of $[-0.391, -0.109]$. The 100 m record had a finite limit and a calculated upper limit was 9.46 s. The calculated upper limit of the gold medalist for men was 9.58 s, which is equal to the record of Usain Bolt.

3) The return level estimate for men in the world was obtained as 9.74, 9.62, and 9.58 s, with a 95% confidence interval of [9.69, 9.79], [9.54, 9.69] and [9.48, 9.67] for 10-, 100- and 350-year return periods, respectively. The probability of occurrence in one year for a new world record of men, 9.58 s (Usain Bolt), was 1/350, while that for women, 10.49 s (Florence Griffith-Joyner), was about 1/100, confirming it was more difficult for men to break records than women.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

- [1] Coles, S. (2001) An Introduction to Statistical Modeling of Extreme Values. Springer-Verlag. <https://doi.org/10.1007/978-1-4471-3675-0>
- [2] Katz, R.W., Parlange, M.B. and Naveau, P. (2002) Statistics of Extremes in Hydrology. *Advances in Water Resources*, **25**, 1287-1304. [https://doi.org/10.1016/S0309-1708\(02\)00056-8](https://doi.org/10.1016/S0309-1708(02)00056-8)
- [3] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997) Modeling Extremal Events for Insurance and Finance. Springer-Verlag. <https://doi.org/10.1007/978-3-642-33483-2>
- [4] Gilli, M. and kéllezi, E. (2006) An Application of Extreme Value Theory for Measuring Financial Risk. *Computational Economics*, **27**, 207-228. <https://doi.org/10.1007/s10614-006-9025-7>
- [5] Dawson, T.H. (2000) Maximum Wave Crests in Heavy Seas. *Journal of Offshore Mechanics and Arctic Engineering-Transactions of the AMSE*, **122**, 222-224. <https://doi.org/10.1115/1.1287039>
- [6] Roberts, S.J. (2000) Extreme Value Statistics for Novelty Detection in Biomedical Data Processing. *IEE Proceedings—Science Measurement and Technology*, **147**, 363-367. <https://doi.org/10.1049/ip-smt:20000841>
- [7] Lavenda, B.H. and Cipollone, E. (2000) Extreme Value Statistics and Thermodynamics of Earthquakes: Aftershock Sequences. *Annali di Geofisica*, **43**, 967-982.
- [8] Maruyama, F. (2020) Analyzing the Annual Maximum Magnitude of Earthquakes in Japan by Extreme Value Theory. *Open Journal of Applied Sciences*, **10**, 817-824. <https://doi.org/10.4236/ojapps.2020.1012057>
- [9] Brown, S.J. (2018) The Drivers of Variability in UK Extreme Rainfall. *International Journal of Climatology*, **38**, e119-e130. <https://doi.org/10.1002/joc.5356>
- [10] Kawas, M.L. and Moreira, R.G. (2001) Characterization of Product Quality Attributes of Tortilla Chips during the Frying Process. *Journal of Food Engineering*, **47**, 97-107. [https://doi.org/10.1016/S0260-8774\(00\)00104-7](https://doi.org/10.1016/S0260-8774(00)00104-7)
- [11] Thomas, M., Lemaitre, M., Wilson, M. L., Vibound, C., Yordanov, Y., Wackernagel, H. and Carrat, F. (2016) Applications of Extreme Value Theory in Public Health. *PLoS ONE*, **11**, e0159312. <https://doi.org/10.1371/journal.pone.0159312>
- [12] Einmahl, J.H.J. and Magnus, J.R. (2008) Records in Athletics through Extreme-Value Theory. *Journal of the American Statistical Association*, **103**, 1382-1391. <https://doi.org/10.1198/016214508000000698>
- [13] Einmahl, J.H.J. and Smeets, S.G.W.R. (2011) Ultimate 100-m World Records through Extreme-Value Theory. *Statistical Neerlandica*, **65**, 32-42. <https://doi.org/10.1111/j.1467-9574.2010.00470.x>
- [14] Ito, H. and Okano, S. (2005) Analysis of Changes in the 100 m Records in Japan and the World. *Bulletin of Athletics Research*, **1**, 61-66.