ISSN 1330-3651 (Print), ISSN 1848-6339 (Online) https://doi.org/10.17559/TV-20180529125449
Original scientific paper

A New Methodology for the Block Maxima Approach in Selecting the Optimal Block Size

Çiğdem ÖZARI, Özge EREN, Hasan SAYGIN

Abstract: The Block Maxima method divides sample data into equal blocks. Predictions are based on the maximum values of the observations. Choosing an efficient and proper block size for the Block Maxima method is an important issue and varies across fields (e.g., flood, rainfall, finance). However, the main problem is deciding which block size is suitable or optimal for the prediction. In the literature, it is a known fact that the selection of a small block size leads to bias, while the selection of a large block size leads to a variance problem. In one respect, this issue is any trade off problem between the bias and the variance. This paper proposes simple and easy computational method to specify the optimal block size selection process for the Block Maxima method.

Keywords: Block Maxima; Extreme Value Theory; Maximum Likelihood

1 INTRODUCTION

Predicting the probability of extreme and rare events is important for making future inferences. The extreme value theory (EVT) is a robust technique used to analyse the tail behaviour of distributions. Fisher and Tippett developed EVT [1]. It was later formalized by Gnedenko [2]. After the theoretical developments between 1930s and 1940s, lots of papers related to the applications of EVT have been used with different scientific fields (e.g., engineering, finance (McNeil, 1999) [3], environment (Smith, 1989; Stephenson et al., 2005) [4], [5]).

With EVT, let X_1, X_2, X_3,…, X_n be identically and independently distributed random variables. The main theory of the extreme data is about the limit behaviour of the max$\{X_1, X_2, X_3,\ldots, X_n\}$ or min $\{X_1, X_2, X_3,\ldots, X_n\}$ as $n \to \infty$ [6]

EVT deals with the stochastic behaviour of maximum and minimum of identically independent, random variables. The distributional properties of extremes (maximum and minimum) and exceedances of over or below threshold are specified underlying distribution (Kotz 2010 [7]).

1.1 EVT Approaches

There are two principal models for extreme values: The Peaks over Threshold (POT) method and the Block Maxima (BM) method. POT focuses on the realisations exceeding a given (high) threshold u. With POT, Balkema and de Haan (1974) [6] and Pickands (1975) [8] state that, for a large enough u, it is well approximated by the Generalised Pareto Distribution (GPD), as shown in Eq. (1).

$$GPD_{\xi,\,\mu,\,\sigma}F(x)=\begin{cases}1-\left(1+\xi\dfrac{x-\mu}{\sigma}\right)^{-\frac{1}{\xi}} & \text{if } \xi\neq 0\\[2ex] 1-\exp\left\{-\dfrac{x-\mu}{\sigma}\right\} & \text{if } \xi=0\end{cases} \tag{1}$$

where μ is the location, σ is the scale, and ξ is the shape of the parameter.

The BM method is widely suitable for applying the Generalized Extreme Value (GEV) distribution. The GEV distribution unites the Gumbel, Fréchet and Weibull distributions into a single family to allow for a continuous range of possible shapes. A single three-parameter model GEV distribution has the following cumulative distribution function:

$$\begin{aligned}F(x)&=\exp\left\{-\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}\right\} & \xi\neq 0\\ F(x)&=\exp\left\{-\left[\exp\left(\frac{x-\mu}{\sigma}\right)\right]\right\} & \xi=0\end{aligned} \tag{2}$$

The GEV distribution encompasses three limiting distributions of extreme value, depending on the value of the parameter shape:

If $\xi > 0$, it suits the Fréchet distribution.
If $\xi < 0$, it suits the Weibull distribution.
If $\xi = 0$, it suits the Gumbel distribution.

Researchers disagree as to which technique is more efficient. According to Ferreira and de Haan [9] the BM method is more efficient than the POT method under the usual practical conditions. On the other hand, the POT method allows for greater flexibility in many cases since it might be difficult to change the block size in practice [9]. POT method is applicable for selecting loads above a threshold. However, the crucial problem becomes how to determine the threshold properly. Although there are many studies related to the threshold selection, a stable and effective method has not yet been established [10]. In a previous work [11], the BM and the POT approaches have been applied to pitting corrosion data from laboratory-simulated buried line pipe steel and they conclude that both approaches have been useful for gaining a better understanding of pitting in low carbon steel.

Bekiros et al compares the predictive ability of value at risk estimates obtained from various estimation techniques such as J. P. Morgan's Riskmetrics, Moving Average, Generalized Autoregressive Conditional Heteroskedasticity and EVT including POT and the BM [12]. For the block size, they analyzed monthly and quarterly minima. They found that EVT models are more suitable for long-run forecasts of the maximum potential losses rather than being a day-to-day tool to measure the market risk.

Generally, on the BM method studies, it can be seen specific time selection like one month, six months, a year (Engeland, 2004 [13]) or used as arbitrary parameter (Santinelli, 2014 [14]. Selecting a proper and optimum block size for the BM is a crucial issue. Varying across field studies different block sizes were used without any explanation (e.g., flood (Mudersbach, 2010) [15], rainfall (Villarini, 2011) [16]). Bystörm also indicated that this can be named as optimal block size problem [17]. Singh et all gives an example for how to convert the rain fall data set to use the BM. They suggest application by dividing the datasets into yearly, semester, quarterly or monthly blocks without other assumptions [18]. The main intent of Cooley's study is to show weather temperature prediction by using EVT on the area of Central England with the data set starting from 1878 up to 2007 and they used the BM approach by getting daily maximum temperatures [19].

The fitting will be inaccurate if block size is too small which may lead to a biased estimation [10], whereas one that is too large may result in a few extracted extreme values, and subsequently, a large variance. Thus, to fully extract the extreme values to constitute the sample loads for fitting a GEV, the block size must be exact. If it is unreasonable, the predictions will be inaccurate.

To sum up; when we examined the literature which uses the BM methodology, we could not reach a method that allows researchers to select the proper or stable block size. In many studies (also in the same area), a lot of researchers use different block size without any assumptions. Because of this, choosing an accurate and proper block size for the BM method is an important issue to make good prediction. The main aim of this study is to propose a simple and easy computational method to specify the optimal block size selection process for the BM method and to pay attention to this important issue.

2 METHODOLOGY AND FINDINGS

The proposed methodology includes seven steps to obtain the optimum block size. To explain the methodology which we propose in this study, we generate a randomly continuous dataset between 0-1, labelled as actual data and the last 10% part of actual data is reserved for testing and is labelled as test data.

A continuous random variable is generated with 2200 observations. The last 10% of the generated dataset (with 220 observations) is reserved for the testing part of the analysis.

Step 1: Create data blocks from the actual dataset with block size k, $k \in \{10, 11, \ldots, 50\}$.

The actual dataset is divided into blocks with different block sizes, between 10 and 50 (Considering the bias and variance problems).

Step 2: For any block size k, first calculate the maximum value of each block, then combining these maximum values construct the k-dataset. In this step, the k-dataset was constructed, where $k = 10, 11, \ldots, 50$.

For instance, if $k = 10$, 200 block sets are constructed, and the maximum values are obtained from each block set. These maximum values get together and are called the 10-dataset. Fig. 1 illustrates the graphs for the various block sizes (i.e., 10, 20, 30 and 40).

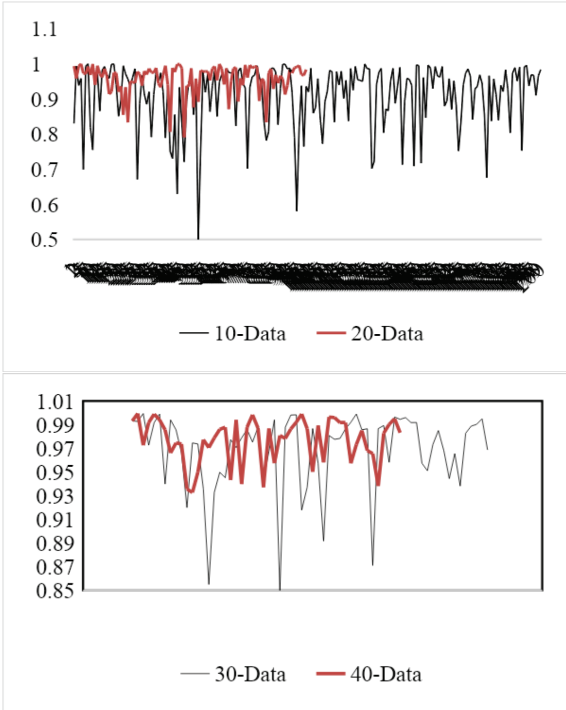

Figure 1 Maximum Values for the Various Block Sizes

Step 3: Check whether the k-dataset fits to the GEV distribution or not and calculate the parameters of the k-dataset (if they are GEV distributed, otherwise, determine the best distribution and calculate its parameter).

The Anderson Darling and Kolmogorov Smirnov tests were used to determine the extent to which the extreme dispersion of all blocks is appropriate.

Although there are many parameter estimation techniques available, in this study, the Maximum Likelihood estimation technique has been used. It is seen to be the most appropriate (Gaines and Denny [20], 1993; Leder, 1998 [21]) technique. Fig. 2 illustrates the parameters of the GEV distribution for all block sizes (i.e., 10 through 50).

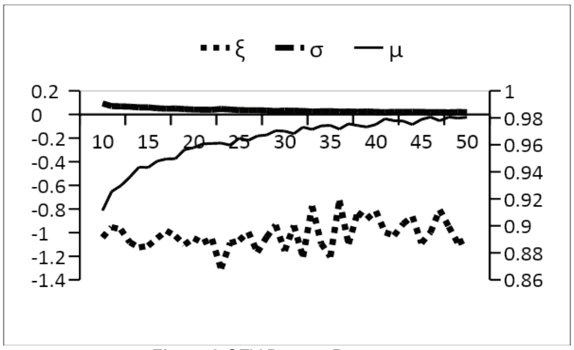

Figure 2 GEV Dataset Parameters

As shown in Fig. 2, there is no trend in the shape of the parameter indicated by ξ. The location parameter has an increasing trend, while the scale parameter σ has a decreasing trend. This decreasing trend conflicts with the literature which says variance problem can occur if the block size is too large. Tab. 1 illustrates the parameters of the GEV distribution for each block size.

Table 1 GEV Distribution Parameters

Block Size	ξ	σ	μ
10	-1.041	0.093	0.911
11	-0.954	0.071	0.925
12	-0.973	0.068	0.930
13	-1.083	0.066	0.936
14	-1.126	0.060	0.943
15	-1.113	0.060	0.943
16	-1.048	0.052	0.948
17	-0.984	0.049	0.950
18	-1.028	0.050	0.950
19	-1.103	0.045	0.956
20	-1.047	0.042	0.958
21	-1.086	0.040	0.961
22	-1.054	0.040	0.961
23	-1.305	0.046	0.961
24	-1.088	0.042	0.960
25	-1.066	0.036	0.965
26	-0.996	0.036	0.963
27	-1.167	0.036	0.967
28	-1.039	0.033	0.967
29	-0.936	0.028	0.970
10	-1.041	0.093	0.911
30	-1.155	0.033	0.970
31	-0.941	0.030	0.968
32	-1.206	0.030	0.973
33	-0.775	0.024	0.971
34	-1.093	0.027	0.974
35	-1.205	0.028	0.974
36	-0.728	0.023	0.972
37	-1.096	0.025	0.976
38	-0.829	0.022	0.974
39	-0.886	0.025	0.973
40	-0.817	0.021	0.975
41	-0.998	0.019	0.979
42	-1.025	0.022	0.978
43	-0.925	0.021	0.978
44	-0.863	0.023	0.975
45	-1.089	0.022	0.979
46	-0.998	0.019	0.981
47	-0.807	0.020	0.978
48	-0.950	0.018	0.981
49	-1.089	0.021	0.980
50	-1.155	0.033	0.970

Step 4: With the parameters calculated in Step 3, for any block size k generate new random variables which are GEV distributed with these parameters. Label this dataset as the k-predicted dataset.

Fig. 3 illustrates the graphs of test data and some predicted data sets which are constructed from different block size.

Step 5: Define term extreme.

Since the BM method is a method used to predict extreme values, we needed to define the term extreme. In other words, we expect to predict extreme data values. Using basic statistical methods and expert opinions, we decided to determine the extreme value by calculating two standard deviations away from the mean. In our study, this value is equal to 0.94.

Step 6: For each k, check the similarity between test data and k-predicted data set for extreme values.

To check the similarity, first we eliminate the data that was smaller than 0.94 as explained in Step 5.

Since the study begins to deal with extreme values only, the number of observations of the predicted data becomes less. To compare two independent groups similarity, we decided to use non-parametric Mann-Whitney U test. In other words, by using Mann-Whitney U test with 95% confidence interval level, the means of the test data set, and the predicted data set were checked if they come from similar population or not. The null hypothesis for this test is that the difference of location between the samples is equal to 0 and alternative hypothesis is not equal. Tab. 2 illustrates the findings of this test.

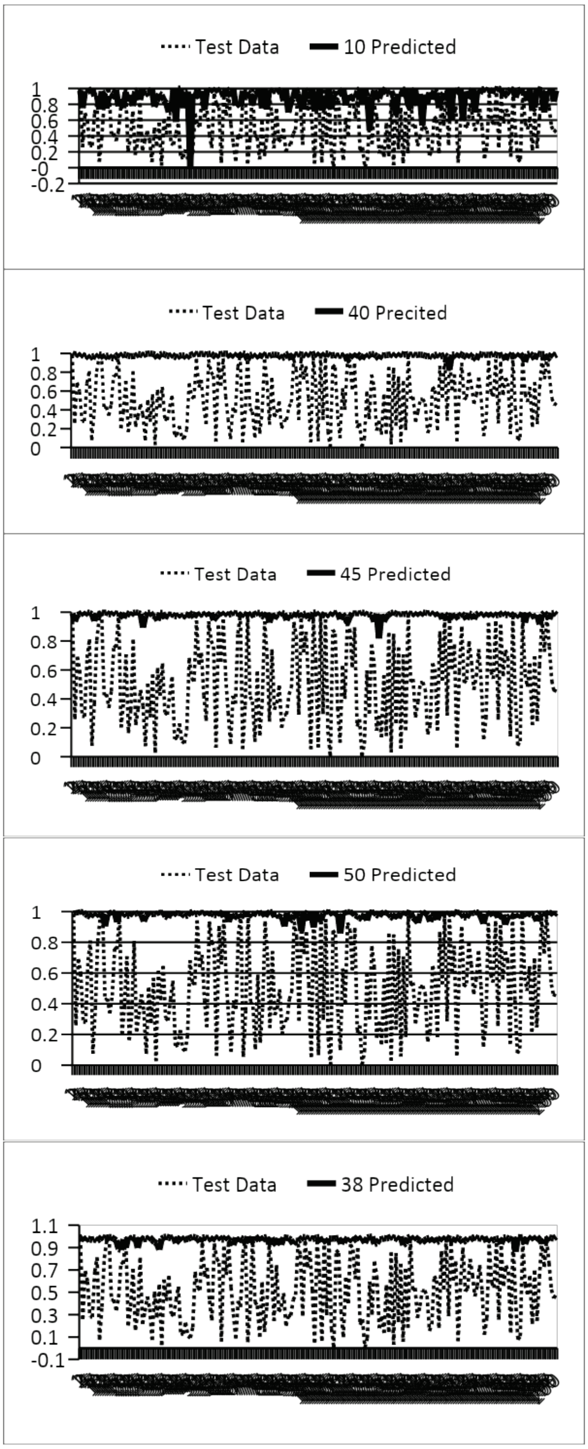

Figure 3 Test and Predicted Data Sets

From Tab. 2 we can conclude that for some of the block sizes, the predicted data and the test data groups are not similar.

Step 6: Eliminate non-similar block size.

In this step, we eliminate the predicted blocks for which Mann-Whitney test says that they are not similar. Tab. 3 illustrates only similar predicted blocks.

Table 2 Significance Values of Mann-Whitney U test

Block Size	p-value	Result	Block Size	p-value	Result
10	0.011	Reject	31	0.009	Reject
11	0.001	Reject	32	0.267	Accept
12	0.020	Reject	33	0.255	Accept
13	0.062	Accept	34	0.683	Accept
14	0.036	Reject	35	0.199	Accept
15	0.397	Accept	36	0.930	Accept
16	0.267	Accept	37	0.884	Accept
17	0.006	Reject	38	0.122	Accept
18	0.220	Accept	39	0.726	Accept
19	0.307	Accept	40	0.243	Accept
20	0.335	Accept	41	0.096	Accept
21	0.280	Accept	42	0.153	Accept
22	0.953	Accept	43	0.267	Accept
23	0.521	Accept	44	0.280	Accept
24	0.884	Accept	45	0.012	Reject
25	0.414	Accept	46	0.096	Accept
26	1.000	Accept	47	0.029	Reject
27	0.953	Accept	48	0.007	Reject
28	0.930	Accept	49	0.179	Accept
29	0.620	Accept	50	0.041	Reject
30	0.006	Reject			

Table 3 Similar Block Size with p-values

Block Size	p-value	Block Size	p-value
13	0.062	32	0.267
15	0.397	33	0.255
16	0.267	34	0.683
18	0.220	35	0.199
19	0.307	36	0.930
20	0.335	37	0.884
21	0.280	38	0.122
22	0.953	39	0.726
23	0.521	40	0.243
24	0.884	41	0.096
25	0.414	42	0.153
26	1.000	43	0.267
27	0.953	44	0.280
28	0.930	46	0.096
29	0.620	49	0.179

Table 4 Predicted and Controlled Dataset Example

Test Data	13 Pred.	13 Control	38 Pred.	38 Control	40 Pred.	40 Control
0.9735	0.9766	0.0031	0.9725	0.0010	0.9821	0.0086
0.9839	0.9804	0.0035	0.9735	0.0105	0.9899	0.0060
0.9697	0.9903	0.0206	0.9763	0.0067	0.9690	0.0006
0.9479	0.9340	0.0139	0.9977	0.0498	0.9989	0.0509
0.9601	0.8364	0.1237	0.9527	0.0074	0.9298	0.0303
0.9722	0.8171	0.1551	0.9958	0.0236	0.9938	0.0215
0.9581	0.9668	0.0088	0.9704	0.0123	0.9360	0.0220
0.9418	0.9562	0.0144	0.9883	0.0465	1.0004	0.0586
0.9499	0.9589	0.0090	0.9257	0.0242	0.9854	0.0355
0.9924	0.9699	0.0225	0.9770	0.0154	0.9642	0.0282
0.9937	0.9699	0.0238	0.9824	0.0113	0.9532	0.0405
0.9839	0.8799	0.1040	0.9943	0.0104	0.9950	0.0111
0.9745	0.9727	0.0018	0.9806	0.0061	0.9747	0.0002
0.9600	0.9621	0.0021	0.9649	0.0049	0.9988	0.0388
0.9791	0.8272	0.1520	0.9888	0.0097	0.9323	0.0468
0.9991	0.9577	0.0413	0.9952	0.0039	0.9857	0.0134
0.9583	0.8568	0.1015	0.9924	0.0341	0.9872	0.0288
0.9946	0.8804	0.1142	0.9616	0.0330	0.9962	0.0016
0.9491	0.9963	0.0472	0.9961	0.0471	0.9631	0.0140

Step 7: Select the appropriate block size

The appropriate block size has the highest relationship between the predicted data and the test data (that we reserved at the beginning from the actual data set). There are many methods available to check the relationship between them. In this study, we used the absolute value difference. In other words; for any block size k, we take the difference of test data and predicted data sets and then take the absolute value of these differences for only extreme values and these data sets are labelled as the k^{th} control data set. Tab. 4 illustrates the test data, predicted data and the control data with block size 13, 38 and 40.

Tab. 5 illustrates the block size with its control blocks' total sum of absolute differences. Block size 38 has the smallest difference, with a value of 0.3575.

Table 5 Block Size and Total Absolute Difference

Block Size	Total Absolute Difference	Block Size	Total Absolute Difference
13	0.9627	32	0.4529
15	0.4831	33	0.4090
16	0.6620	34	0.5193
18	0.7156	35	0.3974
19	0.7318	36	0.5971
20	0.4454	37	0.4961
21	0.4718	38	0.3575
22	0.5912	39	0.4810
23	0.8201	40	0.4575
24	0.7392	41	0.4210
25	0.4615	42	0.4534
26	0.7567	43	0.3861
27	0.6071	44	0.3707
28	0.4571	46	0.4291
29	0.4606	49	0.4588

Fig. 4 illustrates the relationship between block size and the total difference between test and predicted data sets.

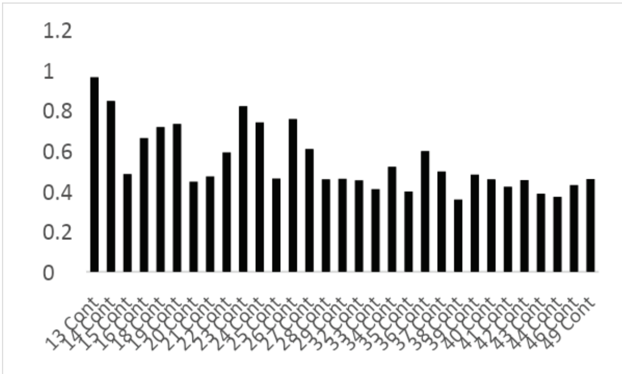

Figure 4 Block Size & Total Difference

Block size 38 has the smallest difference with a value of 0.3575 and also Mann-Whitney U test says that the two groups are similar. The block size with the slightest difference and with the taking acceptance from Mann-Whitney U test is the width of the block; hence, we use this to make a good estimate or prediction for the dataset.

3 CONCLUSIONS

Many researchers have designed methodologies related to the block size dilemma in relation to the BM method. The main purpose of this study is to recommend a method that selects an optimal or most appropriate block size to make better estimation.

The selection of a suitable block size is a critical issue. In most of the studies related to this topic, we could not reach any reasons or assumptions that explain why they chose the block size they used in their studies. How can we know that another block size can estimate better findings than all the other block sizes? Because of this, we decided to propose a simple and useful method to explain the reason behind choosing the appropriate block size.

In the proposed method, after creating the tested data, a predicted data set was constructed. First by the help of Mann-Whitney U test we check the similarity between test data and all predicted data sets for only extreme values. The strength of the relationship between the testing data and predicted data (for each block size) was then assessed. 38 predicted data values were found to be more precise than the other block sizes. To check which block size is more precise, absolute difference technique has been used. There are many methods available to measure this relation (e.g., Root sum square method, Euclidean distance or correlation coefficient). In the future, it is recommended that the same study be conducted for different parameter estimation techniques.

4 REFERENCES

[1] Fisher, R. A. & Tippett, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Mathematical Proceedings of the Cambridge Philosophical Society, 24*(2). https://doi.org/10.1017/S0305004100015681

[2] Gnedenko, B. (1943). Sur la distribution limite du terme maximum d'une serie aleatoire. *Annals of Mathematics, 44*(3), 423-453. https://doi.org/10.2307/1968974

[3] McNeil, A. J. (1999). *Extreme value theory for risk managers*. Departement Mathematik ETH Zentrum.

[4] Smith, R. L. (1986). Extreme value theory based on the r largest annual events. *Journal of Hydrology*, *86*(1-2), 27-43. https://doi.org/10.1016/0022-1694(86)90004-1

[5] Stephenson, A. & Gilleland, E. (2005). Software for the analysis of extreme events: The current state and future directions. *Extremes, 8*(3), 87-109. https://doi.org/10.1007/s10687-006-7962-0

[6] Balkema, A. A. & De Haan, L. (1974). Residual life time at great age. *The Annals of Probability*, 792-804. https://doi.org/10.1214/aop/1176996548

[7] Kotz, S. & Nadarajah, S. (2000). Extreme value distributions: theory and applications. *World Scientific*. https://doi.org/10.1142/p191

[8] Pickands III, J. (1975). Statistical inference using extreme order statistics. *The Annals of Statistics*, 119-131. https://doi.org/10.1214/aos/1176343003

[9] Ferreira, A. & de Haan, L. (2015). On the block maxima method in extreme value theory: PWM estimators. *The Annals of statistics*, *43*(1), 276-298. https://doi.org/10.1214/14-AOS1280

[10] Coles, S., Bawa, J., Trenner, L., & Dorazio, P. (2001). *An Introduction to Statistical Modeling of Extreme Values* (208). London: Springer. https://doi.org/10.1007/978-1-4471-3675-0

[11] Rivas, D., Caleyo, F., Valor, A., & Hallen, J. M. (2008). Extreme value analysis applied to pitting corrosion experiments in low carbon steel: comparison of block maxima and peak over threshold approaches. *Corrosion Sci.* 3193-3204. https://doi.org/10.1016/j.corsci.2008.08.002

[12] Bekiros, S. D. & Georgoutsos, D. A. (2005). Estimation of Value-at-Risk by extreme value and conventional methods: A comparative evaluation of their predictive performance. *Journal of International Financial Markets, Institutions and Money, 15*(3), 209-228. https://doi.org/10.1016/j.intfin.2004.05.002

[13] Engeland, K., Hisdal, H., & Frigessi, A. (2004). Practical extreme value modelling of hydrological floods and droughts: a case study. *Extremes, 7*(1), 5-30. https://doi.org/10.1007/s10687-004-4727-5

[14] Santinelli, L., et al. (2014). On the sustainability of the extreme value theory for WCET estimation. *OASIcs-OpenAccess Series in Informatics*. (39). Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik,

[15] Mudersbach, C. & Jensen, J. (2010). Nonstationary extreme value analysis of annual maximum water levels for designing coastal structures on the German North Sea coastline. *Journal of Flood Risk Management*, *3*(1), 52-62. https://doi.org/10.1111/j.1753-318X.2009.01054.x

[16] Villarini, G., Smith, J. A., Serinaldi, F., & Ntelekos, A. A. (2011). Analyses of seasonal and annual maximum daily discharge records for central Europe. *Journal of Hydrology*, *399*(3), 299-312. https://doi.org/10.1016/j.jhydrol.2011.01.007

[17] Bystörm Hans N. E. (2004). Managing extreme risks in tranquil and volatile markets using conditional extreme value theory. *International Review of Financial Analysis*, 133-152. https://doi.org/10.1016/j.irfa.2004.02.003

[18] Singh, A. K., Allen, D. E., & Robert, P. J. (2013). Extreme market risk and extreme value theory. *Mathematics and computers in simulation*, 94, 310-328. https://doi.org/10.1016/j.matcom.2012.05.010

[19] Cooley, D. (2009). Extreme value analysis and the study of climate change. *Climatic change*, *97*(1-2), 77. https://doi.org/10.1007/s10584-009-9627-x

[20] Gaines, S. D. & Denny, M. W. (1993). The largest, smallest, highest, lowest, longest, and shortest: extremes in ecology. *Ecology*, *74*(6), 1677-1692. https://doi.org/10.2307/1939926

[21] Leder, N., Smirčić, A., & Vilibić, I. (1998). Extreme values of surface wave heights in the Northern Adriatic. *Geofizika*, *15*(1), 1-13.

Contact information:

Çiğdem ÖZARI,
İstanbul Aydın University, İstanbul, Turkey
cigdemozari@aydin.edu.tr

Özge EREN,
Corresponding author
İstanbul Aydın University, İstanbul, Turkey
ozgeeren@aydin.edu.tr

Hasan SAYGIN,
İstanbul Aydın University, İstanbul, Turkey
hasansaygin@aydin.edu.tr