Математички Билтен
47 (LXXIII) 1 (2023), 39–48

UDC: 519.246:519.862]:336.76(797.7)
DOI: 10.37560/matbil23471039gj

DETERMINING VALUE AT RISK USING EXTREME VALUE THEORY ON A FINANCIAL DATA SET

STEVO GJORGIEV [1] AND ANETA GACOVSKA-BARANDOVSKA [2]

Abstract. Extreme value theory has a wide range of applications. The paper considers application of extreme value theory in the area of financial flows. Our data set has been processed using two different methods, block maxima and peak over the threshold method. We compare the obtained results for the risk measures and draw conclusions on the behavior of the financial flows for different time intervals.

1. Introduction

Analyzing extreme events poses a significant challenge in financial forecasting. Conventional methods for predicting stock market indices frequently overlook these events, resulting in misleading conclusions. The Macedonian Stock Exchange has witnessed fluctuations since its establishment, notably during global financial crises and the COVID-19 pandemic. This study concentrates on the widely utilized risk metric, Value at Risk, assessing its efficiency in forecasting the Macedonian stock index amidst the COVID-19 crisis using the Extreme Value Theory.

2. The Theoretical Basis of the Extreme Value Methods

In the theory of extreme values, two main methods are primarily practiced to address differences in the movements of the phenomena considered. This theory is mainly concerned about the tails of the distributions on which the extreme values of the analyzed data belong to. There are two main mehods in this theory Block Maxima (BM) and Peak Over Threshold (POT). The theory of extreme values traces its origins back to the 18th century. One of the most crucial periods, marked by foundational research and the acquisition of fundamental results in this field, is the beginning of the 20th century. One of the most important theorems in this theory is the fundamental theorem for extreme values, which states:

2010 *Mathematics Subject Classification.* Primary:60G70, 62P20 Secondary:91B70, 91G70.
Key words and phrases. Extreme value theory, block maxima, peak over the threshold method, risk measures, Macedonian Stock Exchange, MB10 index.

Theorem 1 (Fisher, Tippett (1928), Gnedenko (1943))**.** [3] *Let* (X_n)*,* $n \in \mathbb{N}$*, be a sequence of independent identically distributed random variables. If there exist normalization constants* $c_n > 0$*,* $d_n \in \mathbb{R}$*, and a non-degenerate distribution function* H *such that the distribution function*

$$\frac{M_n - d_n}{c_n} \xrightarrow{d} H,$$

then H *belongs to one of the following three families of distribution functions:*
Frechet: $\Phi_\alpha(x) = 0$ *if* $x \leq 0$*,* $\Phi_\alpha(x) = e^{-x^{-\alpha}}$ *if* $x > 0$*,* $\alpha > 0$*.*
Weibull: $\Psi_\alpha(x) = e^{-(-x)^{-\alpha}}$ *if* $x \leq 0$*,* $\Psi_\alpha(x) = 1$ *if* $x > 0$*,* $\alpha > 0$*.*
Gumbel: $\Lambda(x) = e^{-e^{-x}}$*,* $x \in \mathbb{R}$*.*

Φ_α, Ψ_α and Λ are distribution functions of extreme values.

Let (X_n), $n \in \mathbb{N}$ be a sequence of independent identically distributed random variables, and $M_n = \max\{X_1, X_2, ..., X_n\}$ be the maximum statistics. The Generalized Extreme Value (GEV) distribution function is given by the function $H_{\xi,\mu,\psi}$ defined by

$$H_{\xi,\mu,\psi}(x) = \begin{cases} e^{-\left(1+\xi\frac{x-\mu}{\psi}\right)^{-\frac{1}{\xi}}}, & \xi \neq 0 \\ e^{-e^{-\frac{x-\mu}{\psi}}}, & \xi = 0 \end{cases},$$

such that $x \in \mathcal{D}$ and

$$\mathcal{D} = \begin{cases} \left(-\infty, -\frac{1}{\xi}\right), & \xi < 0 \\ (-\infty, +\infty), & \xi = 0 \\ \left(-\frac{1}{\xi}, +\infty\right), & \xi > 0 \end{cases}.$$

The parameter $\mu \in \mathbb{R}$ is the location parameter, the parameter $\psi > 0$ is the scale parameter, and ξ is the shape parameter. The shape parameter plays a crucial role because, for different values of this parameter, the generalized extreme value distribution transforms into one of the three types of extreme value distributions. When $\xi = 0$, the distribution is spread across all values on the x-axis, essentially forming the Gumbel distribution. For $\xi > 0$, it corresponds to the Frechet distribution, while for $\xi < 0$, it corresponds to the Weibull distribution.

2.1. **Block maxima.** Using the block maxima (BM) method, we divide the time interval into equal parts or blocks and model only the maximum losses for each block, based on the generalized extreme value distribution. The main idea in modeling with this method is that the distribution of block maxima for a sufficiently large n can be approximated by the Generalized Extreme Value distribution $H_{\xi,\mu,\psi}$. In this case

$$VaR_\alpha = \begin{cases} \mu - \frac{\psi}{\xi}\left(1 - (-n\ln\alpha)^{-\xi}\right), & \xi \neq 0 \\ \mu - \psi\ln(-n\ln\alpha), & \xi = 0 \end{cases}, \tag{2.1}$$

where α is the significance level of VaR_α.

2.2. **Peak Over Threshold.** Unlike the block maxima method, a popular modeling method using extreme value theory is the Peak Over Threshold (POT) method. In modeling with this method, we consider all values that exceed a specified "high level" or threshold. Let u be the high threshold. The distribution function of the values of the random variable X that exceed the threshold u is defined as

$$F_u(y) = P\{X - u \le y | X > u\},$$

where the right side of the equation represents conditional probability, essentially representing the probability that the values of the random variable X exceed the threshold u by at least y, given that X exceeds u. This conditional probability can be expressed as

$$F_u(y) = \frac{P\{X - u \le y, X > u\}}{P\{X > u\}} = \frac{F(y+u) - F(u)}{1 - F(u)}.$$

The distribution function of excesses above the threshold u is defined as

$$F_u(x) = P\{X - u \le x | X > u\} = \frac{F(x+u) - F(u)}{1 - F(u)},$$

for $0 \le x \le r_F - u$, where $r_F \le \infty$ is the right endpoint of the support of F.

Definition 2.1 (Generalized Pareto Distribution (GPD))**.** The distribution function of the Generalized Pareto Distribution is given by

$$G_{\xi,\beta}(x) = \begin{cases} 1 - \left(1 + \frac{\xi x}{\beta}\right)^{-\frac{1}{\xi}}, & \text{if } \xi \ne 0 \\ 1 - e^{-\frac{x}{\beta}}, & \text{if } \xi = 0 \end{cases},$$

where $\beta > 0$, and $x \ge 0$ if $\xi \ge 0$, and $0 \le x \le -\frac{\beta}{\xi}$ if $\xi < 0$. The parameters ξ and β represent the shape and scale parameters, respectively.

Theorem 2 (Pickands–Balkema–de Haan)**.** *Let F_u be the distribution function of excesses above the threshold u, and $r_F \le \infty$ be the right endpoint of the support of F. Then, $F \in MDA(H_\xi)$, $\xi \in \mathbb{R}$ if and only if there exists a positive measurable function $\beta(u)$ such that*

$$\lim_{u \to r_F} \sup_{0 \le x < r_F - u} \left| F_u(x) - G_{\xi,\beta(u)}(x) \right| = 0.$$

This implies that the conditional distribution of the observations exceeding the threshold u can be "asymptotically" modeled using the Generalized Pareto Distribution. In fact, Theorem 2 has widespread applications in the real world as it has been demonstrated that the Generalized Pareto Distribution is an essential distribution for modeling losses that exceed a high threshold. In this case,

$$VaR_\alpha(X) = \begin{cases} u + \frac{\beta}{\xi}\left(\left(\frac{1-\alpha}{\overline{F}(u)}\right)^{-\xi} - 1\right), & \xi \ne 0 \\ u - \beta \ln\left(\frac{1-\alpha}{\overline{F}(u)}\right), & \xi = 0 \end{cases}, \tag{2.2}$$

where $\overline{F}(x) = 1 - F(x) = P\{X > x\}$, n is the number of observations.

3. Data and methodology

In this research, we conduct an analysis of the Macedonian Stock Exchange Index, MB10. The study covers the period from March 2007 to February 2020, encompassing the global economic crisis of 2008 and the onset of the COVID-19 crisis. The primary objective of our analysis is to investigate the behavior of the index during this specified time frame, and determine if these movements can provide early warnings regarding the significant changes that occurred in March 2020 due to the COVID-19 crisis.

The index values we examine are denominated in index points for the MB10 index, which are publicly available on the Macedonian Stock Exchange's official website. Following standard practices in financial data analysis, we work with the logarithm of daily returns. The daily percentage log-returns are calculated by taking the logarithm of the ratio of the daily index values multiplied by 100, i.e. $r_t = 100 \log \left(\frac{I_t}{I_{t-1}} \right)$ where I_t represents the closing value of the index on each trading day. Specifically, we are examining the daily log losses (negative returns), thus reconstructing the series to represent the losses. The series contains data from 3196 traiding days (closing prices).

Table 1. Descriptive statistics of the considered dataset

Statistic	N	Mean	St. Dev.	Min	Max
MBI 10	3,196	2,925.033	1,676.800	1,556.960	10,057.770
Negative returns	3,196	0.001	0.570	−10.740	10.014

Below, we present a graphical representation of the daily dynamics of index values and time series of daily log-losses for the Macedonian Stock Exchange Index.

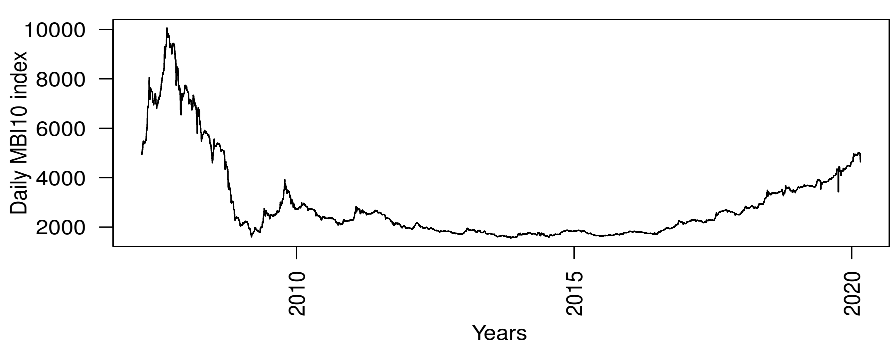

Figure 1. Daily dynamics of MBI10 index

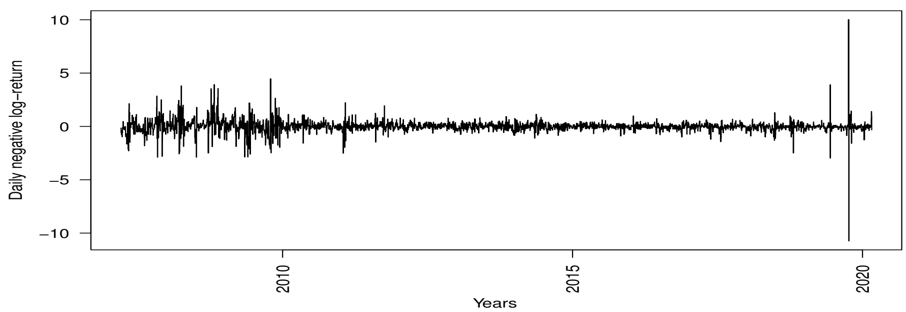

FIGURE 2. Daily negative log returns of MBI10 index

According to the graphical representation of the series in Figure 1, we can observe some deviations from the normal trends of the index in the period from 2007-2008. These notable deviations result from specific investment policies during that period, namely due to increased foreign investments in the country and an influx of foreign capital.

4. EMPIRICAL RESULTS

4.1. **Block maxima method.** In this research, as seen before, we consider data from 3196 traiding days of the MBI10 index. The block maxima method will be used to determine the maxima within specific time intervals. Specifically, we will consider dividing the series into quarters and months. In the quarterly division, we examine 53 quarters, representing blocks with approximately 61 values of the index in each block. In the case of monthly intervals, we consider 156 blocks, each with approximately 21 values of the index. The graphs on Figure 3 and Figure 4 display the block maxima in each of the quarterly and monthly divisions. Additionally, the Figure 5 show histograms for quarterly and monthly maxima. For the obtained maxima in each of the mentioned divisions, we will associate the corresponding extreme value distribution. In fact, using the method of maximum likelihood estimation, we will determine the estimators for the location, scale, and shape parameters, both in the case of dividing the series into quarters and months. The respective values for the estimators, according to the data analysis, are provided in Table 2. In both cases we have a shape parameter larger than zero, $\widehat{\xi} > 0$, so we have Generalized Extreme Value Distribution which corresponds to Frechet distribution.

Table 2. Maximum likelihood estimators

Parameter	Quarterly BM	Monthly BM
$\widehat{\mu}$	0.7095195	0.4362130
$\widehat{\psi}$	0.5662507	0.3363466
$\widehat{\xi}$	0.3422667	0.3473133

Table 3. Value at Risk (VaR) at different significance levels

Significance level	Quarterly BM	Monthly BM
0.95	0.175	0.412
0.99	1.011	1.130
0.999	3.363	3.172

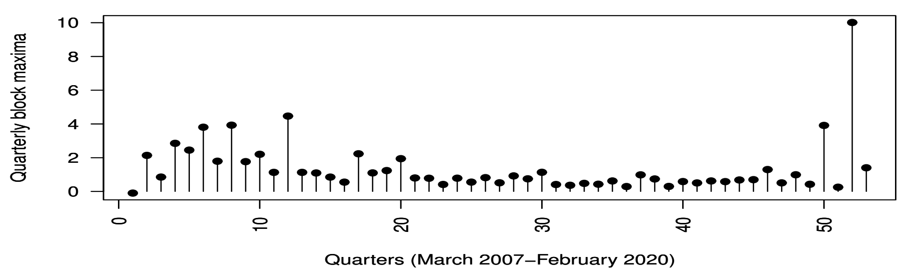

Figure 3. Quarterly block maxima

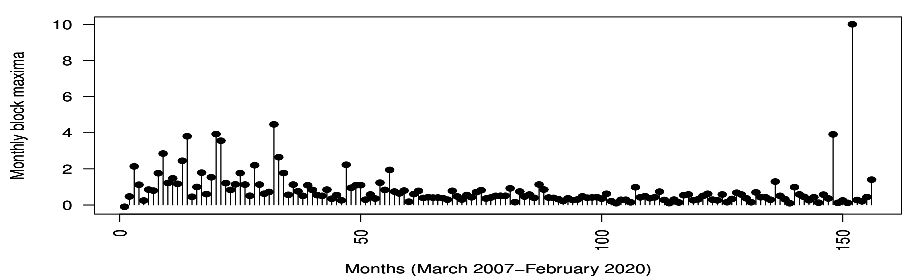

Figure 4. Monthly block maxima

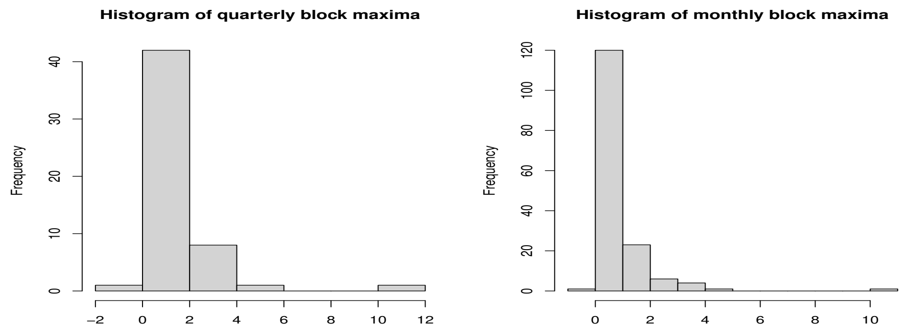

FIGURE 5. Histogram of quarterly (left) and monthly (right) block maxima

4.2. **Peak Over Threshold.** Peak Over Threshold method will be used to determinate the value of the threshold and the observation, and its number that exceed that threshold. According to the Mean Excess function given on Figure 6, the most appropriate value for the threshold u is in the interval $(0.5, 1)$.

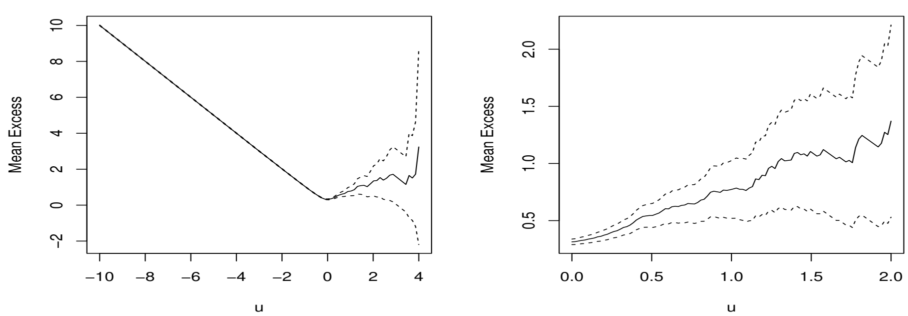

FIGURE 6. Empirical mean residual life plot for log negative returns. The solid line represents the empirical MRL with the dashed lines representing a 95% confidence interval.

In order to determinate the most appropriate value of u, we will consider the Parameter Threshold Stability Plot of the Modified Scale parameter and for the Shape parameter of GPD, given in Figure 7 and Figure 8, respectively.

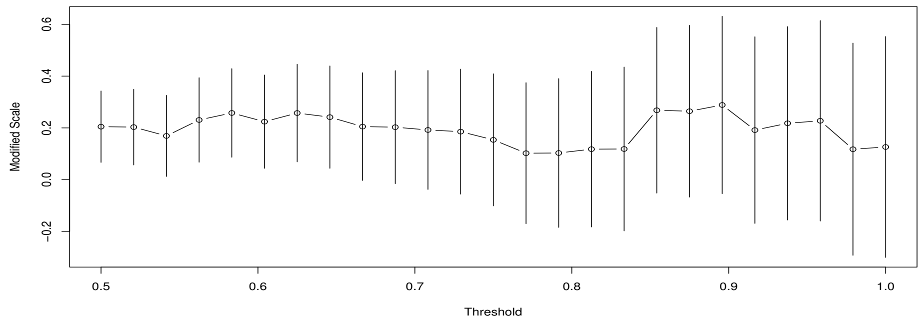

FIGURE 7. Parameter Threshold Stability Plot of the Modified scale parameter

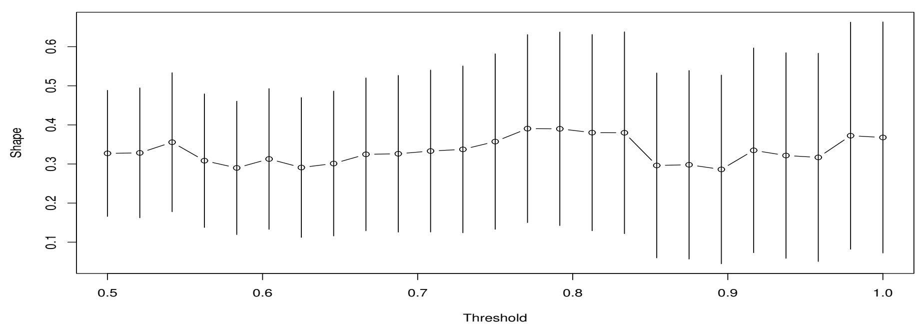

FIGURE 8. Parameter Threshold Stability Plot of the shape parameter

Looking at the shape estimate plot, we observe it behaves like a constant at the neighbourhood of 0.7. Also, the same conclusion can be obtained by the modified scale parameter plot. Hereby we are looking for a threshold value at the neighbourhood of 0.7. This is the reason we are considering the values $u = 0.68$, $u = 0.7$ and $u = 0.72$ for the threshold. In Table 4, the estimates of the parameters of the corresponding Generalized Pareto distribution and its standard errors are given.

So, we are considering the data whose value exceeds the given thresholds, separetly, such that with the maximum likelihood estimation we obtain the parameters of the corresponding Generalized Pareto distribution, given in Table 4.

As one can observe, the parameter estimate values are quite close. The smallest standard errors occur when the threshold is $u = 0.68$, so we consider that value as the most appropriate threshold setting. With the threshold set in this way, we have 167 observations from the examined data set that exceed this value.

TABLE 4. Maximum likelihood estimates and its standard errors for the parameters of GPD

	Maximum	likelihood	estimates	Standard	errors of	estimates
	$u = 0.68$	$u = 0.7$	$u = 0.72$	$u = 0.68$	$u = 0.7$	$u = 0.72$
ξ	0.335	0.332	0.353	0.103	0.105	0.110
β	0.417	0.426	0.415	0.053	0.055	0.055

Therefore, in this case, we are examining the Generalized Pareto distribution with parameters $\xi = 0.335$ and $\beta = 0.417$, whose standard errors are 0.103 and 0.053, respectively. From this associated distribution, we obtain that the Value at Risk in this case is 0.6985232, 1.6015064 and 4.1195725 with significance level $\alpha = 0.95$, $\alpha = 0.99$ and $\alpha = 0.999$, respectively.

5. Conclusions

The theory of extreme values plays a crucial role in modeling stock market index movements, as confirmed by numerous studies. According to the previous analysis of the MB10 index using the Block Maxima method, we can conclude that even in normal trends, without the threat of COVID-19, expected losses could exceed 3%, at a significance level of 0.999. The same conclusion is reached when considering the monthly partition of the series. On the other hand, when the series is examined using the Peak Over Threshold method, we conclude that at a level of 0.999, losses of 4.12% were expected. Thus, such fluctuations in the index were anticipated according to its trends, and the COVID crisis only accelerated and further increased the loss, i.e., the decline in the stock index. According to publicly available data, the stock index experienced a decline of 3.9% and 4.2%, during March 2020, relative to the value of the index from the previous trading day, which are very close values to the obtained values for the value at risk in our analyses. In fact, the peak over threshold method provides very close values to the real ones, unlike the block maxima method. However, in this study, the results obtained with both methods are important, as they serve as good indicators in any case for the trends of the index. This research serves as a basis for further analyses of the MB10 index using additional methods and risk measures to predict future index trends.

Acknowledgement

The second author is supported by Ss. Cyril and Methodius University's project "Mathematical models and applications" (Grant: NIP.UKIM.20-21.6).

References

[1] V. O. Andreeva, S. E. Tinykovb,O. P. Ovchinnikovaa, G. P. Parahin, *Extreme Value Theory and Peaks Over Threshold Model in the Russian Stock Market*, *Journal of Siberian Federal University. Engineering & Technologies 1*, 2012, 111–121.

[2] S. Coles, *An Introduction to Statistical Modeling of Extreme Values*, Springer, 2007.
[3] P. Embrechts, C. Klüppelberg, T. Mikosch, *Modelling extremal events for insurance and finance*, Springer, 1997.
[4] A. J. McNeil, R. Frey, P. Embrechts, *Quantitative risk management concepts, techniques and tools*, Princeton Univeristy Press, 2015.
[5] C. Omari, S. Mundia, I. Ngina, *Forecasting Value-at-Risk of Financial Markets under the Global Pandemic of COVID-19 Using Conditional Extreme Value Theory*, Journal of Mathematical Finance,10,569–597, 2020, https://doi.org/10.4236/jmf.2020.104034.
[6] N. Pratiwi, C. Iswahyudi, R. I. Safitri, G. P. Parahin, *Generalized extreme value distribution for value at risk analysis on gold price*, *Journal of Physics: Conference Series*, 2019, https://dx.doi.org/10.1088/1742-6596/1217/1/012090.
[7] A. Qayyum, F. Nawaz, *Measuring financial risk using extreme value theory: evidence from Pakistan*, *Economics Bulletin*, Vol. 31, 2010.
[8] G. Rodríguez , *Extreme value theory: An application to the Peruvian stock market returns*, Revista de Métodos Cuantitativos para la Economía y la Empresa, Universidad Pablo de Olavide, Sevilla, Vol.23, 48–74, 2017.
[9] l. Shengyuan, *Empirical Analysis of Value at Risk (VaR) of Stock Portfolio Based on Python*, *Proceedings of the 2022 International Conference on mathematical statistics and economic analysis (MSEA 2022)*, Atlantis Press, 2022, 559–567, https://doi.org/10.2991/978-94-6463-042-8_80.
[10] https://www.mse.mk/

STEVO GJORGIEV
Ss. Cyril and Methodius University in Skopje
Faculty of Natural Sciences and Mathematics - Skopje
Arhimedova 3, North Macedonia
Email address: stevogjorgiev@pmf.ukim.mk, stevogjorgiev@gmail.com

ANETA GACOVSKA-BARANDOVSKA
Ss. Cyril and Methodius University in Skopje
Faculty of Natural Sciences and Mathematics - Skopje
Arhimedova 3, North Macedonia
Email address: aneta@pmf.ukim.mk, agacovska@gmail.com

Received 28.11.2023
Revised 5.6.2024
Accepted 11.7.2024