Mitteilungen der Österreichischen Geographischen Gesellschaft, 165. Jg., S. 281–305
(Annals of the Austrian Geographical Society, Vol. 165, pp. 281–305)
Wien (Vienna) 2023, https://doi.org/10.1553/moegg165-081

Physische Geographie

Physical Geography

Determining Streamflow Patterns in North Macedonia Using PCA and AHC Analysis

Ivan Radevski, Svemir Gorin, Vladimir Zlatanoski (all Skopje),
Ana M. Petrović (Belgrade [Beograd]), Arse Kuzmanoski, Emilija Manevska,
and Blagoja Markoski (all Skopje)*

*Initial submission / erste Einreichung: 11/2022; revised submission / revidierte Fassung: 11/2023;
final acceptance / endgültige Annahme: 12/2023*

with 10 figures and 3 tables in the text

Contents

Summary

The main objective of this scientific study is the systematical arrangement of the average monthly streamflow in the Republic of North Macedonia by using adequate methodology.

* Prof. Dr. sci. Ivan Radevski (*corresponding author*), Vladimir Zlatanoski, Ph.D., M.Sc., Prof. Dr. sci. Svemir Gorin, Arse Kuzmanoski, B.Cs., Emilija Manevska, M.Sc., Prof. Dr. Blagoja Markoski, all: Institute of Geography, Faculty of Natural Sciences and Mathematics, Ss. Cyril and Methodius University in Skopje, str. Arhimedova 3, 1000 Skopje, North Macedonia; Ana M. Petrović PhD, Geographical Institute "Jovan Cvijic" of the Serbian Academy of Sciences and Arts, Đjure Jakšića 9, 11000 Belgrade, Serbia. – Emails: ivan.radevski@pmf.ukim.mk; zlatanoski@pmf.ukim.mk; svemir@pmf.ukim.mk; arsekuzmanoski@pmf.ukim.mk.

The river streamflow patterns in North Macedonia have yet to undergo comprehensive analysis, leaving a critical gap in our understanding of their dynamics and potential impact on water resource management. This study aims to address this knowledge deficit by conducting a thorough investigation of river streamflow patterns in the region, shedding light on their variability, trends, and implications for sustainable water resource planning. The prevailing pattern of the streamflow regime in North Macedonia was defined according to their monthly data, for the period 1961–2010, using Principal Component Analysis (PCA) and Agglomerative Hierarchical Clustering (AHC), using Ward's agglomerative method and Euclidean distance.

In this paper, the analysis involved processing data collected from 30 hydrological gauging stations situated along 22 perennial streams. After the PCA analysis, the factor loadings classification was carried via AHC. It results with three different patterns and five sub-patterns. The first pattern (R1) covers the northwestern part of the country, located in the high mountains Šar Mountain, Korab-Deshat, Jakupica Mountain and the Ohrid-Prespa Region. It features strong a magnitude in late spring with 25 percent of the total annual streamflow in May and a minimum in late September. The first half of the summer is still relatively rich in water. The second pattern (R2) covers the largest north, northeastern and southern part of the country (Kriva Reka catchment, Middle Povardarie and Pelagonia Basin) with a not so extreme maximum during April (16 % of the total streamflow) and drier summers with the minimum during August. The third pattern (R3) covers a smaller area and is located in the southeastern part of the country (Strumitsa-Radovish Basin). It contains gauging stations with an earlier maximum during March to April (each around 15 %) and the minimum during August to September, where the annual streamflow amplitude is much lower compared to other regions. From the two main streamflow drivers, snowmelt influence is less significant going eastward, while the rainfall influence is much stronger going eastward.

Keywords: Streamflow patterns, streamflow regimes, main drivers, North Macedonia, Pardé coefficient, principal component analysis (PCA), agglomerative hierarchical clustering (AHC)

Zusammenfassung

Analyse von Abflussmustern im Flusssystem Nordmazedoniens mittels PCA- und AHC-Analyse

Das Hauptziel dieser wissenschaftlichen Studie ist die systematische Bestimmung des durchschnittlichen monatlichen Abflusses in der Republik Nordmazedonien mit Hilfe einer geeigneten Methodik. Die Abflussmuster der Flüsse in Nordmazedonien wurden bisher noch nicht umfassend analysiert, was eine kritische Lücke in unserem Verständnis ihrer Dynamik und ihrer möglichen Auswirkungen auf die Bewirtschaftung der Wasserressourcen hinterlässt. Die vorliegende Studie zielt darauf ab, dieses Wissensdefizit zu beheben, indem sie die Abflussmuster der Flüsse in der Region gründlich untersucht und ihre Variabilität, Trends und Auswirkungen auf die nachhaltige Planung von Wasserressourcen beleuchtet. Das vorherrschende Muster des Abflussregimes in Nordmazedonien wurde

anhand der monatlichen Daten für den Zeitraum 1961–2010 mit Hilfe der Hauptkomponentenanalyse (PCA) und der agglomerativen hierarchischen Clusterbildung (AHC) unter Verwendung der agglomerativen Methode von Ward und der euklidischen Distanz definiert.

In dieser Studie wurden die Daten von 30 hydrologischen Messstationen entlang von 22 mehrjährigen Flüssen ausgewertet. Nach der PCA-Analyse wurde die Klassifizierung der Faktorenladungen mittels AHC durchgeführt. Dabei ergaben sich drei verschiedene Muster und fünf Untermuster. Das erste Muster (R1) deckt den nordwestlichen Teil des Landes ab, der in den Hochgebirgen Šar, Korab-Deshat, Jakupica und der Region Ohrid-Prespa liegt. Es zeichnet sich durch ein starkes Ausmaß im späten Frühjahr mit 25 Prozent des gesamten jährlichen Abflusses im Mai und einem Minimum Ende September aus. Die erste Hälfte des Sommers ist noch relativ wasserreich. Das zweite Muster (R2) umfasst den größten nördlichen, nordöstlichen und südlichen Teil des Landes (Einzugsgebiet der Kriva Reka, mittleres Povardarie- und Pelagonien-Becken) mit einem nicht so extremen Maximum im April (16 % des gesamten Abflusses) und trockeneren Sommern mit einem Minimum im August. Das dritte Muster (R3) umfasst ein kleineres Gebiet und befindet sich im südöstlichen Teil des Landes (Strumitsa-Radovish-Becken). Es enthält Messstationen mit einem frühen Maximum im März bis April (jeweils etwa 15 %) und dem Minimum im August bis September, wo die jährliche Abflussamplitude im Vergleich zu anderen Regionen viel geringer ist. Von den beiden wichtigsten Abflussfaktoren ist der Einfluss der Schneeschmelze vom Westen des Landes gegen Osten hin immer weniger bedeutend, während der Einfluss der Niederschläge in östlicher Richtung zunimmt und viel stärker ist.

Schlagwörter: Abflussmuster, Abflussregime, hauptsächliche Bestimmungsfaktoren („drivers"), Nordmazedonien, Pardé Koeffizient, Hauptkomponentenanalyse (PCA), agglomerative hierarchische Clusterbildung (AHC)

1 Introduction

One of the significant issues in hydrology is the classification of the streamflow patterns. This work belongs to a great number of studies of the streamflow regime patterns, with the objective to gain a wider application in the water management. In a southeast European country like North Macedonia, where reliable measurements and operational gauging stations have become scarcer, particulary after 2010, when the National Hydrometeorological Service (NHMS) had a problem with lack of staff and finances (less number of active gauging stations after 2000 and especially after 2010, small number of staff in the institution NHMS), this study is of importance on a local and a regional Balkan level as it provides essential insights into the river streamflow patterns. Understanding these patterns is vital for informed decision-making in water resource management, helping mitigate the adverse effects of water scarcity, floods, and other hydrological challenges in the face of limited data infrastructure. It would be of significant value, particularly for a country with reduced instrumental and measurement conditions to choose streamflow time series without gaps for better quality analysis.

In the Balkan region there are two studies about the river regime classification, but only with graphical comparison as a methodological approach of maximum and minimum monthly streamflow and it's differenciation, or three peak months with high waters and three months with lowest values in streamflow about river regimes of Bulgaria (Hristova 2007) and the Balkan Peninsula (Stanescu 2004).

The results about Bulgaria show three distinct hydrological regions and seven hydrological districts have been demarcated. The hydrological region characterised by a Mildness-Continental type of a river regime exhibits three discernible streamflow phases, specifically the spring high water phase, the summer-autumn low water phase, and the transitional phase. However, in comparison the region with a Continental-Mediterranean type of river regime manifests two primary streamflow phases, comprising a prolonged winter-spring high water phase and a summer-autumn low streamflow phase. Meanwhile, the region typified by a transitional water regime (located near the border between North Macedonia and Bulgaria) features four distinct streamflow phases, which include a high water phase from March/April to June/July, a prolonged low water phase from July/August to October, and an annual pattern of increased streamflow in November or December, followed by a decline in January/February/March (Hristova 2007).

The article by Stanescu (2004) about river regimes in the Balkan Peninsula includes a graphical comparison analysis about North Macedonia with three distinctive river regimes: The Western North Macedonia upland Šar Mountains region (high streamflows in April to May and low streamflows in September to October, the Eastern-western mountain and central hilly plateau regime (maximum from February to April and a minimum phase from June to September), and the Northeast upland mountain zone (maximum in March/April and minimum phase in August/September).

Streamflow regionalisation is useful for transferring data from gauged to ungauged basins and obtain hydrological information about ungauged basins (Blöschl and Sivapalan 1995). The natural streamflow regime is very important for the precision of the streamflow pattern results (Snelder et al. 2009).

Previous studies analysed streamflow trends, using the Mann-Kendall trend test and Sen's slope (periods from 1961 to 2000 and 1961 to 2010 for several gauging stations) on 13 natural gauges without an anthropogenic influence. The results show a statistically significant decreasing trend ($p<0.05$) of the average annual streamflow in four of the 13 gauging stations, a decreasing trend of the maximum streamflow in also four of the 13 gauging stations, and five of the 13 gauging stations have a decreasing trend in minimum annual streamflows, with very strong trends ($p<0.01$). In the seasonal results, the spring is the season with only one of the 13 stations with a statistically negative significant trend in minimum, average and maximum streamflow, so the streamflow in spring is most stable in the country. In winter there are statistically significant negative trends, particularly for the average streamflow (thre of the 13 gauging stations), in autumn there are statistically significant decreasing trends, particularly for the minimum streamflow (three of 13 gauging stations). The summer results show quite a large number of gauging stations with statistically significant decreasing trends (five of 13 stations). According to the results, there are no statistically significant *increasing* streamflow trends in annual and seasonal results in North Macedonia (Radevski et al. 2018).

The aim of our study is to obtain a new comprehensive classification of the streamflow patterns in the country, which will be useful for the water management institutions, for private bodies working in the sphere of water resources and in addition for the international cooperation in the projects because almost all hydrological network is shared with neighbourhood countries. Elsewise, this study could be a trigger for the future wider Balkan classification with the almost similar methodology. In the future it will be possible to compare streamflow patterns in North Macedonia with other results from the Balkans.

2 Study Area

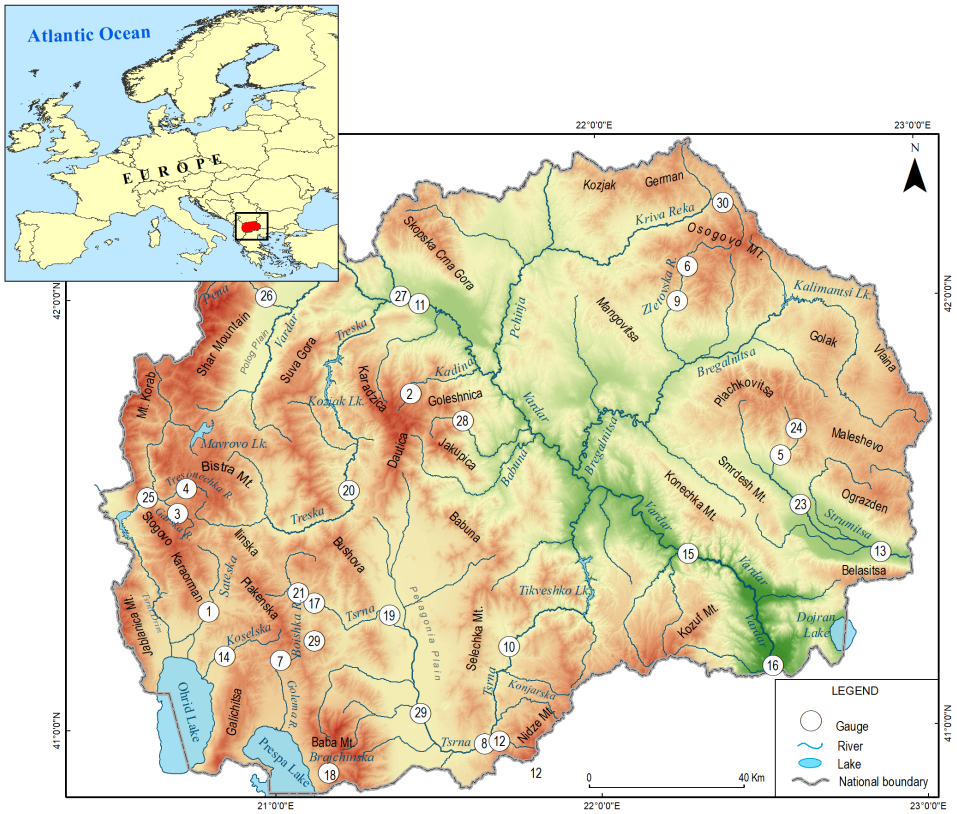
The Republic of North Macedonia is located in Southeastern Europe, in a central part of the Balkan Penninsula and it covers an area of 25,713 km2. The ground forms are predominantly hilly-mountainous and the elevation varies from the lowest point of 54 m a.s.l. (in Gevgelija Basin) to 2,764 m a.s.l. (peak Golem Korab on the Korab Mountain), with almost 30 basins and almost the same number of mountain ridges. According to the geographical latitude and the ground conditions, it is divided to the larger Continental climate region in the north, and a smaller moderate Mediterranean climate region to the Southeast and Southwest (Radevski et al. 2018).

North Macedonia exhibits diverse climatic regions owing to its geographical latitude and relief conditions. The country can be roughly divided into distinct areas with varying climates:

1. The larger region, predominantly characterised by a continental climate, encompasses several plains including the Prespa Plain, Kichevo Plain, Kumanovo Plain, Pijanec and Slavishte Plains, Polog Plain, and the expansive Pelagonia Plain. In this region, the precipitation pattern exhibits two notable peaks in spring and autumn, along with several minima. The results for the summer are showing the primary minimum, while winter has a secondary and less pronounced minimum.
2. Moving south-east and south-west, there are two smaller regions with a Sub-Mediterranean climate. These areas cover Gevgelija-Valandovo, Dojran Plain, Ohrid-Struga, where the influence of the Mediterranean climate is particularly strong. However, there are occasional years when this influence extends even further, affecting the valley of the river Vardar to Skopje and the valley of the river Strumica in the Strumica-Radovish Plain.
3. The country's highest mountains, including Šar Mountain, Korab, Jablanica, Baba, and Jakupica Mountain, situated parallel to the plains in the western part of North Macedonia, are home to regions with mountain climates. At elevations exceeding 2250 meters, an alpine climate prevails. These mountainous and alpine regions receive the highest precipitation levels, often exceeding 1000 mm per year, particularly on the tallest peaks. Winter marks the peak of precipitation in these areas, while late summer is characterised by the lowest rainfall.
4. Notably, the climate changes considerably as elevation increases, shifting from a sub-Mediterranean climate to a continental one and finally to a mountainous climate.

There is one exception in Ohrid, located in the southwest near the large Ohrid Lake. Even at higher elevations here, a sub-Mediterranean climate persists (Radevski et al. 2018).

This climatic diversity is influenced by the complex interplay of geographical factors and is essential for understanding North Macedonia's weather patterns (Kendrovski and Spasenovska 2011).



Source: Own design

Figure 1: Topographic map of North Macedonia showing the 30 gauging stations (see Table 1)

The selection of the gauging stations is mainly "natural" (without larger significant human influence) although several downstream gauges in the river Vardar basin have upstream dams, which significantly have an influence on the downstream streamflow, hence used in the study according to the small number of countinously working gauging stations. Dams and reservoirs play a significant role in regulating water resources. They can mitigate

N	Basin	River (R) Lake (L)	Tributary	Gauging station	Upstream drainage (km^2)	Gauge elevation (m)	Qm (m^3/sec)	Time series period	Impound-ment ratio
1	Adriatic	Ohrid L.	Sateska	Botun	368.00	743.18	5.960	1951–2010	/
2	Aegean	Vardar	Kadina	Krusa	9.46	1,305.91	1.380	1961–2000	/
3	Adriatic	Mala R.	Garska	Gari	25.95	1,011.92	1.170	1961–2000	/
4	Adriatic	Mala R.	Tresonecka	Tresonce	71.24	601.00	1.142	1961–2000	/
5	Aegean	Strumitsa	Plavaja	Podaresh	118.34	393.37	0.917	1961–2004	/
6	Aegean	Zletovska	Emirica	Emirica-Vliv	32.42	559.70	0.493	1961–2010	/
7	Aegean	Koselska R.	Leva R.	Leva Reka	25.12	1,004.51	0.434	1961–2000	/
8	Aegean	Tsrna	Konjarska	Konj. Skocivir	63.00	576.10	0.751	1961–2000	/
9	Aegean	Bregalnica	Zletovska	Zletovo	117.96	519.51	1.980	1961–2000	0.36
10	Aegean	Vardar	Tsrna	R. Most	4,526.00	369.25	22.790	1961–2010	0.17
11	Aegean	Vardar	/	Skopje	4,625.00	239.55	58.900	1951–2010	0.30
12	Aegean	Vardar	Tsrna	Tsrna-Skocivir	3,975.00	564.60	19.410	1961–2010	0.20
13	Aegean	Strumitsa	/	Novo Selo	1,636.00	199.82	3.860	1961–2010	0.61
14	Adriatic	Ohrid L.	Koselska	Kosel	98.50	763.96	1.310	1961–2010	/
15	Aegean	Vardar	/	Demir Kapija	21,350.00	94.27	125.800	1961–2010	0.45
16	Aegean	Vardar	/	Gevgelija	22,310.00	45.10	142.200	1951–2010	0.40
17	Aegean	Vardar	Tsrna	Dolenci	216.50	739.10	2.420	1961–2010	/
18	Adriatic	Pespa L.	Brajcinska	Brajcino	61.50	975.29	0.946	1961–2010	/
19	Aegean	Vardar	Tsrna	Bucin	657.87	597.49	5.550	1961–2000	1.39
20	Aegean	Vardar	Treska	M. Brod	886.00	530.38	11.080	1961–2010	/
21	Aegean	Vardar	Tsrna	Zeleznec	123.00	730.00	1.860	1961–2010	/
22	Aegean	Vardar	Tsrna	Novaci	2,543.00	572.65	12.820	1961–2010	0.31
23	Aegean	Strumitsa	Susheva	Susevo	468.00	253.00	1.620	1961–2010	/
24	Aegean	Bregalnica	Smiljanska	Smiljanci	81.00	445.59	0.686	1961–2010	/
25	Adriatic	Crn Drim	Radika	B. Most	750.90	592.24	17.720	1961–2000	0.64
26	Aegean	Vardar	Pena	Tetovo	170.49	482.00	3.900	1961–2000	/
27	Aegean	Vardar	Lepenec	Lepenec-Vliv	770.00	260.00	8.210	1961–2010	/
28	Aegean	Vardar	Topolka	Drenovo	74.40	540.00	1.200	1961–2000	/
29	Aegean	Tsrna	Boishka	Boishte	95.00	885.00	0.760	1961–2000	/
30	Aegean	Pchinja	Kriva Reka	Zhidilovo	64.60	782.00	1.250	1961–2000	/

Source: Data obtained from the National Hydrometeorological Service of North Macedonia

Table 1: Basic data for 30 gaugung stations in North Macedonia

the impacts of climate change by storing excess water during wet periods and releasing it during dry spells. However, these infrastructure projects can also induce changes in local climates and ecosystems, impacting river systems and their associated biodiversity (Falkenmark and Rockström 2004).

The country's aquatics belong predominantly to the Aegean Basin with 22,319 km^2, and to the Adriatic Basin with 3,350 km^2. The main river system is the Vardar basin with its tributaries: Lepenec, Pchinja, Bregalnica, Treska and the river Tsrna. It has a drainage area in North Macedonia of 20,535 km^2, and a total length of 388 km (301 km in North Macedonia), which streams directly to the Aegean Sea at the gulf of Thermaikos near the city of Thessaloniki in Greece. The second basin belongs to the river Strumitsa with a total length of 114 km, 81 km in North Macedonia (right tributary of the river Struma in Bulgaria). The main river in the Adriatic Basin is Crni Drim with two large lakes, Ohrid and Prespa (connected with an underground inflow to the lake Ohrid) and the largest tributary of Crni Drim from the north, the river Radika (Radevski et al. 2018). The selected gauging stations on the 23 streams are presented on Figure 1 with their basic characteristics in Table 1.

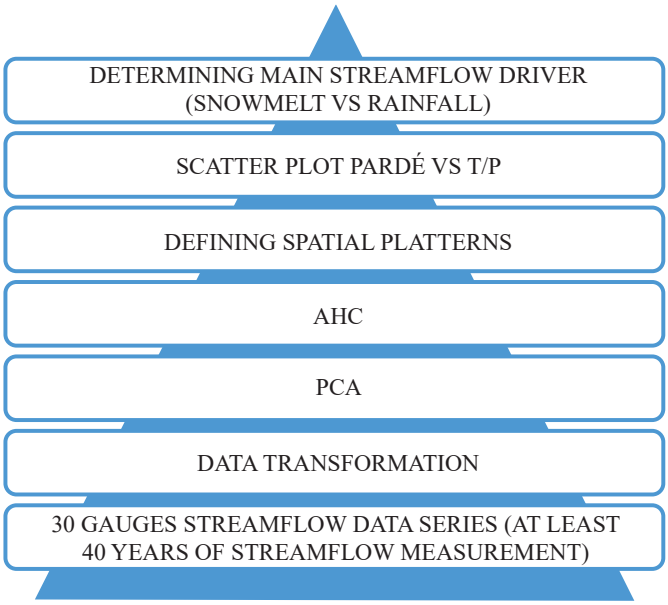
3 Data and Methodology

The geographic catchment classifications are usually presented as mosaic patterns (Olden et al. 2012). This is induced by a predominance of physiographic factors of the gauging stations in the complex relief structures. In contrast to the areas with uniform terrain, regional patterns are expected in the classification procedure. The streamflow data record series should be at least cover a time span of 25 years, otherwise the uncertainity of the results is increased (Kennard et al. 2010). In order to obtain a better spatial covering of the study area, we chose 30 gauging stations in North Macedonia with not concurent data length, but at least 40 years of continous measurement due to the above mentioned conditions. The original streamflow data was kept in the analysis, hence the impaundment ratio was calculated for a clear vision of the human affected streamflow gauging stations (Morán-Tejeda et al. 2011).

Different authors analysed the streamflow regime at the national, regional or worldwide scale, based on various classification procedures as principal components analysis (PCA) (Olden and Poff 2003; Masiokas et al. 2019; Di Prinzio et al. 2011; Kahya et al. 2008), agglomerative hierarchical clustering (AHC) (Stahl 2001; Bower et al. 2004), self organising maps (Di Prinzio et al. 2011) etc. The classification is mainly based on the average monthly streamflow and its timing (Haines et al. 1988). The geographically nearest study of such a streamflow classification was made in Turkey (Kahya et al. 2008). In another study, the physical geographic data of the catchments is also included (Razavi and Coulibaly 2013).

The monthly streamflow analysis allows a measure of magnitude timing and comprehensive explorations of the streamflow regime (Olden et al. 2012; Curran and Biles 2021).

The issue with the streamflow standardisation and classification was solved by Pardé (1955) in detail. He introduced the monthly *Pardé coefficient* (PC; relation between the mean monthly (Q_m) and mean annual (Q_y) streamflow (equation 1) in order to improve the



Source: Own design

Figure 2: Methodological pyramid of the article

comparability of different rivers with high streamflow variation. The coefficient therefore describes the mean monthly distribution of streamflow over the year, depending on the number of the maximum of the monthly Pardé coefficient over the year The difference between the maximum (PC_{max}) and the minimum (PC_{min}) values of the monthly Pardé coefficients is called amplitude (A) or range (R) (equation 2).

$$PC_m = \frac{Q_m}{Q_y} \tag{1}$$

$$A = PC_{max} - PC_{min} \tag{2}$$

Additionally, for a better view of the data, besides the Pardé coefficient, the previously used practice of data tranformation in percents was performed (Masiokas et al. 2019), where the monthly percents were also plotted, so the results and the discussion chapter could be clearly described. For each of the gauging stations involved in this study, an average streamflow regime was determined by averaging the streamflow in each month over all years of record, recalculated as percents of the total annual streamflow.

Before starting the PCA and AHC analysis, testing normality was performed using the Shapiro-Wilk test. All of the analysed gauging stations were normally distributed time series, with $p > 0.05$ except the gauging station Gari on Garska River, $p = 0.009$, which indicates that the PCA analysis has to be performed adequately (Shapiro and Wilk 1965).

N	Location	Atributes (M – magnitude, T – timing, F – frequency, D – duration)	Temporal scale (daily, monthly and annually)	Methodology	Quotation
1	Austria	M, T	M	PCA and partitive cluster analysis	Laaha and Blöschl (2006)
2	Europe	M, T, F, D	D and M	AHC using Ward's method	Stahl (2001)
3	Scandi-navia	M, T	M	Streamflow regime defined by occurrence of the highest (three classes) and the lowest months' values (two cases)	Krasovskaia (1995)
4	Turkey	M	A	Non-hierarchical, k-means clustering	Kahya et al. (2008)
5	France	M, T	M	Streamflow proportion in four seasons and stream-flow origin (rainfall, snow-melt, glacial melt)	Pardé (1955)
6	Quebec, Canada	M, D	M	PCA	Assani et al. (2006)
7	UK	M, T	M	AHC using Ward's method	Bower et al. (2004)
8	USA	M, T	D, M, A	PCA	Olden and Poff (2003)

Source: Own compilation

Table 2: Previous studies of streamflow regime classification

Table 2 contains data from eight regional, continental and worldwide studies using different methodologies in analysing the streamflow regionalisation patterns. The hydrographs show the characterisation of the seasonal behaviour of rivers through parameters including the mean monthly streamflow and its variability (standard deviation). River regimes were identified by plotting the monthly standardised values averaged over the study period. The regimes were compared (z-scores) across seasons and regions. To identify the dominant river regimes in the study area, different classification procedures including PCA and AHC were attempted. A varimax rotated principal components analysis was performed.

Principal components analysis (PCA) is commonly used to identify patterns in climatic and hydrological data series (Kalayci and Kahya 2006), because it allows retention of the common features in the data, but also enables local peculiarities to be identified. The coefficients of such combinations are called "loading factors" and are representing the cor-

relation of the principal component with each original variable. The "maximum loading factor" allowed us to classify each gauging station, and the average monthly streamflow of each class was calculated to obtain representative hydrographs (Morán-Tejeda et al. 2011). The three crucial months with a maximum annual streamflow were March, April and May, so the Pardè coeficient, the coefficcient of variation (Cv) and the Range (R) were used for the PCA analysis with a scope for the results to be more exact.

Additionally, the AHC was performed on the previously mentioned data, using months in order to differentiate seasonal streamflow patterns. The same Pardè coefficients were used. The AHC used Eucledian distance and Ward's minimum-variance agglomerative method in XLSTAT software. The resulting dendrogram should clearly identify patterns and sub-patterns of the monthly streamflow Pardé coefficients. The classification was perfomed with two nested grouplings, one for the patterns, and the other for sub-patterns. The graph allows easy and evident differences between patterns and the sub-patterns (Curran and Biles 2021). The hierarchical clustering method allows the measurement of the similarities/dissimilarities between the gauging stations via the Eucledian distance method:

$$d_{ij} = \sqrt{\sum_{k=1}^{m} (x_{ik} - x_{jk})^2} \tag{3}$$

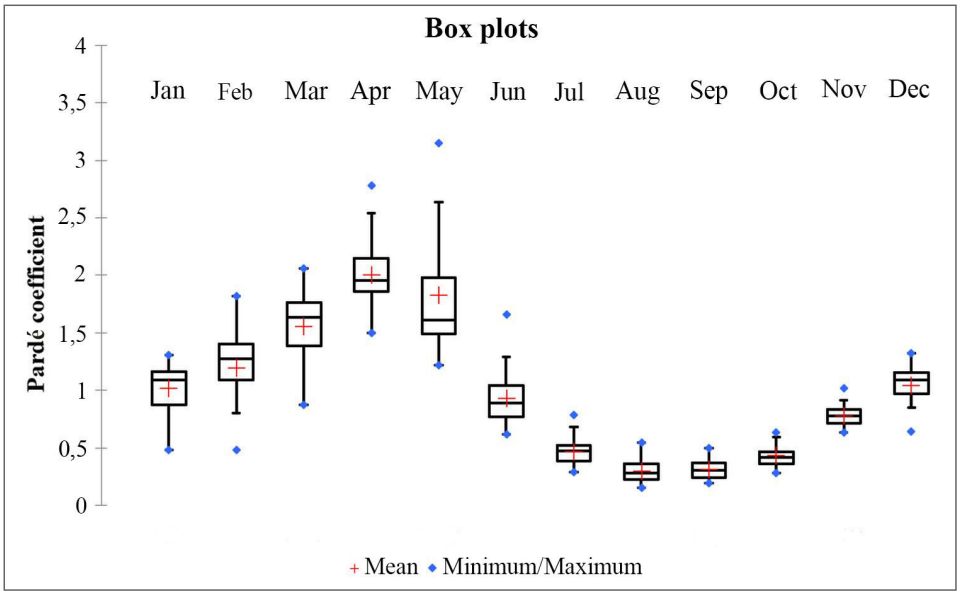
Above in the equation (Bower et al. 2004), the x_{ik} is k_{th} calculated statistic for gauging station i, and denotes the k_{th} calculated statistic for gauging station j, m=30, and d_{ij} is an Euclidian distance between each streamflow gauging station.

In the analysis, the obtained sub-patterns are plotted versus monthly precipitation sums and monthly mean air temperature as the main drivers, with a scope to detect the main factor of peak streamflow in the specific month at each gauging station, so the regime could be more or less rainfall dominant and more or less snowmelt dominant.

4 Results and Discussion

Figure 3 shows the variability in the monthly river streamflows for 30 gauging stations in North Macedonia. Obtained classes varied by overall seasonality of maximum streamflow depending on the month, measured by the range of Pardé coefficients). Thus, after the dry season (June to October) the streamflow increased in November and December, and peak streamflow featured from February to May with a gradually increasing and a sharp decrease occurring in June. Streamflows continued to decrease during the summer months, and the lowest streamflows occurred in August and September. In general, the variability of the monthly streamflow among the various gauging stations was moderate, as variation between the 1st and the 3rd quartiles (Figure 3, boxplot), which is larger in the first half of the year and smaller in the second half, where most months had variations below 0.5 units of the standard deviation. A great degree of variability (up to two units of standard deviation) was evident. The extremes are presented in the lengths of the whiskers (10th and 90th percentiles) and the blue dots (5th and 95th percentiles).

The shape of the Figure 3 shows a common river streamflow pattern of moderate Mediterranean influence with a higher mean elevation located in hilly mountanous area, with a general spring maximum in April. In May, the streamflow is still high, even higher at the high mountainous gauging stations than in April. The streamflow is abruptly decreasing in June with more than one Pardé unit of variation. Thus, after the minimum late summer season (July–October), the streamflow is gradually increasing from November to the peak streamflow which is detected in April and May.The streamflow variability during the summer months is very small, under the 0.2 units of variation (Figure 3, 1st and 3rd quantile presented with boxes). Hence, the spring variability is much higher, up to 1,5 units of variability in May (10th and 90th percentiles). Nevertheless, the variability of the extremes (blue dots) is much higher, almost two units of variability in May, and above one in February, March, April and June.



Solid horizontal line = median, the red cross = mean, box = 1st and 3rd quantile, whiskers 10th and 90th percentiles and blue dots = 5th and 95th percentiles.

Source: Own data analysis, own design

Figure 3: Boxplot of Pardé coefficients showing the monthly streamflow data dissemination.

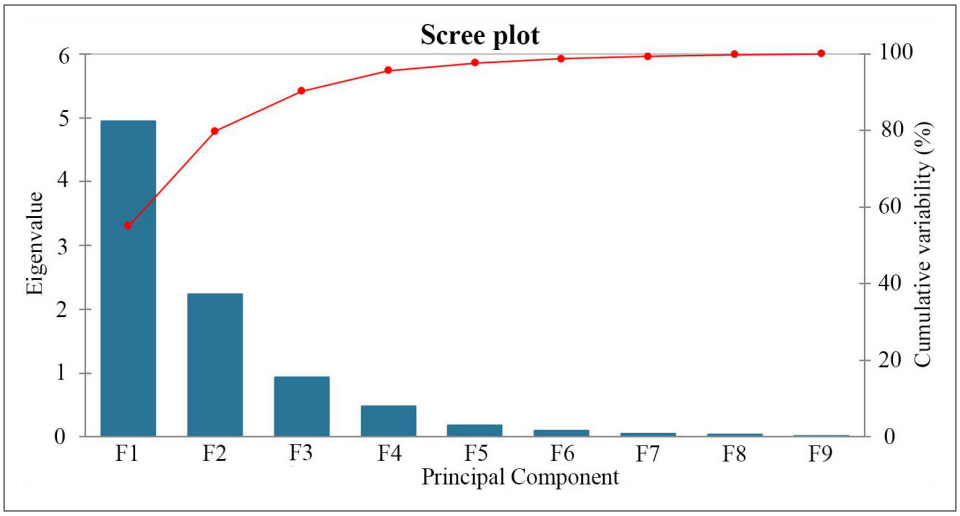
The principal components analysis has indicated that the first three principal components (F1, F2, and F3) explained above 90 percent of variance (55 %, 24 %, and 10 %, respectively; Table 3 and Figure 4). The selection of the maximum loading factor allowed us to identify which gauging stations are strongly correlated with each component.

The main problem in the study is the similarity of the regimes, considering the relatively small territory for analysis. These gauging stations have a maximum streamflow in the late winter, early, middle and late spring. Thus, April, May and March have the highest

Pardé coefficients, where the monthly streamflow is up to four times higher than the average annual streamflow. In that case, the values in these months will be crucial for the PCA.

Component	F1	F2	F3	F4
Eigenvalue	4.954	2.236	0.934	0.487
Variability (%)	55.040	24.839	10.383	5.407
Cumulative %	55.040	79.879	90.262	95.669

Table 3: Rotated results of the PCA of monthly river streamflows



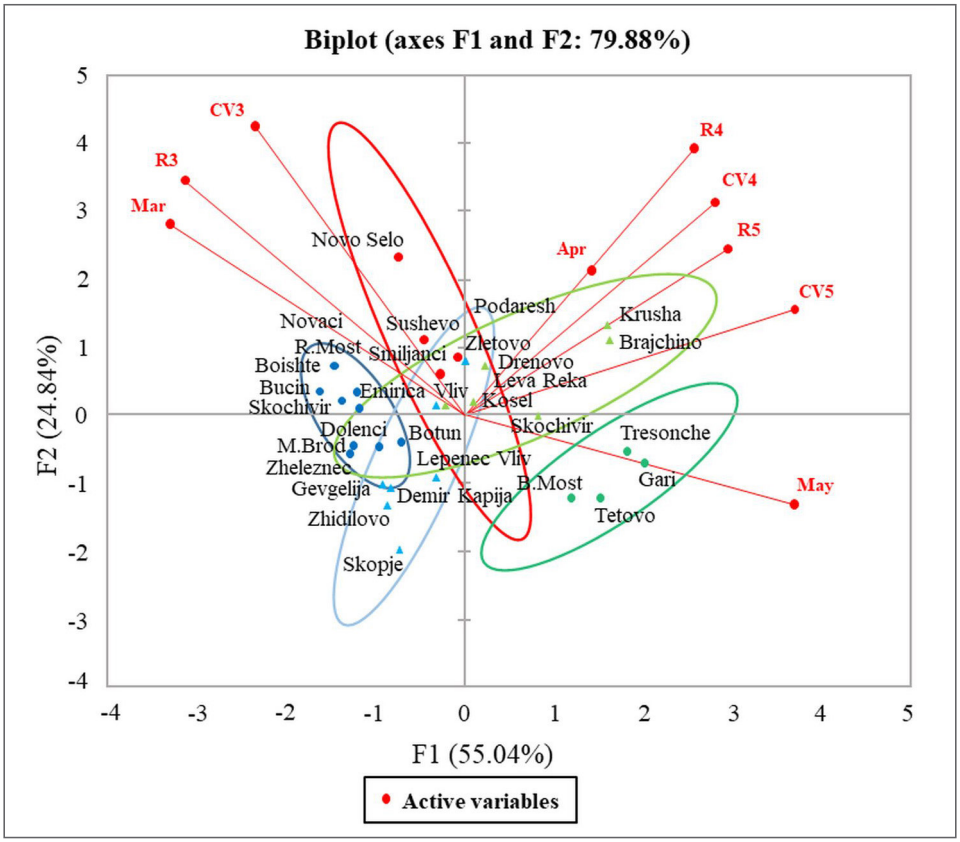
Source: Own data analysis, own design

Figure 4: Percentage of the total variability explained by each principle component (PC) on axis

Going from west to east, Pardé amplitudes decrease, and the maximum appears earlier. A PCA with orthogonal varimax rotation was used to identify spatial differences in river regimes and to classify the regimes into different types. This indicated the three principal components (F1, F2, and F3). The analysis performed with variables: a) mean monthly streamflow for 30 gauging stations transferred to Pardé's coefficient; b) monthly Pardé's coefficient of variation and c) monthly range of Pardé's coefficient, and the names of the gauging stations as observations.

The two biplots show PCA analysis for 30 gauging stations in North Macedonia. The previously mentioned spring maximum is varying from March to May, so these three months with its coefficients of variation (Cv) and ranges are plotted on Figure 5a and 5b. Figure 5a is more representative, covering 95 percent of the variance on both axes. Figure

5b (covering 65 % of the total variance) is taken into consideration because the April vector axis has a better rotation, showing a better comparing with the specific gauging stations having an April peak flow, whose red axis is much longer than Figure 5a.

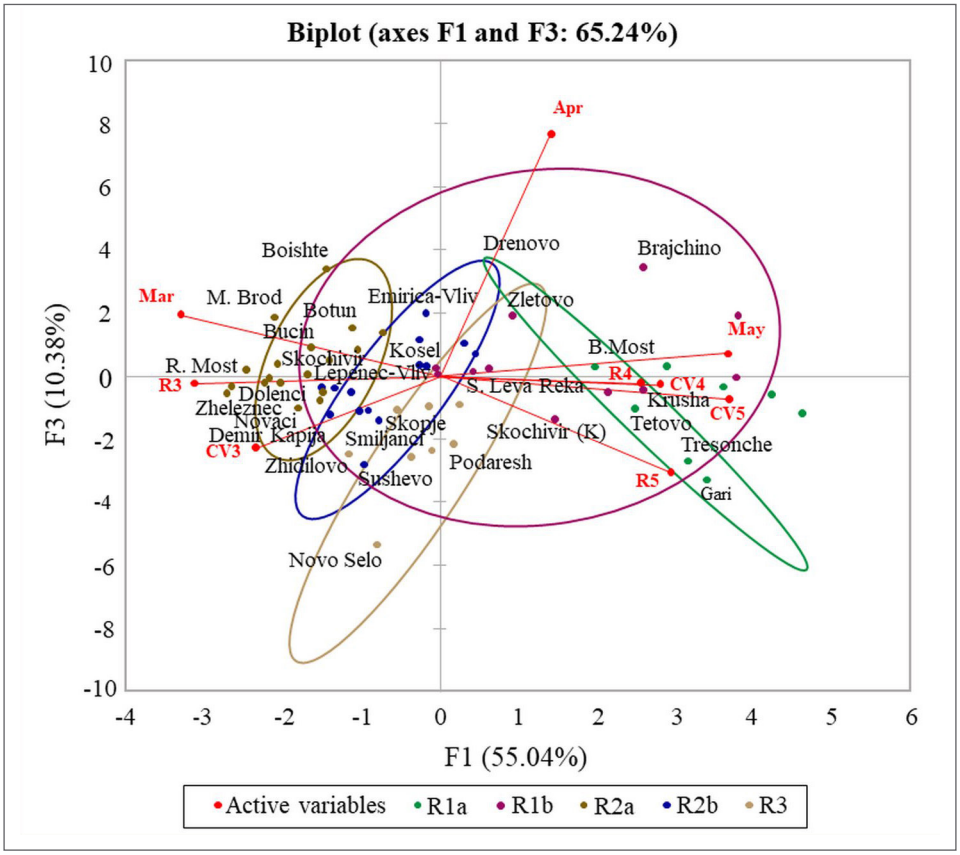


Source: Own data analysis, own design

Figure 5a: PCA biplots for: First two principal components, F1 and F2.. Gauging station colour and concentration ellipses indicate a concrete cluster group (R1a, R1b, R2a, R2b, R3)

According to Figure 5, the April maximum corresponds with the R1b, the May maximum is complementary with R1a, while the March is a crucial month for the streamflow patterns R2a and R2b (Figure 5b), excellently correlated, while R3 is well correlated with March, but entirely uncorrelated with the May axis on the PCA biplot (Figure 5b).

Figure 5a shows a clear cluster group of four R1a sub-patterns, it has a redundancy with the peak flow featuring in May (red vector axis), the month with a very intense snowmelt process (four gauging stations in the western part of the study area marked with a dark green colour). The R3 sub-pattern, plotted with four red dots and a red ellipse is

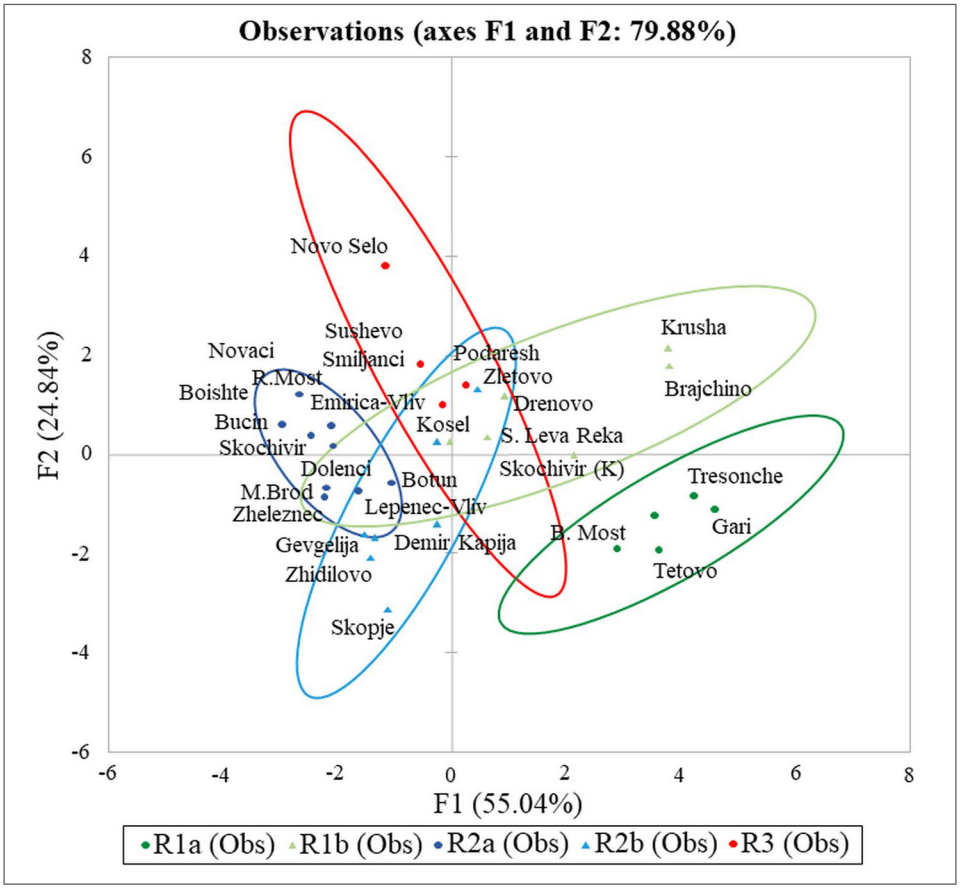


Source: Own data analysis, own design

Figure 5b: PCA biplots for: First and third principal component F1 and F3. Gauging station colour and concentration ellipses indicate a concrete cluster group (R1a, R1b, R2a, R2b, R3)

located opposite the R1a on the biplot, having higher values and it belongs to the F2 component, which means it has a significantly different streamflow pattern. The R1b is nearly located on the biplot, hence positively correlated with both axes F1 and F2, different from when compared to R1a. The three other sub-patterns have a significant overlapping of each other and show a weak subclass fidelity. These gauging stations are near a subclass boundary, so the AHC method will clearly distinct these patterns.

On Figure 6 there is a more evident presentation of the five streamflow sub-patterns in North Macedonia and their redundance on the plot. The dark green and light green areals and dots present the R1a and R1b streamflow sub-patterns. The dark and light blue areals and dots represent R2a and R2b streamflow sub-patterns. The R3 streamflow pattern is depicted with red colour.

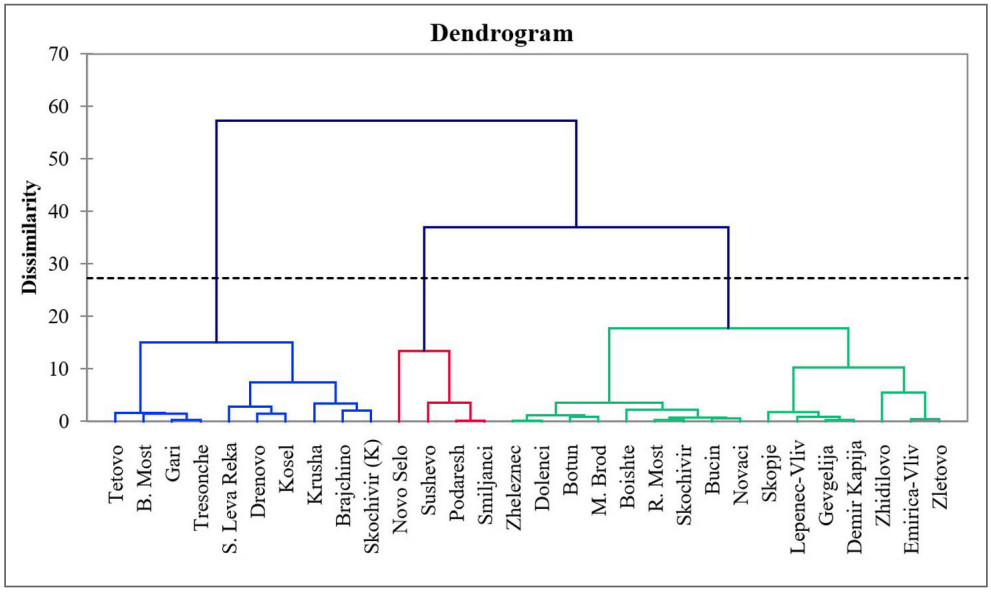


Source: Own data analysis, own design

Figure 6: Principal components analysis (PCA) of Pardè coefficient observations resulted in five subregions marked by concentration ellipses indicating each membership

The result of the hierarchical agglomerative clustering (AHC) shows three statistically significant different patterns below the black dashed "dissimilarity line", which are statistically significant, and clades and leaves shown on the dendrogram (Figure 7). It shows a closer connection between R2 and R3, compared with R1, proven with a smaller dissimilarity measure. Besides the non-significant dissimilarity, we separate *five different streamflow patterns* dividing R1 and R2 on two sub-patterns R1a, R1b (the first and the second blue group on the dendrogram), R2a and R2b (the first and the second green group in the dendrogram). In the third pattern R3, there are no sub-patterns.

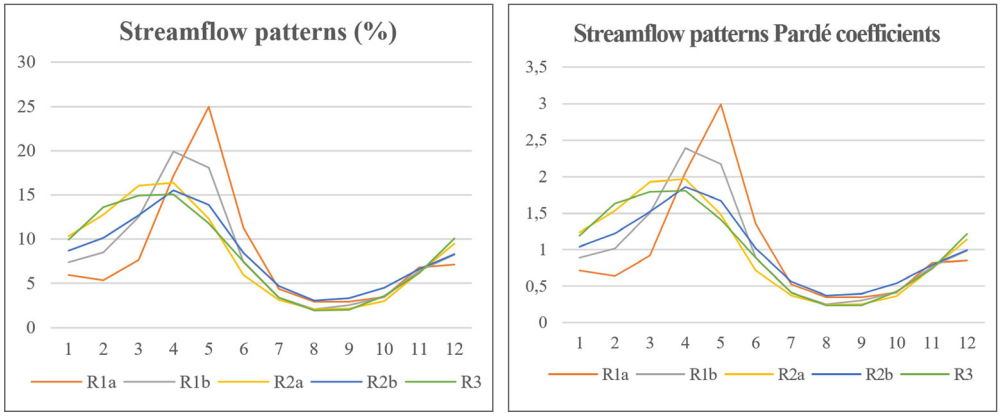
From these five streamflow patterns we made the mean hydrographs (Figure 8a and 8b), showing the streamflow pattern (SP) in percentages and in Pardé's coefficient. The results were mapped with a scope for better spatial review of the streamflow patterns (Figure 9).



Dendrogram order from left to right: blue for R1a and R1b; magenta for R3 and green for R2a and R2b. The dashed black line shows the significant dissimilarity between the three main streamflow patterns R1, R2 and R3

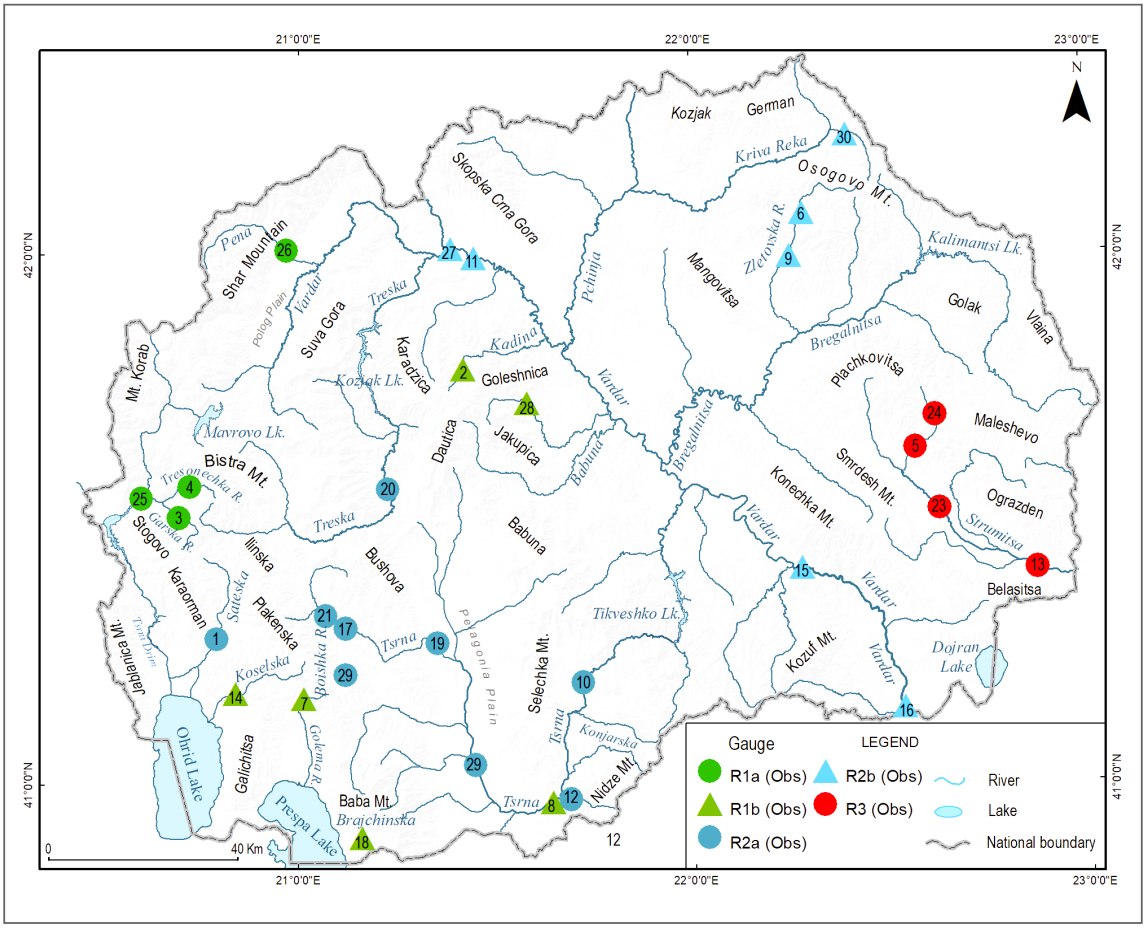
Source: Own data analysis, own design

Figure 7: AHC dendrogram of monthly Pardé coefficients for 30 gauging stations in North Macedonia



Source: Own data analysis, own design

Figure 8a,b: Averaged streamflow patterns in North Macedonia in percent and Pardé coefficients



Source: Own data analysis, own design

Figure 9: Map of different streamflow patterns in North Macedonia

R1a The streamflow pattern is a stable regime without significant differences between the gauging stations. It is located in the north-western part of the country, at the border between the mountain headwaters of the stream and the transitional zone of their catchments in the high mountain areas (around Korab, Šara and Bistra Mountain, mapped on Figure 8), and characterised with more than 25 percent of the annual streamflow (AS) in May, predominantly resulted by a snow melting process and two minimums. The summer minimum is below 5 percent monthly in August and September, and the winter minimum around 5 percent monthly in February, which is logically after a long frost period on these mountain gauging stations.

R1b It is a more variable streamflow pattern, particularly in the late winter and spring. It has a lower peak in spring than R1a, with a percentage of the annual streamflow

higher than 20 percent, occurring in April and May, two minimums, the first more intensive minimum around 3 percent of the annual streamflow in August and September, and a weaker streamflow minimum in January and February. In this pattern, the nival influence is weaker than R1a. This second subregion is located eastward from the R1a, and it covers three mountain streams in the southwestern and central part of North Macedonia (area around the lakes Ohrid and Prespa, Jakupica and Nidze Mountain, mapped on Figure 8).

In both R1 patterns there is an evident secondary winter minimum in January which is 7 percent of the total annual streamflow. The main differences between these two streamflow patterns are a later and weaker maximum of R1b compared with R1a, and an earlier January minimum, so the nival component of R1a is significantly stronger.

R2a This is a very stable defined pattern. It is located on the east of the R1b pattern and covers mainly the basins of Tsrna and Treska Rivers, which are characterised with the largest flat plain areas, as Pelagonia Plain (Figure 8). It has a spring maximum in April and March, in both months above 16 percent of the total annual streamflow, with earlier mountain snowmelt, and a much lower maximum than R1a and R1b. In this case, there is no significant difference between the seasons, and one summer minimum in August/September with around 3 percent per month of the total annual streamflow.

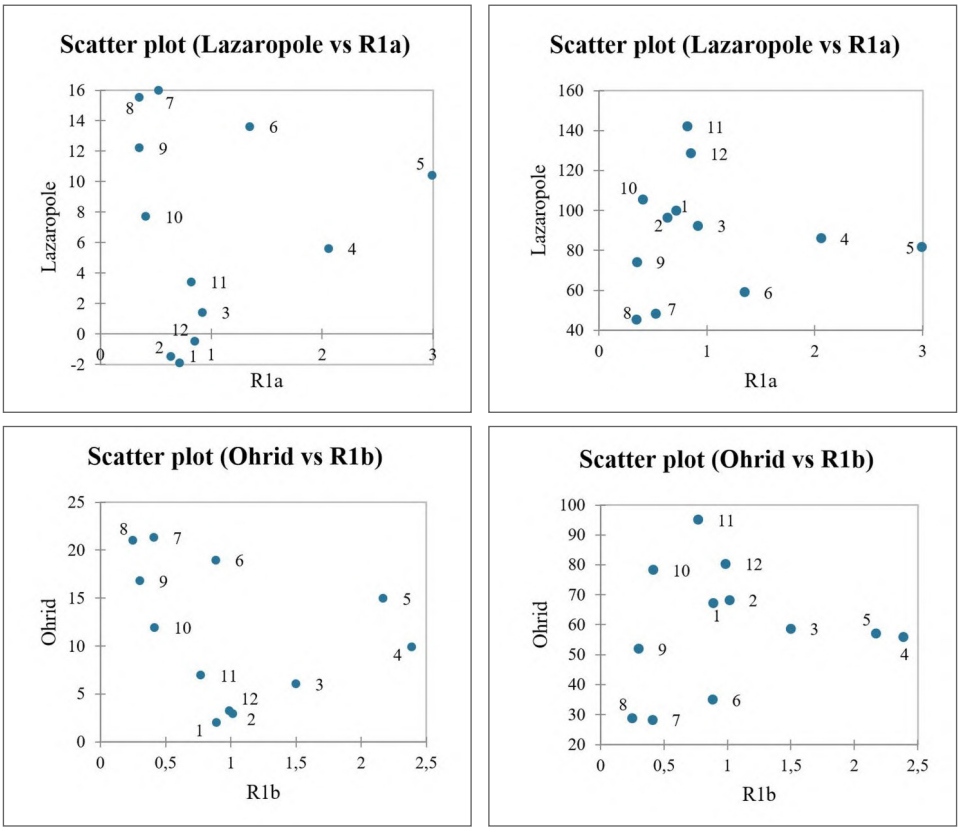
R2b This is a stable pattern, except the variability in April. It has a spring maximum in April with around 15 percent of the total annual streamflow, much lower March values than R2a, with a later mountain snowmelt than $R1_a$, and a much lower maximum. In this case, there is no significant difference between the seasons, and one summer minimum in August with below 5 percent per month of the total annual streamflow. It is located in the middle Vardar section and the Pchinja River and Bregalnica River basins (Figure 8) We have to note that in this streamflow pattern there is a partial anthropogenic influence on several gauging stations, which reflects with higher values of Pardé's coefficient in the summer months. That is the reason why this pattern has an artificially higher percentage of above 7 percent of August and September streamflow amount.

In both R2 streamflow patterns there is no secondary winter minimum.

R3 R3 is a very stable defined pattern, characterised with a spring maximum in April and March, comprising around 15 percent of the total annual streamflow, although without a clear monthly peak, and a highest percentage of the monthly streamflow in February (12 %) and January (10 %). This streamflow pattern indicates the earlier snowmelt and Mediterranean influence and rainfall featuring in the coldest months, without any signs of winter stagnation or streamflow decreasing, – a different situation compared with the western and the central part of North Macedonia. It covers only four gauging stations in the east far part of North Macedonia in the basin of the river Strumitsa, a right tributary of the Struma River in Bulgaria. In this case, there is not a significant difference between the seasons, and one summer minimum in August with around 3 percent per

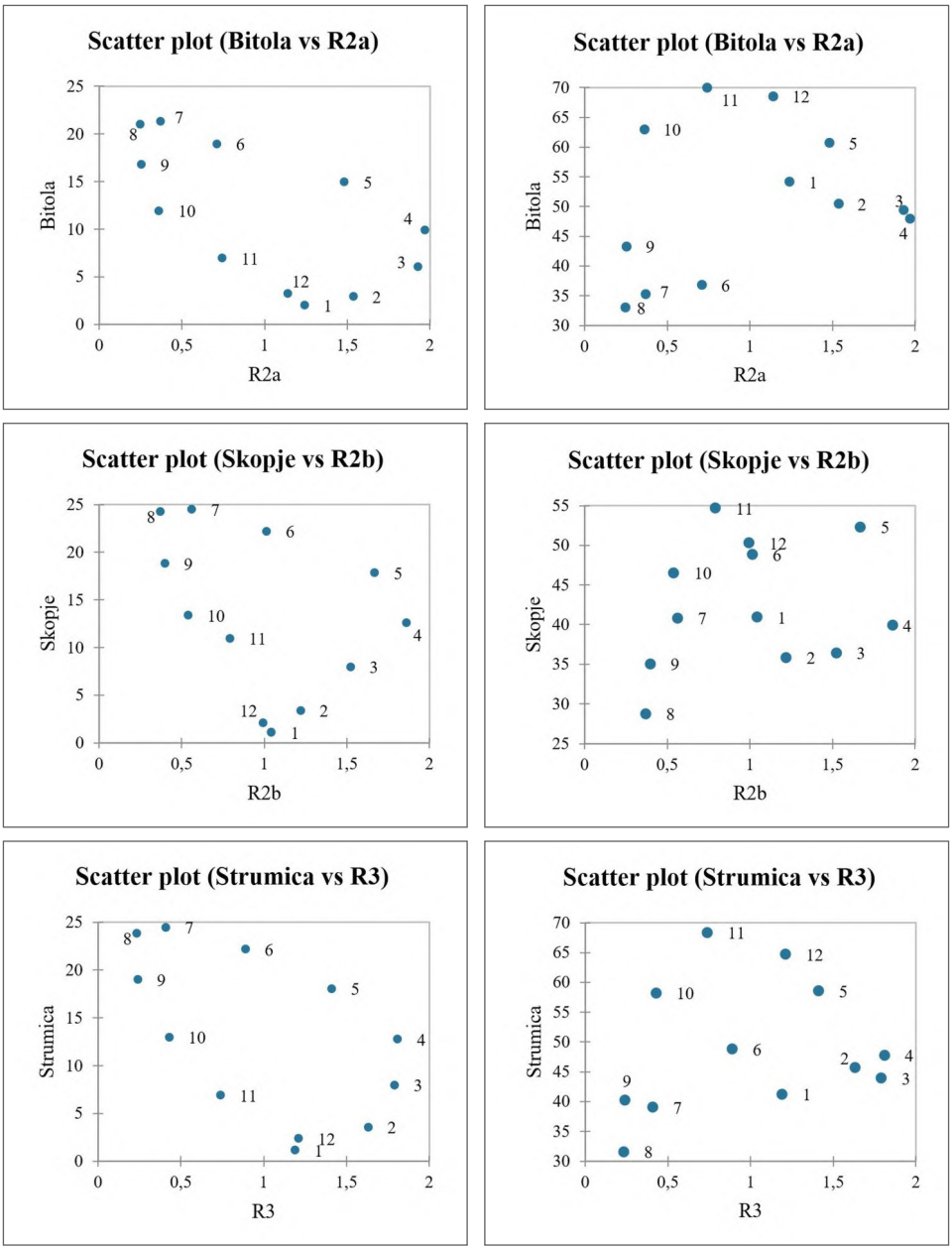
month of the total annual streamflow. The nival component affects R3 less, but the rainfall influence is higher.

With a scope to have a better understanding of the climatic factor influence on the streamflow regime, we plotted each streamflow pattern versus air temperature (the main driver of the snowmelt process) and precipitation, namely in the adequately and centrally located meteorological gauging station among the concrete streamflow pattern. Therefore, for the R1a pattern we used the monthly air temperature and monthly precipitation sums data in Lazaropole gauging station, for R1b – Ohrid gauging station, for R2a – Bitola gauging station, for R2b – Skopje gauging station and for R3 – Strumitsa gauging station. This enabled us to conclude which climatic factor is predominant in each SSP peak flow creation (see Figure 10).



Source: Own data analysis, own design

Figure 10: Five streamflow sub-patterns (SSP) in North Macedonia versus temperature (left column) and precipitation (right column) in five selected gauging stations of the SSPs. The numbers represent the concrete months



Source: Own data analysis, own design

Figure 10 (continuation): Five streamflow sub-patterns (SSP) in North Macedonia versus temperature (left column) and precipitation (right column) in five selected gauging stations of the SSPs. The numbers represent the concrete months

1. R1a. The snowmelt induced by a higher temperature has a stronger influence than the monthly precipitation sum in May (5) peak flow in the R1a streamflow sub-pattern, when the Pardé coefficient (PC) = 3, so the streamflow responds positively on snowmelt. The negative temperature (T) in January and February indicates the secondary winter minimum, when the snowmelt and the rainfall are restricted.
2. R1b. The snowmelt induced by a higher monthly temperature has almost the same influence with the monthly precipitation sum in April (4) peak flow for the R1b streamflow sub-pattern, when the PC = 2.4. In May (5), the snowmelt influence is significantly higher than the precipitation. The January air temperature is close to 0, that's why the secondary minimum is weaker than in R1a.
3. R2a. The snowmelt induced by a higher monthly temperature has almost the same influence with the monthly precipitation sum in the April (4) peak flow in the R2a streamflow sub-pattern, when the PC = 2. In March (3), the precipitation influence is significantly higher than the snowmelt, so the streamflow responds positively on rainfall. Although, the January air temperature is close to 0, the February has a higher value of the Pardè coefficient because of a higher rainfall amount than in both R1 streamflow patterns.
4. R2b. The snowmelt induced by a higher monthly temperature has almost the same influence with the monthly precipitation sum in the April (4) peak flow for the R2b streamflow sub-pattern, when the PC = 1.9. In May (5), the precipitation influence is significantly higher than the snowmelt, so the streamflow responds positively on rainfall. In March (3) the Pardé values are much lower than R2a. Although, the January air temperature is close to 0, the February has a higher value of Pardé because of a higher rainfall amount than in both R1 streamflow patterns.
5. R3. The snowmelt induced by a higher monthly temperature has almost the same influence with the monthly precipitation sum in the April (4) and March (3) peak flow for the R3 streamflow pattern, when the PC = 1.8. In February (5), the precipitation influence is significantly higher than the snowmelt, so the streamflow responds positively on rainfall. February has the highest value of Pardé coefficient (>1.5) from all above mentioned streamflow patterns, induced by the rainfall and the Mediterranean influence.

Overall, going eastward, the streamflow due to snowmelt process declines, but the precipitation influence is increasing. Otherwise, in the same direction, the month with a peak flow is going backward, from the highest values in May for R1a, towards high values in April, March and even closer to the highest Pardé values in February for the R3 streamflow sub-pattern.

5 Conclusion

The study already mentioned at the beginning of this article about river regimes in Bulgaria (Hristova 2007) shows compatible results in the border region between North Macedonia and Bulgaria. It is covered by streamflow pattern R3 in this study and the "transitional water regime zone". Both regions have maximum streamflow in March and April, but,

differently, the minimum occurs in North Macedonia in August/September, while in Bulgaria the minimum is in June and July, because of earlier autumn rainfall on the high Pirin and Rila mountains.

The research on the Balkan hydrological regimes of Stanescu (2004) includes an analysis of North Macedonia, but only used data from five gauging stations. According to his results, there are three river regime types: Western upland mountain zone (maximum in May and minimum in August), which is corresponding to the R1a sub-pattern in general. The second regime "Western-Eastern and Cental mountain and hilly plateau" has a maximum from February to April and a minimum streamflow usually occurring from July to September. Spatially this pattern is very disseminated, different from our study where in this area there are three different sub-patterns (R2a and R2b and partially R3). Probably this difference is because of the low density of stations in the previous study about Balkan river rigimes. Elsewise, the author used graphical comparison, not one of the above mentioned automatic methodologies (PCA, AHC, SOM etc).

The performed analysis on 30 streamflow gauging stations resulted with a systematic countrywide classification of the streamflow patterns in North Macedonia. The monthly streamflow variability was higher in the spring months, while lower during the summer period. Otherwise, the controlling factors average air temperature and precipitation were plotted versus Pardè coefficients.

Our classification generated five sub-patterns (R1a, R1b, R2a, R2b, and R3), nested in the three main streamflow patterns (R1, R2, and R3), and dominated by rainfall, snowmelt or both. Their spatial distribution is correctly located from the west to east. The month of a maximum streamflow is changing backward from May in the western part of North Macedonia to March and February on the southeastern part of the country. The second (winter) minimum, which occurred in February, was clearly evident only in the R1a and partially in R1b streamflow patterns, where air temperatures are low during the winter months and there is no input of snowmelt water in these streams.

The obtained sub-patterns show increasing rainfall predominance going eastward, and decreasing snowmelt influence also going eastward from the streamflow patterns R1a to R3. The snowmelt dominance was identified in the high mountain region westward in the spring area of the rivers Vardar and Radika, while the rainfall predominance was crucial in the Strumitsa River catchment. Besides the relatively small territory of the country, the results are quite robust and different patterns show a clear distinction between each other. It could help authorities in recognising the key possible maximum streamflow drivers and help focusing expectations on hydrological changes in North Macedonia and wider on the Balkan Peninsula.

6 References

Assani A. A., Tardif S., Lajoie F. (2006): Statistical Analysis of Factors Affecting the Spatial Variability of Annual Minimum Flow Characteristics in a Cold Temperate Continental Region (Southern Québec, Canada). In: Journal of Hydrology, 328 (3–4), pp. 753–763. – *https://doi.org/10.1016/j.jhydrol.2006.01.015*.

Blöschl G., Sivapalan M. (1995): Scale Issues in Hydrological Modelling: A Review. In: Hydrological Processes, 9 (3–4), pp. 251–290. – *https://doi.org/10.1002/hyp.3360090305.*

Bower D., Hannah D. M., Mcgregor G. R. (2004): Techniques for Assessing the Climatic Sensitivity of River Flow Regimes. In: Hydrological Processes, 18 (13), pp. 2515–2543. – *https://doi.org/10.1002/hyp.1479*.

Curran J. H., Biles F. E. (2021): Identification of Seasonal Streamflow Regimes and Streamflow Drivers for Daily and Peak Flows in Alaska. In: Water Resources Research, 57 (2), article e2020WR028425. – *https://doi.org/10.1029/2020WR028425*.

Di Prinzio M., Castellarin A., Toth E. (2011): Data-driven Catchment Classification: Application to the Pub Problem. In: Hydrology and Earth System Sciences, 15 (6), pp. 1921–1935. – *https://doi.org/10.5194/hess-15-1921-2011*.

Falkenmark M., Rockström J. (2004): Balancing Water for Humans and Nature: The New Approach in Ecohydrology. London / Sterling, VA: Earthscan.

Haines A. T., Finlayson B. L., Mcmahon T.A. (1988): A Global Classification of River Regimes. In: Applied Geography, 8 (4), pp. 255–272. – *https://doi.org/10.1016/0143-6228(88)90035-5*.

Hristova N. (2007): Geographical Specificity of the River's Regime in Bulgaria. In: Journal of the Geographical Institute Jovan Cvijic of the Serbian Academy of Science and Arts (SASA). Collection of Papers, 57 (1), pp. 71–77. – *https://doiserbia.nb.rs/img/doi/0350-7599/2007/0350-75990757071H.pdf*.

Kahya E., Kalayci S., Piechota T. C. (2008): Streamflow Regionalization: Case Study of Turkey. In: Journal of Hydrologic Engineering, 13 (4), pp. 205–214. – *https://doi.org/10.1061/(ASCE)1084-0699(2008)13:4(205)*.

Kalayci S., Kahya E. (2006): Assessment of Streamflow Variability Modes in Turkey: 1964–1994. In: Journal of Hydrology, 324 (1–4), pp. 163–177. – *https://doi.org/10.1016/j.jhydrol.2005. 10.002*.

Kendrovski V., Spasenovska M. (2011): The Effects on Health of Climate Change in the Republic of Macedonia. Skopje: Ministry of Health of the Republic of Macedonia. – *https://www.researchgate.net/publication/270889739_the_effects_on_health_of_climate_change_in_the_republic_of_macedonia*.

Kennard M. J., Mackay S. J., Pusey B. J., Olden J. D., Marsh N. (2010): Quantifying Uncertainty in Estimation of Hydrologic Metrics for Ecohydrological Studies. In: River Research and Applications, 26 (2), pp. 137–156. – *https://doi.org/10.1002/rra.1249*.

Krasovskaia I. (1995): Quantification of the Stability of River Flow Regimes. In: Hydrological Science Journal, 40 (5), pp. 587–598. – *https://doi.org/10.1080/02626669509491446*.

Laaha G., Blöschl G. (2006): A Comparison of Low Flow Regionalisation Methods – Catchment Grouping. In: Journal of Hydrology, 323 (1–4), pp. 193–214. – *https://doi.org/10.1016/j.jhydrol.2005.09.001*.

Masiokas M. H., Cara L., Villalba R., Pitte P., Luckman B. H., Toum E., Christie D. A., Le Quesne C., Mauget S. (2019): Streamflow Variations across the Andes (18°–55° S) during the Instrumental Era. In: Scientific reports, 9 (article No. 17879). – *https://doi.org/10.1038/s41598-019-53981-x*.

Morán-Tejeda E., López-Moreno J. I., Ceballos-Barbancho A., Vicente-Serrano S. M. (2011): River Regimes and Recent Hydrological Changes in the Duero Basin (Spain). In: Journal of Hydrology, 404 (3–4), pp. 241–258. – *https://doi.org/10.1016/j.jhydrol.2011.04.034*.

Olden J. D., Kennard M. J, Pusey B. J. (2012): A Framework for Hydrologic Classification with a Review of Methodologies and Applications in Ecohydrology. In: Ecohydrology, 5 (4), pp. 503–518. – *https://doi.org/10.1002/eco.251*.

Olden J. D., Liermann C. A. R., Pusey B. J., Kennard M. J. (2009): Protocols for Hydrologic Classification and a Review of Australian Applications. In: Pusey B. J., Kennard M.,

Hutchinson M., Sheldon F. (eds.): Ecohydrological Regionalisation of Australia: A Tool for Management and Science. Canberra: Land & Water Australia, pp. 17–44 (= Appendix 2). – *https://openresearch-repository.anu.edu.au/bitstream/1885/136163/5/pusey-ecohydrological-regionalization.pdf*.

Olden J. D., Poff N. L. (2003): Redundancy and the Choice of Hydrologic Indices for Characterizing Streamflow Regimes. In: River Research and Applications, 19 (2), pp. 101–121. – *https://doi.org/10.1002/rra.700*.

Pardé M. (1955): Fleuves et rivières. Paris: Collection Armand Colin [in French].

Radevski I., Gorin S., Taleska M., Dimitrovska O. (2018): Natural Regime of Streamflow Trends in Macedonia. In: Geografie, Journal of the Czech Geographical Society, 123 (1), pp. 1–20. – *https://doi.org/10.37040/geografie2018123010001*.

Razavi T., Coulibaly P. (2013): Classification of Ontario Watersheds Based on Physical Attributes and Streamflow Series. In: Journal of Hydrology, 493, pp. 81–94. – *https://doi.org/10.1016/j.jhydrol.2013.04.013*.

Shapiro S. S., Wilk M. B. (1965): An Analysis of Variance Test for Normality (Complete Samples). In: Biometrika, 52 (3–4), pp. 591–611. – *https://doi.org/10.2307/2333709*.

Snelder T. H., Lamouroux N., Leathwick J. R., Pella H., Sauquet E., Shankar U. (2009): Predictive Mapping of the Natural Flow Regimes of France. In: Journal of Hydrology, 373 (1–2, pp. 57-67. – *https://doi.org/10.1016/j.jhydrol.2009.04.011*.

Stahl K. (2001): Hydrological Drought: A Study across Europe. PhD Dissertation, Albert Ludwig University, Freiburg, Germany. – *https://freidok.uni-freiburg.de/data/202*.

Stanescu V. A. (2004): Hydrological Regimes in BALWOIS Area. In: Morell M., Todorovik O. (eds.): BALWOIS. Conference on Water Observation and Information Systems for Decision Support. Proceedings. Skopje: Ministry of Education and Science of Republic of Macedonia, pp. 45–65.