ISSN 0354-8724 (hard copy) | ISSN 1820-7138 (online)

Climatic Regionalization of Montenegro by Applying Different Methods of Cluster Analysis

Dragan Burić[A*], Jovan Mihajlović[B], Vladan Ducić[B]

Received: April 02, 2023 | Revised: June 08, 2023 | Accepted: June 18, 2023
doi: 10.5937/gp27-43776

Abstract

To carry out an "objective" regionalization of the climate of Montenegro for the period 1961–2020, this paper used cluster analysis, which is a multivariate technique that classifies a sample of subjects (objects) based on a set of variables into a single number. Based on the results (score), several groups were separated, and similar classes (groups) were grouped into the same cluster. Annual data for mean temperature and total precipitation from 18 meteorological stations were utilized. Temperature and precipitation cluster regions were separated using three different hierarchical agglomerative methods (*Unweighted Pair Group Method with Arithmetic Mean (UPGMA)*, *Single linkage*, and *Ward's*) and one non–hierarchical method (*K–means*). The *Euclidean distance* was used as a measure of distance for hierarchical methods, and the results were represented graphically in the form of dendrograms and thematic maps. The obtained results indicate that the singled–out temperature and precipitation cluster regions largely coincide with the established climate types in Montenegro. The cluster results further showed that the distribution of meteorological stations clearly reflects the largest part of the climatic diversity of Montenegro and indicates the spatial dimension of temperature and precipitation.

Keywords: cluster regions *UPGMA, Single linkage, Ward's, K–means,* Montenegro

Introduction

Cluster analysis is a multivariate technique that classifies a certain sample of subjects (objects) into different groups based on a set of measurable variables and their results (scores), such that similar classes (groups) form the same cluster. There are several reasons for using cluster analysis in statistical data analysis. In addition to uncovering/suggesting hidden structures in the dataset, the primary purpose is reduction, specifically reducing the number of input data, which is crucial for classification purposes. Investigating the climate regionalization of the continental United States (US) based on multiyear monthly values of temperature and precipitation, Fovell and Fovell (1993) reported that solutions of 14, 25, and 8 cluster levels were the best chosen. Together these clusters are called "*reference clusterings*". At the 14–cluster level, most of the US is divided into four major climate zones: the Southeast, the Central East, the Northeast, and the inner West cluster zone. Previously, this cluster analysis was less frequently used (Wolter, 1987), but there is a growing number of scientific articles in atmospheric science journals that use this technique (Fovell & Fovell, 1993). Clustering of Iran, based on the use of Ward's method and 10 climate components, for the period from 1980 to 2005 based on 64 meteorological stations (MS), distinguishes 4 basic cluster regions (Arbabi, 2011). It can be concluded that the use of multivariate analysis methods (such as factor analysis, *principal component analysis*

[A] Department of Geography, Faculty of Philosophy, University of Montenegro, Danila Bojovica bb, 81400 Nikšić, Montenegro; buric.d@ucg.ac.me
[B] Faculty of Geography, University of Belgrade, Studentski trg 3/3, 11000 Belgrade, Serbia; millennijum@hotmail.com; vladanducic@yahoo.co
* Corresponding author: Dragan Burić, e-mail: buric.d@ucg.ac.me

(*PCA*), cluster analysis, etc.) is very common in modern climate research, because they are used to reduce the amount of climate variables, in order to obtain new variables through their combination and to use the results for subsequent analyses (Appendini et al., 1994; DeGaetano et al., 1990; Johnson, 1998; Moreira et al., 2006; Nassiri et al., 2006; Szép et al., 2005; Van Groenewoud, 1984; Warrington, 1977).

To date, there has been no application of different methods of the multivariate cluster analysis technique in the regionalization of the climate of Montenegro. Three hierarchical and one non–hierarchical method will be used to establish a theoretical model of the distribution of temperature and precipitation cluster regions. Since it is more difficult to find an appropriate pattern for grouping the observed clusters if more variables are included, in this case, only the two most important climatic elements, temperature, and precipitation, will be observed. The disadvantage of this type of multivariate technique is the problem of determining the final number of clusters or the so–called "stopping" rule. There is no statistical criterion or objective standard procedure for this determination. During each object grouping step, a distance plot is generated. However, ultimately, the decision is made by the researcher. There are two equally subjective approaches to determining the final number of clusters: formal tests and the heuristic approach. The first approach refers to the possibility of interpreting the obtained solution, while the second approach concerns the analysis of fusion coefficients (Papić–Blagojević & Bugar, 2009). In this research paper, a strategy of using hierarchical agglomerative methods was initially applied to gain insight into the number of obtained clusters, and then the results were applied in a non–hierarchical method. The idea is to apply several hierarchical and one non–hierarchical method to determine the complementarity of the obtained results. If they agree to a reasonable extent, the end result will be considered convincing.

This study has two main objectives. The first is to group precipitation and temperature regions into a specific, smaller, and "reasonable" number of clusters, based on which the "objective" theoretical assumptions of climate regions in Montenegro will be established. It will suggest the relationships that exist between the obtained cluster regions in the observed period and how the results of different methods of cluster analysis affect the already known spatial distribution of temperature and precipitation, conditioned by the dominant physical–geographical factors in the area. According to the Köppen climate classification (*KCC*), Burić et al. (2014) distinguished two basic climates, three climate types, and five climate subtypes in Montenegro for the period 1961–1990, using data from 23 meteorological stations. The second goal of this study is to perform a comparative analysis with the previously mentioned climate regionalization in Montenegro. Therefore, in the case of cluster analysis, the regions are distinguished by means of joining algorithms, while in the case of climate classification, the regions are based on physical characteristics, such as the influence of physical–geographical factors on the regime of temperature and precipitation.

Research area, database and methodology

Research area

The area of research is Montenegro, a country located in Southeastern Europe and which extends to the southernmost part of the Adriatic Sea in a length of about 100 km. This small Mediterranean country (area 13,812 km^2) stretches between 41°50'–43°50'N, i.e. between 18°26'–20°21'E. Apart from the mathematical–geographical position, the main factors that influence the climate of this country are: relief, air mass variations and the influence of the Mediterranean Sea (Burić et al., 2013). The significant relief dissection has made it possible to distinguish several types and subtypes of climate in this small area. Montenegro is predominantly mountainous, with numerous valleys and a terrain that extends from 0 m a.s.l. to 2534 m a.s.l.

Database and methodology

To determine the cluster regions for temperature and precipitation in Montenegro, data from a 60–year period (1961–2020) were used from 18 meteorological stations. This data included mean annual values of air temperature and annual precipitation sums, as shown in Figure 1. The homogeneity of the temperature and precipitation data was tested using two software packages, *MASH v3.02* and *MISH v1.02*, which are recommended by the World Meteorological Organization (WMO) and developed by Szentimrey (2003) and Szentimrey and Bihari (2007).

While applying cluster analysis, it is very important to carry out grouping analysis, i.e., dispersion of MS included in the analysis. Therefore, the Point Pattern Analysis–nearest neighbors procedure was used in Past4.12 software. The procedure tests overdispersion of points as two–dimensional (2D) coordinate values (Davis, 2002). The calculation of this statistic is based

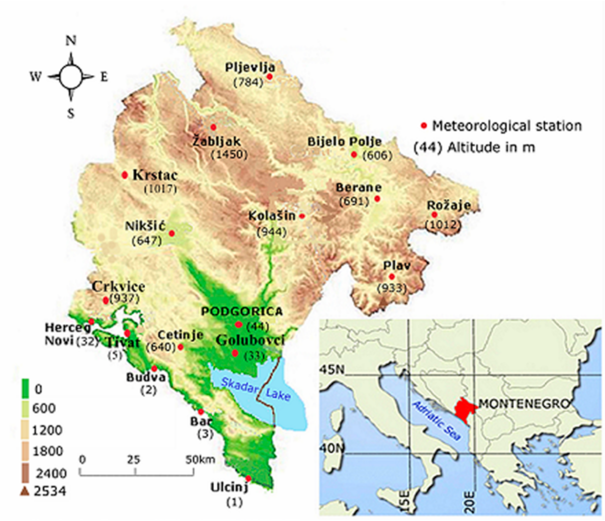

Figure 1. Presentation of meteorological stations used in the analysis
Source: Burić & Doderović, 2022

on nearest neighborhood analysis[1]. The null hypothesis (h_0) is a random Poisson process, giving a modified nearest–neighbor exponential distribution with mean:

$$\eta = \frac{\sqrt{R/t}}{2} \tag{1}$$

- where R is the area and t is the number of points.

The probability that the distribution is random (a Poisson process, given as an exponential distribution) is represented by the value of P when:

$$P = \frac{\overline{w}}{\eta} = \frac{2\overline{w}}{\sqrt{R/t}} \tag{2}$$

- where P is the nearest neighbor value, and $\overline{w}$ is the observed mean distance between nearest neighbours. Table 1 shows the value of the coefficient P in the Poisson distribution.

Table 1. Values of coefficient P offered by PAST

P	Distribution
<1	Clustered points
~1	Poisson patterns
>1	Overdispersed points

The theoretical distribution under the null (h_0) hypothesis is plotted as a continuous curve with a histogram of observed distances. The expected probability density function as a function of distance r is (Clark & Evans, 1954):

$$g(r) = 2\rho\pi r \exp(-\rho\pi r^2) \tag{3}$$

- where ρ, that is, r is point density.

The obtained results, shown on the x/y graph, indicate that the points are not clustered in an area (Figure 2). The value of the coefficient R in this case is 1.39, which means that the points (meteorological stations) are scattered or dispersed in an area, indicating a statistically significant overdispersion of points.

Temperature and precipitation cluster regions were separated using three different hierarchical agglom-

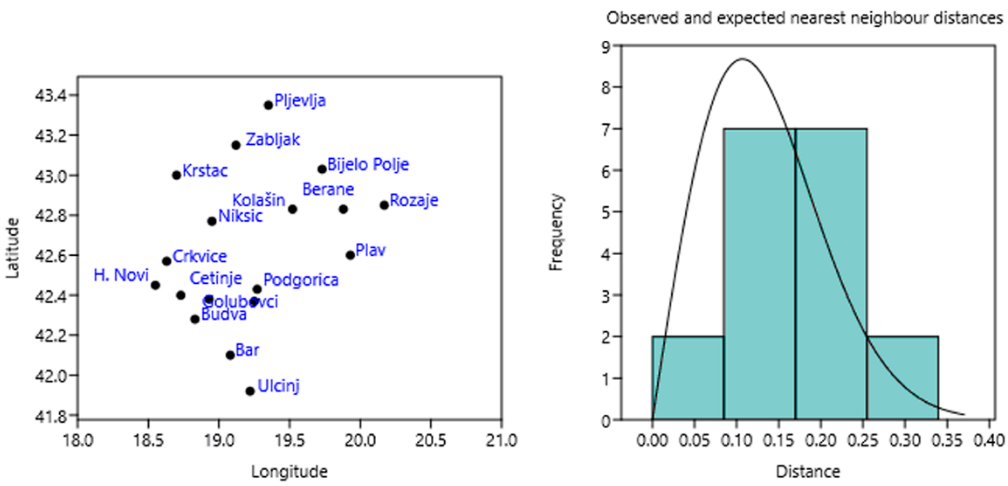

Figure 2. Display of meteorological stations on the XY graph in *PAST* (left) and histogram of observed and expected distances between nearest neighbors (right)

[1] https://www.nhm.uio.no/english/research/resources/past/downloads/past4manual.pdf

erative methods and one non–hierarchical (*K–means*) method. The *Euclidean distance* was used as a measure of distance for hierarchical methods[2], which is a robust and widely applicable measure. When a measure such as the *Euclidean distance* is used, it should be decided before the actual clustering whether the variables will be transformed, or whether the original values will be kept. Standardization of the results is done if the variables used were measured on different measurement scales[3]. *Euclidean distance* is converted to similarity by changing the sign:

$$Euclidean\ dist._{jk} = \sqrt{\sum_{i=1}^{s}(x_{ij} - x_{ik})^2} \quad (4)$$

The choice of variables (in this case, temperature and precipitation) included in the cluster analysis must be determined based on the assumed conceptual model since the analysis itself does not distinguish important from irrelevant variables, and this can greatly affect the final result. Hierarchical agglomerative methods are those in which the variables begin the clustering in their own separate clusters. Then, the two most similar clusters are grouped together, and this is repeated until all variables are in one cluster. The optimal number of clusters is obtained from all cluster solutions. The hierarchical cluster routine produces a "dendrogram" that shows how the data points (rows) can be grouped into clusters. For this analysis, three different algorithms were used by selecting the appropriate options in the software *Past4.12* (Table 2): *UPGMA* (**U**nweighted **P**air **G**roup **M**ethod with **A**rithmetic Mean), *Single linkage and Ward's method*, as well as the non–hierarchical clustering method (*K –means*).

The set of symbols used is as follows: Let X_{ijk} the value for variable **k** in observation **j** belonging to cluster **i**. Furthermore, for this particular method it must be defined like this (The Pennsylvania State University, 2004):

$$ESS = \sum_i \sum_j \sum_k \left| X_{ijk} - \overline{X}_{i\cdot k} \right|^2 \quad (5)$$

- where *ESS* is the Error Sum of Squares;

$$TSS = \sum_i \sum_j \sum_k \left| X_{ijk} - \overline{X}_{\cdot\cdot k} \right|^2 \quad (6)$$

- where *TSS* is the Total Sum of Squares;

$$r^2 = \frac{TSS - ESS}{TSS} \quad (7)$$

The value r^2 is interpreted as the proportion of variation that is explained by a particular clustering of observations. Referring to the explanations given in the *PAST* manual (*V.4.12*) (https://www.nhm.uio.no/english/research/resources/past/downloads/past4manual.pdf), one method is not necessarily better than the other, although some do not recommend *Single linkage*, which can be useful in comparing dendrograms obtained by different algorithms to informally determine clustering strength.

K–means clustering[4] is a non–hierarchical cluster method that was first mentioned under this name in 1967 (MacQueen, 1967). The number of clusters to be used is predetermined, usually according to some hypothesis such as the existence of two meteorological variables (temperature and precipitation), 4 climate regions, or three types of objects in the data set. It is sometimes preferred because it allows subjects to move from one cluster to another (this is not possible in hierarchical cluster analysis where a subject, once placed, cannot move to another cluster) (Everitt et al., 2001; Rencher, 2002). Today, the well–known algo-

Table 2. Joining algorithms used in the analysis

UPGMA	Single linkage	Ward's method
Unweighted pair–group average (UPGMA). Clusters are joined based on the average distance between all members in neighboring clusters.	Single linkage (nearest neighbor). Clusters are joined based on the smallest distance between 2 groups. This method is relatively simple, but it is better than other methods when natural clusters are not spherical or elliptical in shape.	Clusters are joined so that within–group variance growth is minimized. It is also very sensitive to outliers. Nevertheless, this is one of the most popular methods, along with the average link between groups method. This method includes an agglomerative cluster algorithm. Ward's method* starts with *n* clusters of size 1, continues until all observations are included in one cluster. This method is best suited for quantitative variables, but not for binary ones.

Source: Everitt et al., 2001; Rencher, 2002

* https://online.stat.psu.edu/stat505/lesson/14/14.7

2 For *Ward's* method, the *Euclidean distance* is inherent to that algorithm. (https://www.nhm.uio.no/english/research/resources/past/downloads/past4manual.pdf)
3 http://www.statstutor.ac.uk/resources/uploaded/clusteranalysis.pdf
4 https://www.nhm.uio.no/english/research/resources/past/downloads/past4manual.pdf

rithm using *K–means* clustering has been defined by many authors (Bensmail et al., 1999; Fraley & Raftery, 1998; Hartigan & Wong, 1979):

Given a set of observations $X = (x_1,x_2,...,x_n)$, partition the n observations into k partition $S = \{s_1,s_2,...,s_k\}$, such that:

$$\arg\min \sum_{i=1}^{k} \sum_{x \in S_i} \|x - \mu_i\|^2 \quad (8)$$

The digitization of the temperature and precipitation cluster regions obtained using the aforementioned methods of cluster analysis was carried out using Voronoi diagrams. The appropriate options were run in the *QGIS 2.8.1* package to produce the results shown in Figure 3.

Voronoi diagrams were first mentioned in the 17th century when René Descartes argued that the solar system consists of vortices, whose decay produces convex regions rotating around fixed stars, in the field of computational geometry. Their next use was by the mathematician Dirichlet in 1850, and Voronoi gave them a wider meaning in 1907, 1908, and 1909. They have different names and uses in various scientific disciplines, such as the transformation of the mean axis in biology and physiology, *Wigner–Seitz* zones in chemistry and physics, domains of action in crystallography, and Thyssen polygons in meteorology and geography. A Voronoi diagram is also known as a Dirichlet square plate. The cells are called Dirichlet regions or Voronoi diagrams, Delaunay tessellation or Delaunay triangulation (Aurenhammer & Klein, 2000; Barber et al., 1996; Guibas & Stolfi, 1985; Okabe et al., 2000; Preparata & Shamos, 1985).

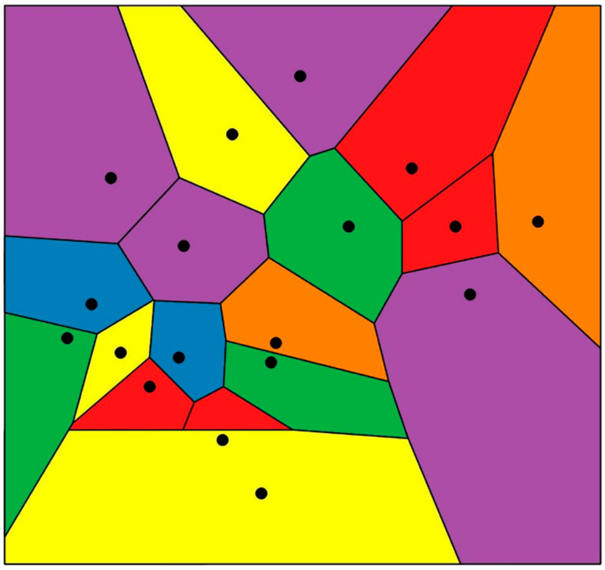

Figure 3. Voronoi diagram of 18 points (meteorological stations included in the analysis) in the *Euclidean* plane

Results and discussion

First of all, it should be emphasized that understanding the term “climate cluster” requires a good (or excellent) knowledge of the climate diversity in Montenegro. Only with such knowledge can high–quality theoretical propositions for an adequate climate regionalization of this country be presented. In order to implement such an analysis, hierarchical agglomerative methods that are dominantly used in research (*UPGMA*, *Single linkage* and *Ward’s* method) and non–hierarchical clustering method (*K–means*) were applied. The standard *Euclidean distance* for hierarchical agglomerative methods was chosen as a measure of similarity/distance. The two most important variables of the climate system, average annual values of temperature and precipitation, were observed for a series of 60 years (1961–2020) from 18 meteorological stations in Montenegro.

Applying the above–mentioned hierarchical methods and *Euclidean distance*, by running the appropriate options in the *PAST* software, 6 dendrograms were obtained. In this particular case, it means that the variability of temperature and precipitation has its own spatial dimension, as does every variable of the climate system, and this influenced the determination of the distance, that is, the separation of temperature and precipitation regions. Based on the non–hierarchical *K–means* method (the default number of clusters is 3), appropriate clusters were formed, as in the previous case. Based on the grouping of MS and by running the appropriate options in the *QGIS* software, a total of 8 thematic maps were obtained on which the digitized regions were previously defined by clusters using the method of Voronoi diagrams.

Hierarchical agglomerative cluster methods for temperature (UPGMA, Single linkage i Ward)

According to the *UPGMA* method, 8 cluster temperature regions were distinguished. **The first cluster** consists of 4 MS located in the mountainous southwestern and northeastern parts of Montenegro, mostly at altitudes of 600–700 m. According to Burić et al. (2014) these are places with a modified Mediterranean climate, dry and hot summer (*Csb*). This climate cluster accounts for 22.2% of the total number of MS included in the analysis. **The second cluster** included 3 MS, mostly located in the altitude zone between 800–900 m. According to the Köppen climate classification, it is a belt with a moderately warm climate (*C*), which is represented by two subtypes of climate: a moderately warm and humid climate with hot summers in the far north of Montenegro (*Cfb*) and a transitional variant of the Etesian climate (*Csb*). Both regions are mountainous and should be considered transitional between Mediterranean and moderately warm and humid climates. This cluster makes up 16.7% of stations. Only one MS (Nikšić) is classified in **the third cluster**, and it is located at an altitude of about 650 m and has a *Csb* subtype of climate. Therefore, MS classified in the three mentioned clusters have moderately warm summers, while winters are due to the influence of altitude and distance from the sea.

The fourth cluster consists of 2 MS (11.1%), located at an altitude of 944 m and 1012 m. Burić et al. (2014) just mention an altitude of about 1000 m as a transition between a moderately warm (*C*) and a mountainous or moderately cold (*D*) climate. At higher altitudes, the *D* climate dominates, so **the fifth cluster** includes only one MS (Žabljak), the only one at an altitude of 1450 m. It is a humid boreal climate with fresh summer (*Dfc* climate subtype). The southern part of Montenegro includes the Adriatic coast and the plain area around the capital (Podgorica). In this part, all MS are located at a low altitude (1–44 m) and are under the strong influence of the Adriatic (Mediterranean Sea), with a typical Mediterranean climate, characterized by dry, sunny and hot summers, while winters are rainy and mild (*Csa* climate subtype). Nevertheless, the *UPGMA* method separates three clusters (Figure 4) with: 3 MS (**sixth cluster**, 16.7%) and with 2 MS each (**seventh and eighth cluster**, with 11.1% each).

Similar results were obtained using the other two methods (*Single linkage and Ward's method*), which confirms the fact that temperature has its own spatial dimension and that there is a correlation between the hierarchical methods in cluster analysis. It could be concluded that this type of multivariate techniques is suitable for research in climatology. Comparing the results of cluster analysis for temperature obtained by *UPGMA* and *Single linkage*, i.e. *Ward's method*, the following differences are observed:

- Compared to *UPGMA*, there are 7 cluster regions in the *Single linkage* method.
- The *Single linkage* method classifies 5 MS in one cluster: Herceg Novi, Ulcinj, Podgorica, Bar and Budva (Figure 5), making it the largest cluster region in percentage terms (27.8%), while according to *UPGMA* the mentioned MS are classified into 2 cluster regions.
- Compared to the *UPGMA* method, there are also 7 cluster regions in the *Ward's method*.
- According to the *Ward's method*, MS Cetinje and Nikšić now constitute a separate cluster region with 11.1%. Also, MS Kolašin, Rožaje and Žabljak form one cluster region with 16.7% of the total number of stations (Figure 6). As a reminder, according to the *UPGMA* method, MS Nikšić and MS Žabljak each form one cluster region.
- In contrast to the results obtained using the *UPGMA* and *Single linkage* methods, in the *Ward's method* there is no cluster of regions consisting of only one MS.
- Percentage–wise, the largest cluster region obtained by the *UPGMA* method includes 22.2% of stations, by the *Single linkage* method 27.8%, while with the *Ward's method*, four regions make up 16.7% of the total number of stations.

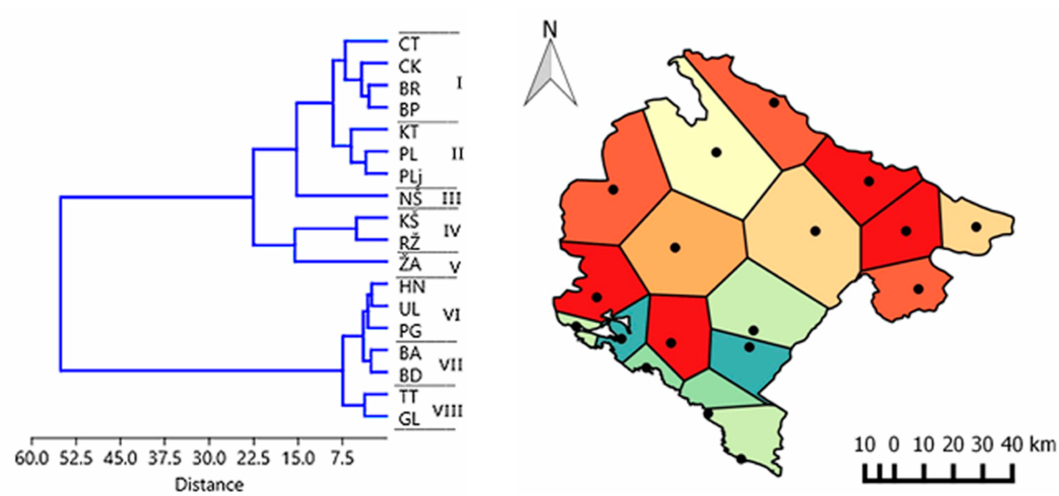

Figure 4. Results of the *UPGMA* method for the period 1961–2020: dendrogram of meteorological stations (left) and temperature regions (right).
(CT – Cetinje, CK – Crkvice, BR – Berane, BP – Bijelo Polje, KT – Krstac, PL – Plav, PLj – Pljevlja, NŠ – Nikšić, KŠ – Kolašin, RŽ – Rožaje, ŽA – Žabljak, HN – Herceg Novi, UL – Ulcinj, PG – Podgorica, BA – Bar, BD – Budva, TT – Tivat, GL – Golubovci)

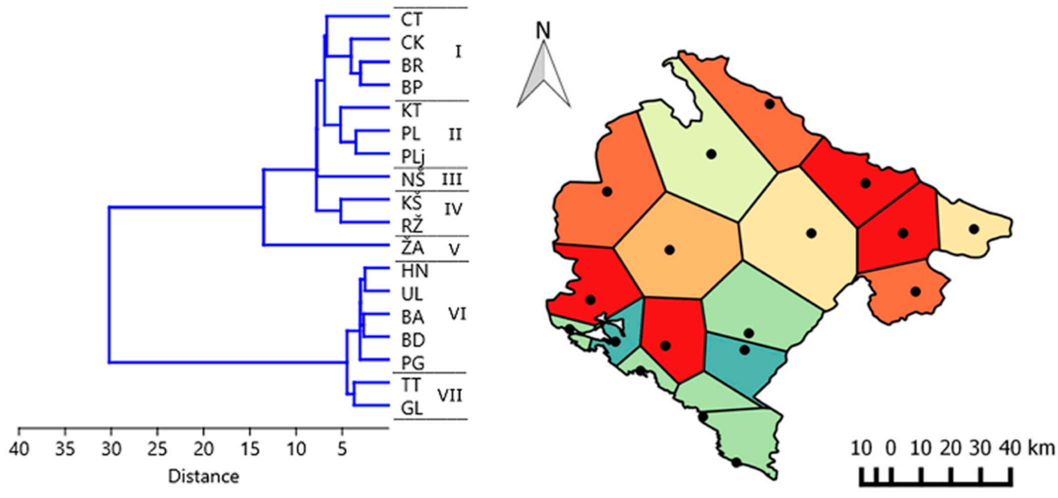

Figure 5. Results of the *Single linkage* method for the period 1961–2020: dendrogram of meteorological stations (left) and temperature regions (right)

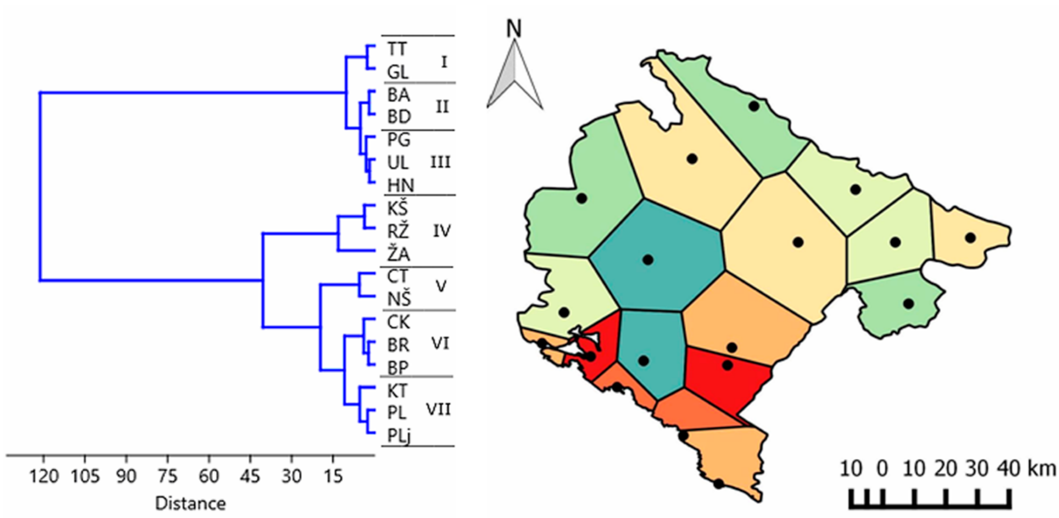

Figure 6. *Ward's method* results for the period 1961–2020: dendrogram of meteorological stations (left) and temperature regions (right)

A non–hierarchical K–means cluster method for temperature

Unlike the previous hierarchical agglomerative methods, the non–hierarchical *K–means* cluster procedure does not produce a dendrogram as a graphical representation of the MS included in the analysis. The default number of clusters in this case is 3, and the algorithm connects the stations to the cluster with the smallest distance to its centroid. The given number of clusters in this case will facilitate the analysis because only similar cluster regions will be observed in all three cases, regardless of the numerical designation of the stations that make it up.

The results of the analysis indicate the following:

I temperature cluster region groups the mountainous MS Rožaje, Žabljak and Kolašin, making up 16.7% of the share (Figure 7). Percentage–wise, this is the smallest cluster region. Common for the mentioned MS is that they are located at altitudes of about 1000 to 1450 m, located at a relatively short distance from each other, as well as in the mountain climate zone, in general. In addition to these common physical–geographical features, the similarity can also be observed in the thermal aspect, so all three MS register a negative mean temperature in winter, while the average summer is around 14–15°C.

II temperature cluster region includes 8 MS, i.e. it is the largest in percentage (44.4%). These are the stations located in the zone of moderate–continental climate and at an altitude of 650 m to 1000 m. In the thermal regime, the predominant influence in these places is the physical–geographic features of the surrounding area, first of all the elevation, the dissection of the relief and the distance from the sea. Therefore, the stations have moderately warm summers and moderately cold winters. Average summer temperatures range from 17–20°C, while average winter temperatures range from –1°C to 3°C.

III temperature cluster region groups 38.8% of MS included in the analysis (7 out of 18 MS). What these stations have in common is that they are located in the zone of predominant influence of the Mediterranean, and two climatic areas are distinguished: Adriatic–Mediterranean and modified Mediterranean. A modified Mediterranean climate is represented in the Podgorica–Skadar basin and the Bjelopavlići plain (MS Podgorica and Golubovci). Other MS are located along the Montenegrin coast of the Adriatic, and it is a narrow zone with a typical Mediterranean climate (short, mild and rainy winters with rare frosts, long and warm summers, average annual insolation is about 2600 h). The average summer temperature is 24–26°C, and the winter temperature is around 6–9°C.

ing the non–hierarchical agglomerative method and the hierarchical method, it is concluded that the multivariate technique is suitable for research in climatology, because the analysis indicated a logical grouping of MS, primarily based on real physical–geographical factors in geographical area.

Hierarchical agglomerative cluster methods for precipitation (UPGMA, Single linkage i Ward)

The *UPGMA* method distinguishes 6 clusters of precipitation regions. **The first cluster** is the largest, because it groups five MS (27.8% of the total number of stations). These are places in the extreme north and northeast of Montenegro, which have a continental (MS Pljevlja and Rožaje) and a Mediterranean–continental pluviometric regime (MS Bijelo Polje, Berane and Plav). These are the most continental regions of Montenegro with the lowest annual rainfall (average

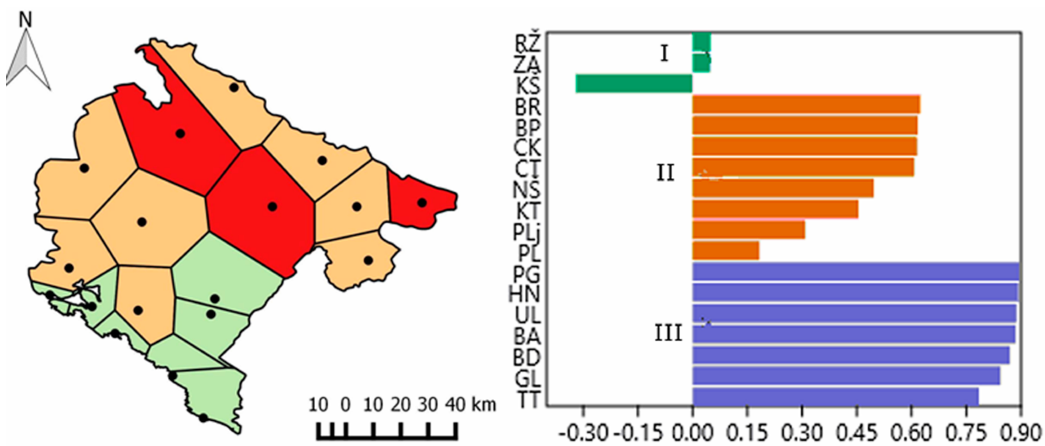

Figure 7. Temperature regions for the period 1961–2020 obtained by the *K–means* method (left) and Silhouette plot (right)
(x–axis: on a scale from –1 to 1, where 1 means a perfectly suitable assignment to a group; –1 means that the object would be better placed in another group; 0 means that the object is on the border between two clusters)

The results obtained using the *K–means* cluster method correspond to a large extent with previous research by Burić et al. (2014). Namely, in the I temperature cluster region there are mountain stations that belong to the *D* climate (*Dfb, Dfc*), as well as MS Kolašin, which is at the transition between *C* and *D* climates. The II temperature cluster region consists of the largest number of stations distributed in climate class *C* and in different climate types *Cs* and *Cf*, depending on the physical–geographical features there is also a climate differentiation. Also, the stations in the III temperature cluster region belong to the same climate class, type and subtype (*C*, *Cs*, *Csa*). In general, based on the comparative analysis of the results of climate regionalization according to the Köppen climate classification, given by the aforementioned authors, and presented in this paper, which were obtained us-

800–1000 mm). The difference in seasonal rainfall is smaller than in the rest of the country. **The second cluster** consists of three MS (Herceg Novi, Nikšić and Kolašin) or 16.7%. This cluster includes the parts of Montenegro that are closer to the sources of moisture (the Atlantic and the Mediterranean Sea) and that have a typical Mediterranean pluviometric regime. The average annual amount of precipitation is about 1900–2100 mm, of which about 32–34% falls in winter, in general. The least precipitation is in summer, only about 10–12% of the annual average. And **the third cluster** groups three MS (Figure 8): Tivat, Golubovci and Podgorica. These are regions with a Csa subtype of climate and a Mediterranean precipitation regime. However, compared to the previous cluster, there is less precipitation in these places – the annual average is around 1600 mm.

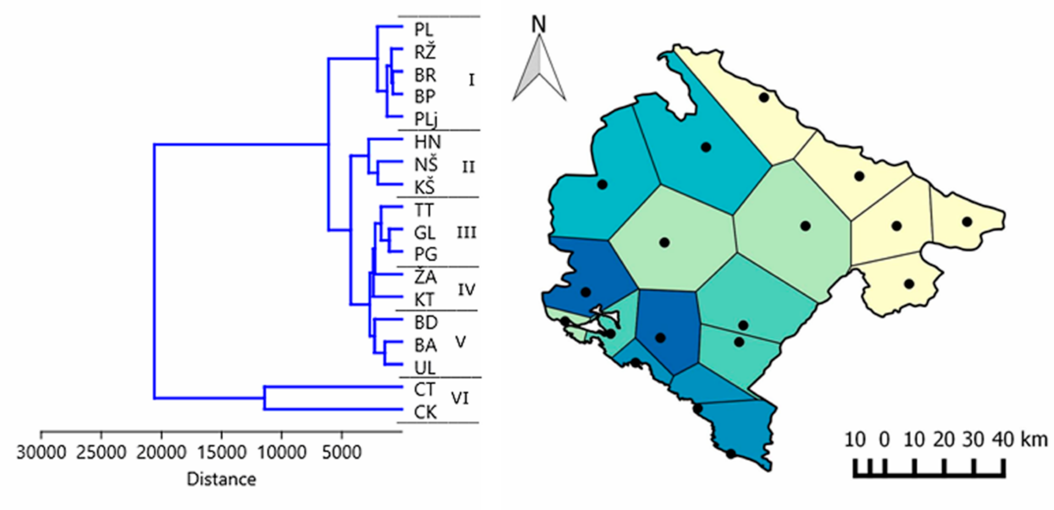

Figure 8. Results of the *UPGMA* method for the period 1961–2020: dendrogram of meteorological stations (left) and precipitation regions (right)

The western part of Montenegro with two MS belongs to **the fourth cluster**. These are the mountain MS (Žabljak and Krstac) with an annual average of about 1500–1650 mm of precipitation. This cluster is also characterized by minimal precipitation in summer, but still in this season about 15–17% of precipitation falls compared to the annual average. Compared to the previous clusters, the difference is that MS Žabljak receives maximum precipitation during the fall, and not in the winter season. **The fifth cluster** consists of three MS on the Montenegrin coast (Budva, Bar and Ulcinj) or 16.7% of the total number of stations (18). The coastal region of Montenegro has a true Mediterranean climate (*Csa*) with a typical Mediterranean pluviometric regime. In this part of Montenegro, the annual average precipitation is around 1300–1500 mm. In summer it is about 10%, and in winter about 33% of the annual precipitation. The last **sixth cluster** includes the rainiest part of Montenegro, which is the southwestern part where MS Crkvice and Cetinje are located (11.1%). For the period 1961–2020, the annual average precipitation in Crkvice is about 4600 mm, and in Cetinje about 3320 mm. It is one of the rainiest regions of Europe, with a Mediterranean precipitation regime, and due to the altitude, the climate formula is *Csa*.

In relation to the clusters obtained by *UPGMA*, the number of cluster regions obtained by the *Single linkage* method is twice as large (12 cluster regions). This fact supports the justification of using multiple methods for extracting cluster regions, especially for precipitation, because it is a very variable climatic element, especially in an area with a dissected relief, as is the case with Montenegro. According to the *Single linkage* method, the changes in the cluster regions are as follows:

- Clusters group the Žabljak and Krstac stations into special rainfall regions with 5.6% of the total number of stations. The *UPGMA* method grouped these two MS into one region.
- Plav and Pljevlja now form a separate cluster region without Rožaje, Berane and Bijelo Polje, so in this case there were changes to 5.6% compared to 27.8% from the first method.
- Budva and Tivat now form separate precipitation regions with 5.6% participation.
- Also, both Cetinje and Crkvice now form separate clusters, and in the first case (*UPGMA* method) those two MS were an integral part of one cluster.

Therefore, the main difference is that the *UPGMA* method does not distinguish clusters with one MS each, while according to *Single linkage* there are as many as 8 cluster regions with one MS each (Figure 9). In other words, out of a total of 12 cluster regions, the percentage share is mostly occupied by clusters with one MS each (8/12 clusters), followed by clusters with two stations each (2/12 clusters) and clusters with three stations each (2/12 clusters).

The smallest number of cluster regions was obtained using the *Ward's method* (5 regions in total), and the results are very similar to the UPGMA method. Comparing the results of these two methods, *UPGMA* and *Ward's*, for the period 1961–2020, the following facts can be observed:

- *Ward's method* separates 5 cluster regions (Figure 10), i.e. one less than UPGMA.
- According to the *Ward's method*, MS Krstac and Žabljak do not now form a common cluster region, but are part of other cluster regions (station Krstac is part of the fourth and Žabljak of the fifth clus-

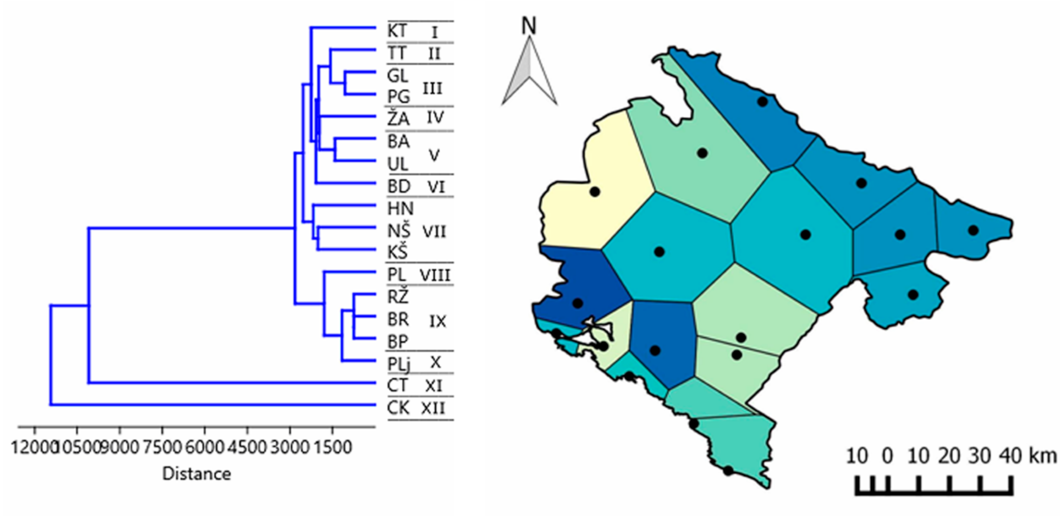

Figure 9. Results of the *Single linkage* method for the period 1961–2020: dendrogram of meteorological stations (left) and precipitation regions (right)

ter region). Recall that the *UPGMA* method groups these two MS into one cluster region.

- In contrast to the results obtained using the *UPGMA* and *Single linkage* methods, in the *Ward's method* clusters with four stations each dominate (22.2% participation).
- Percentagewise, the largest cluster region obtained by the *UPGMA* and *Ward's* methods is 27.8% of the stations (1 cluster region with 5 MS each), and 16.7% by the *Single linkage* method (2 cluster regions with 3 MS each).

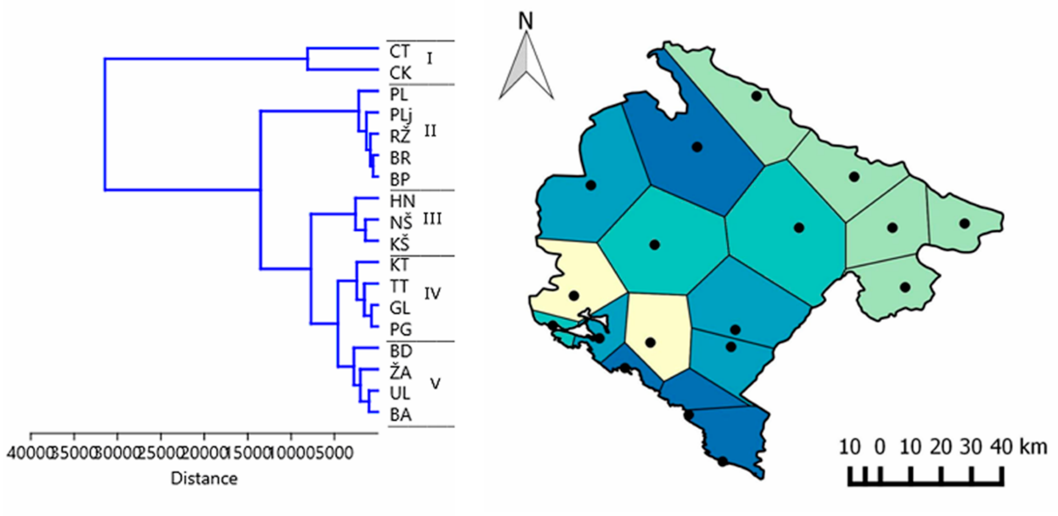

Figure 10. *Ward's* method results for the period 1961–2020: dendrogram of meteorological stations (left) and precipitation regions (right)

A non–hierarchical K–means cluster method for precipitation

In contrast to the results obtained by applying hierarchical methods of regionalization, the *K–means* cluster procedure produces a uniform number of clusters (in this case, the default number of clusters is 3, as in the case of temperature regions), which will facilitate comparative analysis (see Figure 11). The following is a comparison of the cluster regions that include the largest number of the same MS in both groups of methods.

The results of the comparative analysis of the obtained cluster regions in the first and second group of methods indicate the following:

I precipitation cluster region groups 5 stations with 27.8% share, located in the far north and northeast of Montenegro. It is an area with a moderate continental and mountain climate. This part of Montenegro is the farthest from the Adriatic Sea, so the influence of continentality is felt the most. Continental and Mediterranean–continental precipitation regimes are represented. This part has the lowest annual precipitation in Montenegro, an annual average of 800–1000 mm.

II precipitation cluster region includes only two spatially close MS: Cetinje and Crkvice. Percentage–wise, it is the smallest cluster region, and what both stations have in common is that they are located in a mountainous region, in the hinterland of the Adriatic Sea. Due to the specific morphology of the terrain, the two mentioned MS register the highest amount of precipitation in Montenegro (Crkvice about 4600 mm, Cetinje 3320 mm). Therefore, this smallest cluster region logically groups stations that differ in precipitation sums from all other MS included in the analysis, that is, they form a separate precipitation cluster region.

Comparing the results of the formed precipitation cluster regions presented in this paper with the researches of climatic regionalization of Montenegro, given by the mentioned authors, it can be concluded that the precipitation cluster region consists of MS located in the area with a continental precipitation regime and with similar annual sums precipitation. The border of this cluster region coincides with the border between the Mediterranean and continental pluviometric regime defined by the mentioned authors. Around that conditional border, the influences of continentality and maritimeness on the pluviometric regime are interwoven. Therefore, these results undoubtedly indicate the validity of the multivariate techniques used, both the hierarchical and non–hierarchical cluster methods, which were the focus of this

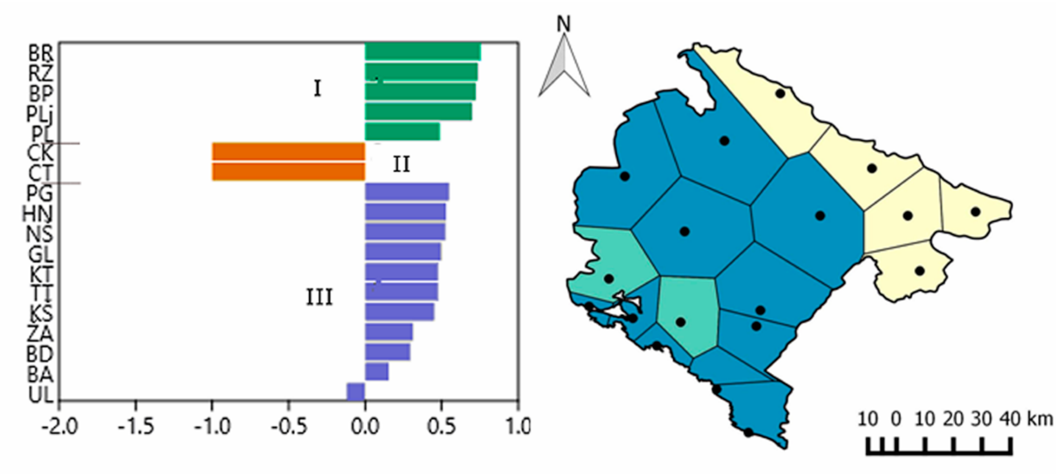

Figure 11. Precipitation regions for the period 1961–2020. obtained by the *K–means* method (right) and Silhouette plot (left)
(x–axis: on a scale from –1 to 1, where 1 means a perfectly suitable assignment to a group; –1 means that the object would be better placed in another group; 0 means that the object is on the border between two clusters)

III precipitation cluster region groups the other 11 MS with a total of 61.1% participation, i.e. this is the largest cluster precipitation region in percentage terms. It is a region that has several climate types (Mediterranean, modified Mediterranean, temperate–continental and mountain climate), but air masses within the cyclone from the Mediterranean Sea and from the west have the dominant influence on precipitation. Within this region, therefore, the Mediterranean precipitation regime is common to all MS. The average annual precipitation in this cluster region ranges from about 1300 mm to 2000 mm.

study. The II precipitation cluster region is comprised of the rainiest stations in Montenegro, which belong to the *Csbx*" climate subtype. Finally, the III precipitation cluster region comprises the largest number of stations characterized by a Mediterranean pluviometric regime and relatively similar annual precipitation sums. Therefore, it is important to note that the separated climate regions based on two different criteria (physical factors used by Burić et al. (2014) in their climatic regionalization of Montenegro and the results of the statistical technique used in this study based on only two climatic variables) match each other.

Conclusion

In this research paper, one of the multivariate techniques used in modern climatology research was applied to achieve the set goals. Hierarchical agglomerative methods and the non–hierarchical *K–means* method were utilized, with *Euclidean distance* as the distance measure. A series of data from the period 1961–2020 was observed, and multi–year values of temperature and precipitation from 18 MS were used as variables. Furthermore, a comparative analysis was carried out based on the results obtained in this study and the climatic regions identified by Burić et al. (2014), which are based on the dominant physical factors in the geographical area of Montenegro.

The research results indicate that the distribution of stations reflects the majority of the climatic diversity of Montenegro. The cluster regions extracted using *UPGMA, Single linkage*, and *Ward's* methods show spatial similarity. As these methods agree to a certain extent, the presented research results can be considered convincing. In other words, the temperature and precipitation cluster regions extracted using different hierarchical agglomerative methods show a high degree of similarity. In both cases (for both temperature and precipitation), the clustering results (i.e., selection of cluster regions) presented in this paper are highly compatible with the climatic regionalization of Montenegro, i.e., the identified climatic types and subtypes given by Burić et al. (2014). Regarding the *K–means* analysis, noticeable logical grouping of meteorological stations is based on real physical–geographical factors in the area. Particularly, this method (*K–means*) clearly separates the precipitation cluster regions in correlation with annual sums and pluviometric regime. Further research on this issue should include other climatic variables (e.g., relative humidity, number of hours of sunshine, average effective precipitation, average wind speed, number of frosty days, etc.) and various multivariate techniques, such as factor analysis, analysis of principal components (*Principal Component Analysis–PCA*) of empirical orthogonal functions, etc.

References

Appendini, K., & Liverman D. (1994). Climate change and food security in Mexico. *Food Policy*, 19(2), 149–164. https://doi.org/10.1016/0306-9192(94)90067-1

Arbabi, A. (2011). Cluster–based method for understanding the climactic diversity of Iran. *African Journal of Agricultural Research*, 6(31), 6525–6529. https://doi.org/10.5897/AJAR11.1321

Aurenhammer, F., & Klein, R. (2000). Voronoi Diagrams. Ch. 5. In J.R. Sack & J. Urrutia (Eds.), *Handbook of Computational Geometry* (pp.201–290 pp). Amsterdam: Elsevier. https://www.researchgate.net/publication/279959776

Barber, C. B., Dobkin, D. P., & Huhdanpaa, H. (1996). The quickhull algorithm for convex hulls. *ACM Transactions on Mathematical Software (TOMS)*, 22(4), 469-483. https://doi.org/10.1145/235815.235821

Bensmail, H., Celeux, G., Raftery, A., & Robert, C. (1999). Inference in Model–Based Cluster Analysis. *Statistics and Computing*, 7(1), 1–10. https://www.researchgate.net/publication/2844935

Blaney, H.P., & Cridle, W.D. (1950). *Determining water requirement in Irrigated areas climatological and Irrigation Data*. Washington, USA: USDA Soil Conservation Service. https://ia800300.us.archive.org/4/items/determiningwater96blan/determiningwater96blan.pdf

Burić D., Ducić V., & Mihajlović, J. (2013). The climate of Montenegro: Modificators and types – part one. *Bulletin of the Serbian Geographical Society*, 93(4), 83–102. https://doi.org/10.2298/GSGD1304083B

Burić D., Ducić V., & Mihajlović, J. (2014). The climate of Montenegro: Modificators and types – part two. *Bulletin of the Serbian Geographical Society*, 94(1),73–90. https://doi.org/10.2298/GSGD1401073B

Burić D., & Doderović M. (2022). Trend of Percentile Climate Indices in Montenegro in the Period 1961–2020. *Sustainability*, 14(19), 12519. https://doi.org/10.3390/su141912519

Clark, P. J., & Evans, F. C. (1954). Distance to nearest neighbor as a measure of spatial relationships in populations. *Ecology*, 35, 445–453. https://doi.org/10.2307/1931034

Davis, J. C. (2002). *Statistics and Data Analysis in Geology, 3rd Edition*. New York: John Wiley & Sons, pp. 656. https://www.bookdepository.com/Statistics–Data-Analysis-Geology-John-C-Davis/9780471172758

DeGaetano, A. T., & Shulman, M. D. (1990). A climatic classification of plant hardiness in the UnitedStates and Canada. *Agricultural and Forest Meteorology*, 51(3–4), 333–351. https://doi.org/10.1016/0168-1923(90)90117-O

Everitt, B.S., Landau, S., & Leese, M., Stahl, D. (2011). *Cluster Analysis, Fifth edition*. UK: John Wiley & Sons. https://cicerocq.files.wordpress.com/2019/05/cluster-analysis_5ed_everitt.pdf

Fovell, R. G., & Fovell, M. Y. C. (1993). Climate zones of the conterminous United States defined using cluster analysis. *Journal of Climate*, 6(11), 2103–2135. https://www.jstor.org/stable/26198599

Fraley, C., & Raftery, A. (1998). How Many Clusters? Which clustering method? Answers via Model–Based Cluster Analysis. *The Computer Journal*, 41(8), 578–588. http://dx.doi.org/10.1093/comjnl/41.8.578

Guibas, L., & Stolfi, J. (1985). Primitives for the manipulation of general subdivisions and the computation of Voronoi. *ACM transactions on graphics* (TOG), 4(2), 74-123. https://doi.org/10.1145/282918.282923

Hartigan, J. A., & Wong, M. A. (1979). Algorithm AS 136: A K–Means Clustering Algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics),* 28(1), 100–108. https://doi.org/10.2307/2346830

Johnson, D. E. (1998). *Applied multivariate methods for data analysts, edit 2,* New York: Duxbury Press. ISBN–10: 0534237967; ISBN–13: 978–0534237967.

Moreira, E. E., Paulo, A. A., Pereira, L. S., & Mexia, J. T. (2006). Analysis of SPI drought class transitions using loglinear models. *Journal of Hydrology,* 331(1–2), 349–359. https://doi.org/10.1016/j.jhydrol.2006.05.022

MacQueen, J. B. (1967). Some methods for classification and analysis of multivariate observations. In L. M. Le Cam & J. Neyman (Eds.), *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability* (pp. 281–297). California: University of California Press. https://digitalassets.lib.berkeley.edu/math/ucb/text/math_s5_v1_article-17.pdf

Nassiri, M., Koocheki, A., Kamali, G. A., & Shahandeh, H. (2006). Potential impact of climate change on rainfed wheat production in Iran: (Potentieller Einfluss des Klimawandels auf die Weizenproduktion unter Rainfed-Bedingungen im Iran). *Archives of agronomy and soil science*, 52(1), 113-124. https://www.researchgate.net/publication/233020291

Okabe, A., Boots, B., & Sugihara, K. (2000). *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams, 2nd ed.* New York: Wiley. ISBN: 978-0-471-98635-5

Papić-Blagojević, N., & Bugar, D. (2009). Osnovne premise analize grupisanja. *Škola biznisa: Naučnostručni časopis,* 4, 166-173. http://www.vps.ns.ac.rs/SB/2009/4.18.pdf

Preparata, F. R., & Shamos, M. I. (1985). *Computational Geometry: An Introduction*. New York: Springer–Verlag, pp. 390. ISBN: 978-1-4612-1098-6

Rencher, A. C. (2002). *Methods of Multivariate Analysis, Second edition*. Wiley, pp. 715. https://doi.org/10.1002/0471271357

Szép, I. J., Mika, J., & Dunkel, Z. (2005). Palmer drought severity index as soil moisture indicator: physical interpretation, statistical behavior and relation to global climate. *Physics and Chemistry of the Earth*, 30(1–3), 231–243. https://doi.org/10.1016/j.pce.2004.08.039

Szentimrey, T. (2003). Multiple analysis of series for homogenization (MASH); Verification procedure for homogenized time series, in: Fourth seminar for homogenization and qualitycontrol in climatological databases. Budapest, Hungary, *WMO–TD No. 1236, WCDMP No. 56*, 193–201.

Szentimrey, T., & Bihari. Z. (2007). Mathematical background of the spatial interpolation methods and the software MISH (Meteorological Interpolation based on Surface Homogenized Data Basis), *Proceedings of the Conference on Spatial Interpolation in -Climatology and Meteorology* (pp. 17-27), Budapest, Hungary.

The Pennsylvania State University (2004). *Ward's method*. Available at: https://online.stat.psu.edu/stat505/lesson/14/14.7

Van Groenewoud, H. (1984). The climatic regions of New Brunswick: A multivariate analysis ofmeteorological data. *Canadian Journal of Forest Research*, 14(3), 389–394. https://doi.org/10.1139/x84-069

Warrington, A. (1977). Crop phonological stages. *Australian Journal of Agricultural Research,* 28(1), 11–27. https://doi.org/10.1071/AR9770011

Wolter, K. (1987). The Southern Oscillation in surface circulation and climate over the tropical Atlantic, Eastern Pacific, and Indian Oceans as captured by cluster analysis. *Journal of Applied Meteorology and Climatology,* 26(4), 540-558. https://doi.org/10.1175/1520-0450(1987)026<0540:TSOISC>2.0.CO;2