Coğrafi Bilimler Dergisi / Turkish Journal of Geographical Sciences
2026; 24(1): 321-336 DOI: 10.33688/aucbd.1695834
Research Article / Araştırma Makalesi

Robust panel clustering of station-scale precipitation in Türkiye

Türkiye'deki şehirlerin yağışları için robust panel kümeleme analizi

Çağlar Sözen*[a], Hasan Bulut[b]

[a] Giresun Üniversitesi, Uygulamalı Bilimler Yüksekokulu, Giresun/Türkiye

[b] Ondokuz Mayıs Üniversitesi, Fen Fakültesi, Samsun/Türkiye

Abstract

This study uses station-level annual precipitation series for Türkiye to objectively delineate homogeneous precipitation regimes. We first applied k-means, PAM, and hierarchical (agglomerative) clustering for K=2"–" 10 and evaluated the solutions using the APN, AD, ADM, FOM, Connectivity, Dunn, and Silhouette indices. Although classical algorithms favored K=2, Rize acted as a high-leverage observation, exerting a disproportionate influence on centroid estimation, producing a singleton cluster, and obscuring meaningful substructure among the remaining stations. We therefore adopted a robust trimmed k-means approach. A multi-index assessment identified K=4 as optimal for APN/ADM/FOM, yielding four structural regimes (C1–C4) while treating Rize as a trimmed singleton. The resulting partition quantifies—at the station scale—the well-known contrast between coastal/orographic enhancement and continental drying, producing clusters that are more compact, less overlapping, and climatologically interpretable.

Keywords: *Robust clustering, Trimmed k-means, Cluster validation indices, Precipitation*

Öz

Bu çalışma, Türkiye genelinde istasyon düzeyi yıllık yağış serilerini kullanarak homojen yağış rejimlerini nesnel biçimde ortaya koyar. Önce k-ortalamalar, PAM, ve hiyerarşik (agglomeratif) yöntemleri $K = 2$–10 aralığında uyguladık; çözümleri APN, AD, ADM, FOM, Connectivity, Dunn, ve Silhouette endeksleriyle doğruladık. Klasik algoritmalar $K = 2$'yi işaret etse de, Rize'nin yüksek kaldıraçlı bir gözlem olarak centroid tahminini orantısız etkilemesi, geri kalan istasyonlarda mevcut anlamlı alt yapıyı maskeledi ve tekil bir küme ortaya çıkardı. Bu nedenle robust budanmış k-ortalamalar yaklaşımına geçtik. Çok ölçütlü değerlendirme, APN/ADM/FOM'da $K = 4$'ün en uygun olduğunu gösterdi; Rize trimmed tekil olarak ayrılırken kalan ağ dört yapısal rejime ayrıldı (C1–C4). Elde edilen rejimler, Türkiye yağış klimatolojisindeki kıyısal/orografik güçlendirme ile kıtasal kuruma arasındaki bilinen dikotomiyi istasyon ölçeğinde nicel olarak doğrular; kümeler daha kompakt, daha az örtüşen ve klimatolojik olarak yorumlanabilir niteliktedir.

Anahtar Kelimeler: *Robust kümeleme, Budanmış k-ortalamalar, Küme geçerlilik endeksleri, Yağış*

***Sorumlu Yazar/Corresponding Author:** caglar.sozen@giresun.edu.tr

Geliş/Received: 09.05.2025 **Kabul/Accepted:** 10.04.2026 **Yayın Tarihi/Online Published**:23.04.2026

Atıf/To Cite: Sözen, Ç., Bulut, H. (2026). Robust panel clustering of station-scale precipitation in Türkiye, *Coğrafi Bilimler Dergisi/ Turkish Journal of Geographical Sciences,* 24(1), 321-336, https://doi.org/10.33688/aucbd.1695834

1. Introduction

The accelerating pace of climate change is reshaping natural and human systems across multiple dimensions—freshwater availability, agricultural sustainability, ecosystem resilience, and public health and safety—through both gradual warming and the rising frequency and intensity of extremes (IPCC, 2021). In particular, shifts in precipitation regimes—manifested as altered seasonality, more frequent droughts, convective downpours, and compound flood events—pose a direct risk to water resources and to the infrastructure and livelihoods that depend on them. These risks are magnified in developing and geographically diverse settings where topography, distance from coasts, and air-mass interactions create sharp spatial gradients in climate.

Türkiye exemplifies such heterogeneity. Orographic uplift along the Pontic and Taurus ranges, coastal influences on the Black Sea and Mediterranean littorals, and the continental interior basins together generate pronounced regional contrasts: persistently wet conditions in parts of the Black Sea region versus low precipitation, recurrent drought, and high evapotranspiration in Central and Southeastern Anatolia (Demircan et al., 2017; Türkeş, 1996). Designing credible adaptation and risk-management policies therefore requires spatially resolved evidence on how precipitation varies across space and time—evidence that cannot be obtained by aggregating to administrative units alone.

Long observational records from meteorological stations remain the backbone of such evidence. Station-level precipitation series enable the detection of spatial correlations, regime boundaries, and temporal anomalies at resolutions relevant to hydrology, agriculture, and urban planning (Trenberth, 2011; Wilks, 2011). Yet, extracting coherent spatial structure from high-dimensional, noisy climate panels is non-trivial: classical clustering methods can be distorted by outliers (e.g., extreme events, measurement artifacts) and by uneven network coverage, which in turn can bias cluster centers and blur boundaries.

In this study, we analyze station-level annual precipitation across Türkiye for the period 1965–2022 from a panel-clustering perspective. We implement three widely used classical clustering algorithms—k-means, Partitioning Around Medoids (PAM), and hierarchical (agglomerative) clustering—chosen for their complementary strengths in environmental applications (Jain et al., 1999; Kaufman & Rousseeuw, 1990; MacQueen, 1967). To mitigate sensitivity to extreme values and non-Gaussian tails that commonly arise in hydroclimate data, we also consider a robust trimmed k-means variant, which excludes a small fraction of the most discordant observations during centroid estimation and thereby improves cluster compactness and stability (Cuesta-Albertos et al., 1997; García-Escudero et al., 2008; Hubert & Arabie, 1985).

We evaluate algorithmic performance using a comprehensive set of seven objective cluster validity indices that span measures of internal compactness and separation, as well as stability under feature perturbations: Silhouette, Dunn, Connectivity, Average Proportion of Non-overlap (APN), Average Distance (AD), Average Distance between Means (ADM), and Figure of Merit (FOM) (Brock, Pihur et al., 2008; Mahmud et al., 2022). Framing the problem as panel clustering allows us to respect the time-series structure at each station while grouping stations by similarity in their long-term precipitation behavior.

Taken together, this paper makes the following contributions: (i) We provide a station-level climatology of long-term precipitation regimes for Türkiye, avoiding administrative aggregation and

using province labels only for geographic orientation; (ii) We benchmark classical versus robust clustering under realistic hydroclimate noise, showing when trimming yields more interpretable and statistically consistent groupings; (iii) We implement a multi-criterion validity assessment that integrates internal and stability indices to support defensible choices of the number of clusters and paradigm. The remainder of the paper is organized as follows: Section 2 details the data and methods; Section 3 presents empirical results and sensitivity analyses; Section 4 discusses implications for adaptation policy and risk management; Section 5 concludes.

1.1. Literature Review

Although numerous studies have sought to characterize Türkiye's climatic diversity, research that objectively delineates homogeneous climatic or precipitation regions from long-term station-based observations remains comparatively limited. Early work largely relied on statistical summaries or geographic proximity rather than data-driven clustering of extended meteorological records.

One of the first comprehensive attempts to redefine Türkiye's climate regions using objective methods was conducted by Ünal et al. (2003). Using mean, maximum, and minimum temperatures and total precipitation from 113 meteorological stations for 1951–1998, they tested five hierarchical clustering techniques and identified seven climate regions using Ward's method. While the resulting partition broadly resembled traditional regional divisions, the study did not incorporate robustness checks for outliers or investigate temporal variability in climate signals.

In Türkeş and Tatlı (2011), spectral clustering was applied to annual precipitation totals from 96 stations (1929–2007). The hybrid SVD + k-means approach yielded eight coherent precipitation regions (namely: Black Sea, Northwestern Türkiye, South Aegean–Western Mediterranean, Mediterranean, Western Continental Central Anatolia, Eastern Continental Central Anatolia, Eastern Anatolia, and Southeastern Anatolia). Although the method captured the roles of synoptic systems and topography in shaping precipitation, it was univariate and did not use cluster-validity indices for model selection.

Kartal et al. (2011) standardized long-term temperature and precipitation (1950–2006) and, using Ward's linkage, obtained seven climatic regions whose boundaries diverged from conventional geographic partitions—highlighting the value of objective classification. As with earlier work, however, sensitivity to noise or outliers was not assessed, nor was stability across alternative algorithms explored.

Compared with classical approaches, later studies began to employ fuzzy or other robust clustering methods. Dikbas et al. (2012) analyzed precipitation series from 188 stations (1967–1998) using fuzzy c-means, identifying six homogeneous rainfall regions and demonstrating, via L-moment tests, that fuzzy, hydrologically defined regions outperform administrative or purely geographic boundaries in regional frequency analysis.

Building on such insights, İyigün et al. (2013) combined temperature, precipitation, and relative humidity data from 244 stations (1970–2010) and tested 12- and 14-cluster scenarios with Ward's method; they argued that 14 clusters better represented Türkiye's physiographic and climatic gradients. Despite these advances, the approach remained sensitive to anomalies and did not explicitly treat noise or extremes.

Zeybekoğlu and Ülke Keskin (2020) normalized 30–78-year records of maximum rainfall intensity (augmented with latitude, longitude, and elevation) from 95 stations and applied fuzzy c-means clustering, identifying five intensity clusters (four distinct zones plus one transition zone). The result—stations from disparate administrative regions falling into the same cluster—illustrated the suitability of FCM for national-scale rainfall-intensity analyses.

More recently, Kömüşcü et al. (2022) applied k-means clustering to monthly and annual precipitation data from 234 stations (1980–2020), delineating five principal precipitation regimes and noting their potential modulation by large-scale teleconnections such as the NAO and MO. The study, however, relied on a single algorithm and did not benchmark robustness or employ multi-index validation.

Beyond clustering, alternative classification frameworks have also been used. Using Holdridge life zones, Tekin et al. (2021) analyzed data from 337 stations and identified 12 bioclimatic zones, emphasizing increased biodiversity and zone diversity where physical geographic factors vary over short distances and noting mismatches between potential and actual vegetation in some areas. Similarly, Taşoğlu et al. (2024) produced a high-resolution Köppen–Geiger map by correcting CHELSA fields with station observations, identifying 14 climate types and numerous microclimates—including the co-occurrence of multiple types over short distances in the Eastern Black Sea and the presence of EF (ice-cap) conditions at the summit of Mount Ararat.

1.2. Contributions and Methodological Innovations of the Present Study

Despite valuable foundations, the literature exhibits recurring limitations:

Outliers and measurement noise in long meteorological records are rarely addressed, risking biased centroids and blurred boundaries.

Most studies rely on a single algorithm (e.g., k-means or Ward) and do not consider robust alternatives such as trimmed k-means or fuzzy-robust hybrids.

Multi-index validation, for objective assessment of the number and quality of clusters (e.g., Silhouette, Dunn, Connectivity), is seldom employed.

To address these gaps, the present study analyzes annual precipitation from meteorological stations across all 81 provinces of Türkiye for 1965–2022, comparing classical algorithms (k-means, PAM, hierarchical/agglomerative) with a robust trimmed k-means variant. Data are standardized, and a suite of seven validity indices—Silhouette, Dunn, Connectivity, APN, AD, ADM, and FOM—is used to evaluate clustering quality and stability. This design aims to yield more reliable and interpretable station-level homogeneous precipitation regions in Türkiye, providing a sound basis for climate-adaptation planning and policy.

2. Materials and Methods

Cluster analysis is one of the most popular multivariate statistical methods used to group units according to their similarities and differences (Rencher, 2002). This section introduces the classical and robust clustering algorithms used in this study, as well as cluster validity indices and panel clustering approaches.

2.1. Classical Clustering Algorithms

2.1.1. Hierarchical Clustering Algorithm

In hierarchical clustering analysis, all observations are considered a single cluster, and new clusters are formed by bringing together the most similar observations (Rencher, 2002). As a similarity measure, the distance matrix is used (Bulut, 2023). Calculations for hierarchical clustering were performed using the hclust() function in the stats package in R (R Core Team, 2023).

2.1.2. K-means clustering algorithm

In the k-means algorithm, the most popular clustering method, the number of clusters should be determined a priori. The objective function of the k-means algorithm is given in Equation (2.1).

$$\min_{m_1,\dots,m_k} \sum_{i=1}^{n} j = \min_{1,2,\dots,k} \|x_i - m_j\|^2 \qquad (2.1)$$

Where m_j is the centre of the jth cluster ($j = 1,\dots,k$) and $\boldsymbol{x}_i$ is the ith observation vector ($i = 1,\dots,n$) (Aggarwal and Reddy, 2014; Bulut, 2023). The calculations for the k-means algorithm were performed using the kmeans() function in the stats package in R (R Core Team, 2023).

2.1.3. Partitioning Around Medoids (PAM) Clustering

Similar to the k-means algorithm, the PAM clustering algorithm requires an initial number of clusters. The PAM algorithm randomly selects k observations as medoids and assigns the remaining observations to the cluster of the nearest medoid. Then, it checks whether there is an observation that minimises the total distance to the other observations in the cluster, and if such an observation is found, this observation is selected as the medoid. When medoids are updated, all observations are reassigned to the cluster with the closest medoid, and the process continues until the medoids remain unchanged and no further transitions between clusters occur (Kaufman & Rousseeuw, 1990). The calculations related to the PAM algorithm were performed using the pam() function from the cluster package in R (Maechler, 2018).

2.2. Cluster Validation

Cluster validity indices are frequently used to determine the optimal number of clusters and the optimal clustering algorithm. In this study, they were used to determine which clustering algorithm and number of clusters to use. These indices can be divided into two categories: internal indices and stability indices. All calculations related to the indices were performed using the clValid package in R (Brock et al., 2008).

2.2.1. Internal Indices

The most popular of these indices is the Silhouette width index. The silhouette width index is the average of the silhouette values of each observation. The silhouette value of observation i is calculated as shown in Equation 2.2. The silhouette width index takes a value between [-1,1] and is interpreted as the clustering result improves as it approaches 1 (Bulut, 2023; Maechler, 2018).

$$S(i) = \frac{b_i - a_i}{max(b_i, a_i)} \qquad (2.2)$$

Another internal index is the Dunn index. This index can take values from zero to infinity and, like the Silhouette index, larger values of the index are used as evidence for better clustering results (Brock et al., 2008; Bulut, 2023). Finally, the connectivity index can take values from zero to infinity, just like the Dunn index. However, unlike the Dunn index, smaller values of the connectivity index indicate greater clustering success (Brock et al., 2008).

2.2.2. Stability Indices

Stability indices make judgements about clustering success by examining the changes in clustering results obtained when each column is removed in turn from the data, compared with the clustering result obtained from the full data. Optimal results for all indices are valid for small index values.

The average proportion of non-overlap (APN) index measures the average proportion of observations assigned to different clusters when comparing clustering based on the full data with clustering obtained after removing a column. APN takes values in [0, 1]. The average distance (AD) measures the mean separation between observations that remain in the same cluster when clustering is performed on the full dataset and on the dataset with one column removed. AD takes values in the range $(0, \infty)$. The average distance between means (ADM) is the average of the Euclidean distances between the vectors $\overline{X}_{C^{i,l}}$ and $\overline{X}_{C^{i,0}}$, where $\overline{X}_{C^{i,0}}$ is the average vector of observations in the cluster with the ith observation in the clustering based on the full data and $\overline{X}_{C^{i,l}}$ is the average vector of observations in the cluster with the ith observation in the clustering result obtained when the lth column is deleted from the data. ADM takes values in the range $(0, \infty)$. Finally, the figure of merit (FOM) is used to measure intra-cluster variability when a column is removed from the data. FOM is calculated as given in Equation 2.3.

$$FOM = \frac{1}{N}\sum_{h=1}^{k}\sum_{i \in C_h(l)} dist\left(x_{ij}, \overline{x}(l)\right) \qquad (2.3)$$

Where x_{ij} is the value of the ith observation with respect to the jth variable and $\overline{x}(l)$ is the cluster mean. FOM takes values in the range $(0, \infty)$ (Brock et al., 2008; Mahmud et al., 2022).

2.3. Panel Clustering

Classical cluster analysis algorithms are used in multivariate data sets where observations are in rows and variables are in columns. However, because climate data are measured over time, they constitute panel data. Therefore, these panel data should be prepared for cluster analysis. This type of clustering is called panel clustering analysis. Panel clustering analysis is also divided into two types, univariate and multivariate, depending on the number of variables used. In this study, a univariate panel clustering approach is introduced because clustering is performed solely on the precipitation variable.

In univariate panel data, observations occupy rows, and columns contain values of the same variable measured at different time periods. We refer to this data structure as single-index panel data. In single-index panel data, if the number of observations is N and the number of time-dependent measurements is T, $X_i(t)$ is the measurement value of the ith observation ($i = 1,2,\ldots,N$) at time t ($t = 1,2,\ldots,T$) for the relevant variable. The data matrix can be represented as follows.

$$\begin{bmatrix} X_1(1) & X_1(2) & \cdots & X_1(T) \\ X_2(1) & X_2(2) & \cdots & X_2(T) \\ \vdots & \vdots & \ddots & \vdots \\ X_N(1) & X_N(2) & \cdots & X_N(T) \end{bmatrix} \quad (2.4)$$

The univariate panel clustering algorithm operates similarly to classical clustering algorithms. In the data used here, N observations are arranged in rows. However, unlike the classical approach, where p variables occupy the columns, the data consist of values measured at T different times for the same variable (Wang & Lu, 2021).

3. Findings and Discussion

Türkiye, located between 36°–42° N and 26°–45° E at the interface of continental and maritime influences, exhibits marked climatic heterogeneity. In this study, we use the precipitation values from 81 stations shown in Figure 1.

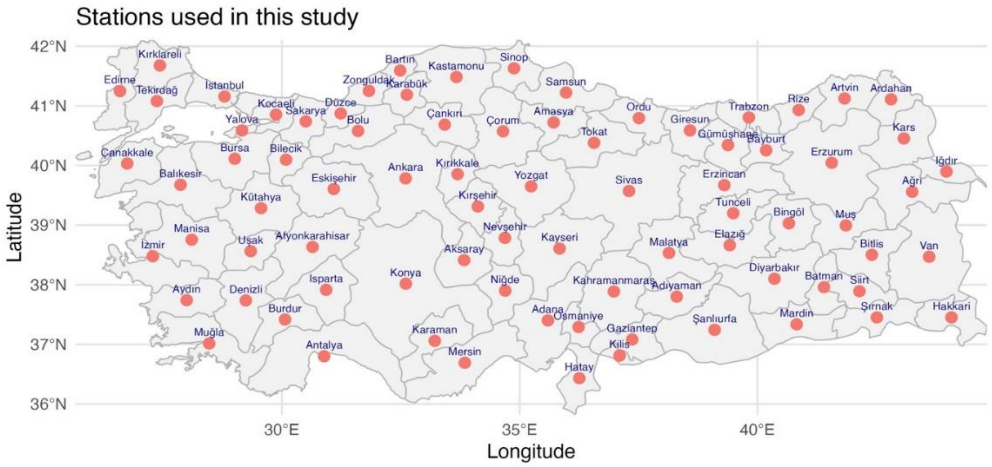

Figure 1. Stations used in this study

To summarize heterogeneity in annual precipitation at the station level, we applied k-means, PAM (Partitioning Around Medoids), and hierarchical (agglomerative) clustering to the TSMS (2023) annual precipitation series for 1965–2022 over $K = 2$–10. For each algorithm–K combination, we computed seven validity indices—APN, AD, ADM, FOM, Connectivity, Dunn, and Silhouette—using the convention that lower APN/AD/ADM/FOM/Connectivity and higher Dunn/Silhouette indicate better clustering.

According to Table 1, for k-means, three of seven indices recommend $K = 2$ as optimal. At that cut, separation and compactness are pronounced (Silhouette = 0.779, Dunn = 0.945) and connectivity is low (2.929). PAM likewise performs best overall at $K = 2$, but with weaker separation (Silhouette 0.426; Dunn 0.092) and poorer neighborhood coherence (Connectivity 10.434). In the hierarchical approach, five of seven indices agree; at this cut ($K = 2$), the Silhouette and Dunn values are virtually identical to those from k-means (0.779 and 0.945), and Connectivity is again 2.929. By contrast, several stability-oriented criteria (e.g., AD, FOM) tend to improve monotonically as K increases; indeed, AD attains its minimum at PAM (3.613), and FOM attains its minimum at k-means ($K = 10$, 0.404). This expected monotonicity implies such criteria are not, by themselves, decisive for a parsimonious K; they should be weighed alongside separation (Silhouette, Dunn) and local structure (Connectivity), which reflect the climatological interpretability of the map.

Table 1. Cluster validity indices for classical clustering results

Method	Validation Indices	Cluster number									Optimum
		2	3	4	5	6	7	8	9	10	
k-means	APN	0.052	0.014	**0.005**	0.013	0.115	0.013	0.011	0.023	0.097	0.005
	AD	7.647	5.575	4.551	4.405	4.355	4.109	4.002	3.676	**3.639**	3.639
	ADM	0.828	0.162	**0.039**	0.117	0.608	0.068	0.055	0.186	0.588	0.039
	FOM	0.811	0.583	0.476	0.468	0.449	0.444	0.433	0.406	**0.404**	0.404
	Connectivity	**2.929**	15.129	30.181	38.754	40.786	42.903	43.903	45.969	49.090	2.929
	Dunn	**0.945**	0.229	0.221	0.254	0.262	0.290	0.296	0.283	0.283	0.945
	Silhouette	**0.779**	0.495	0.365	0.345	0.337	0.337	0.336	0.273	0.238	0.779
pam	APN	**0.010**	0.044	0.044	0.061	0.017	0.097	0.124	0.150	0.138	0.010
	AD	6.090	5.196	4.639	4.240	4.000	3.940	3.856	3.815	**3.613**	3.613
	ADM	0.121	0.332	0.313	0.389	**0.100**	0.486	0.656	0.925	0.706	0.100
	FOM	0.759	0.623	0.483	0.457	0.435	0.428	0.425	0.419	**0.407**	0.407
	Connectivity	**10.434**	26.230	29.159	32.668	44.774	63.931	66.833	58.937	73.187	10.434
	Dunn	0.092	0.078	**0.231**	0.211	0.211	0.139	0.146	0.160	0.160	0.231
	Silhouette	**0.426**	0.354	0.370	0.261	0.251	0.139	0.135	0.159	0.135	0.426
hierarchical	APN	**0.000**	0.002	0.002	0.002	0.016	0.009	0.014	0.013	0.009	0.000
	AD	7.501	5.896	5.749	5.620	5.521	4.335	4.182	4.059	**3.914**	3.914
	ADM	**0.000**	0.023	0.034	0.028	0.119	0.184	0.108	0.096	0.070	0.000
	FOM	0.827	0.615	0.610	0.603	0.595	0.462	0.454	0.444	**0.431**	0.431
	Connectivity	**2.929**	10.462	14.525	17.409	18.409	28.458	30.536	34.106	37.530	2.929
	Dunn	**0.945**	0.353	0.353	0.353	0.353	0.372	0.372	0.372	0.372	0.945
	Silhouette	**0.779**	0.543	0.509	0.497	0.483	0.364	0.355	0.346	0.329	0.779

This assessment is reinforced by Table 2: four of seven indices identify k-means as the best-performing method, and five of seven indicate $K = 2$ as an optimal cluster count. Accordingly, we adopt k-means at $K = 2$ as the classical baseline; the hierarchical result at the same cut yields nearly identical separation metrics and thus serves as corroboration. The k-means $K = 2$ partition is visualized in Figure 2. On that map, the larger cluster contains 80 stations, while a singleton corresponds to Rize. Figure 3 corroborates this: the ordination (Dim1 = 83.3%, Dim2 = 3.7%) places Rize at the extreme positive end of Dim1, far from other stations. Within the large envelope, the remaining 80 stations show evident internal dispersion and substructure: coastal Black Sea sites lie nearer the high-precipitation extreme; many interior and western stations occupy center–left positions; and a few southern orographic sites

extend along the weaker second axis. The elongated hull and visible substructure indicate that the classical k-means $K = 2$ partition is dominated by a single high-leverage observation, suppressing finer splits among the remaining stations. This motivates the adoption of a robust alternative—trimmed k-means—to attenuate outlier influence and to allow substantively meaningful secondary groups to emerge.

Table 2. Optimal number of clusters and clustering algorithm according to cluster validity indices

cluster validity indices	Score	Method	Clusters
APN	0	hierarchical	**2**
AD	3.6131	pam	10
ADM	0	hierarchical	**2**
FOM	0.4044	**k-means**	**10**
Connectivity	2.929	**k-means**	**2**
Dunn	0.9447	**k-means**	**2**
Silhouette	0.7793	**k-means**	**2**

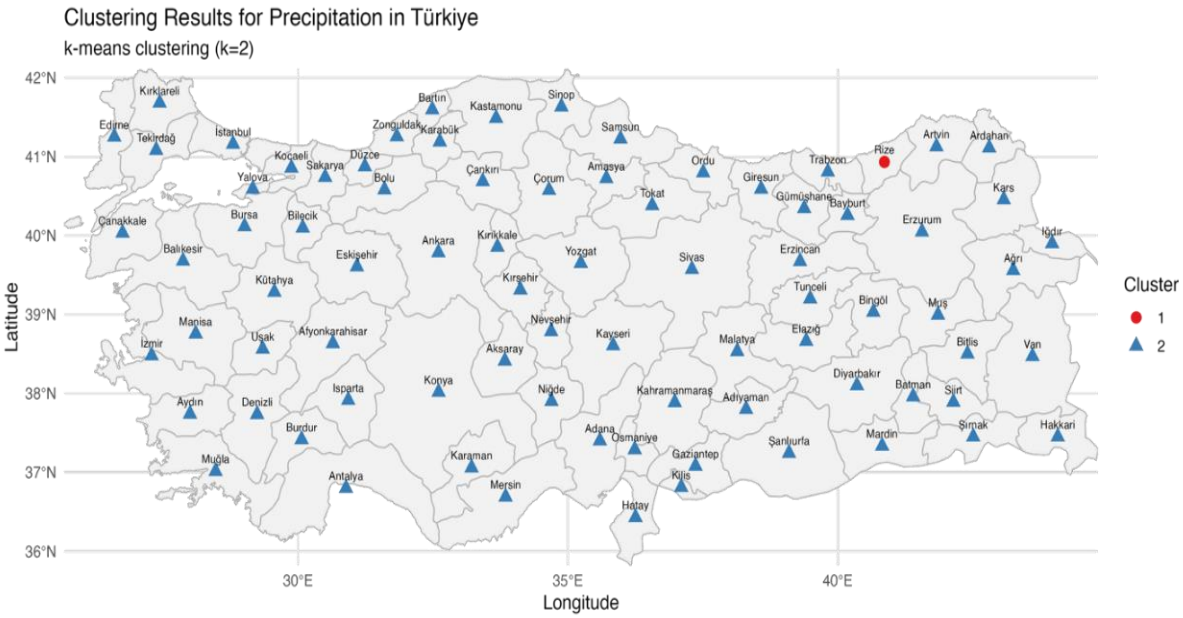

Figure 2. Clustering result for k-means algorithm

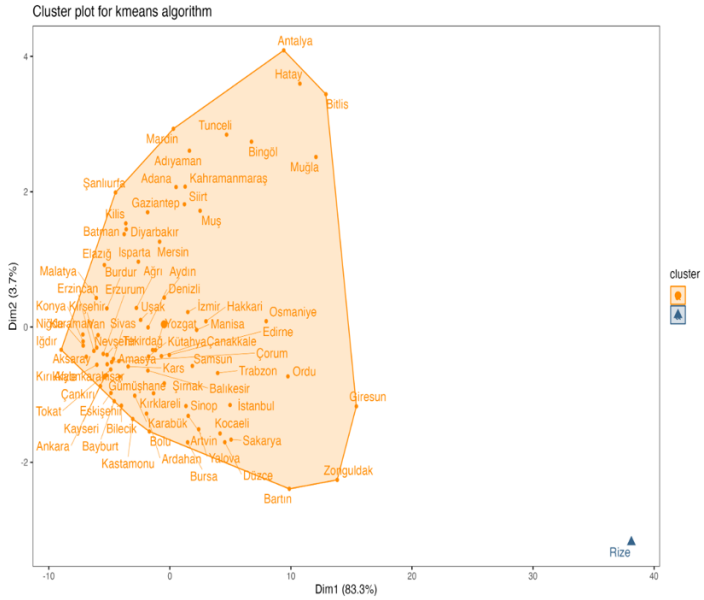

Figure 3. Scatter plot of k-means clustering result

Table 3 shows the trajectories of the seven indices for trimmed k-means over $K = 2$–10. Three indices—APN (minimum 0.09 at $K = 4$), ADM (minimum 259.99 at $K = 4$), and FOM (minimum 174.51 at $K = 4$)—converge on $K = 4$, indicating reduced non-overlap, smaller shifts in cluster means, and lower within-cluster dispersion under feature deletion at that resolution. Connectivity is minimized at $K = 2$ (23.29) and then rises as K increases; Silhouette likewise peaks at $K = 2$ (0.43), consistent with separation-oriented criteria favoring more parsimonious partitions. AD decreases nearly monotonically, reaching its minimum at $K = 9$ (1350.05), while Dunn is maximized at $K = 7$ (0.26). Because some stability and dispersion metrics improve mechanically as K grows, they are not individually decisive for a parsimonious K.

Table 3. Determination of the optimal number of clusters for the trimmed k-means method

Cluster Validation	2	3	4	5	6	7	8	9	10	Optimum
APN	0.17	0.10	**0.09**	0.16	0.28	0.25	0.40	0.34	0.41	0.09
AD	1983.67	1617.94	1474.02	1529.05	1521.19	1426.76	1430.71	**1350.05**	1366.43	1350.05
ADM	803.90	288.33	**259.99**	491.25	578.23	506.45	597.53	602.66	626.94	259.99
FOM	194.49	182.37	**174.51**	204.42	208.83	196.70	186.33	187.06	176.96	174.51
Connectivity	**23.29**	41.54	46.71	56.42	64.22	60.12	84.83	90.83	92.24	23.29
Dunn	0.15	0.08	0.08	0.07	0.08	**0.26**	0.06	0.06	0.06	0.26
Silhouette	**0.43**	0.36	0.25	0.24	0.23	0.25	0.14	0.14	0.17	0.43

Table 4 summarizes the index-wise optima; by majority rule anchored in interpretability, we adopt $K = 4$ for the trimmed stage.

Table 4. Optimal scores of cluster validity indices and number of clusters for trimmed k-means

Cluster Validation	Score	Clusters
APN	0.09	**4**
AD	1350.05	9
ADM	259.99	**4**
FOM	174.51	**4**
Connectivity	23.29	2
Dunn	0.26	7
Silhouette	0.43	2

Figure 4 shows the spatial imprint of the trimmed partition. After trimming the high-leverage outlier (Rize), the remaining stations resolve into four structural clusters (C1–C4) following clear geographic controls: C1 comprises humid/orographic margins (notably the eastern–central Black Sea and parts of the southwestern Mediterranean) and has the highest annual totals; C2 forms a transitional arc between maritime and continental influences, with intermediate totals and stronger seasonality; C3 links segments of the Marmara–Aegean corridor with selected eastern uplands, grouping stations that exhibit comparable annual magnitudes despite differing local drivers; C4 spans the continental interior of Central, Eastern, and Southeastern Anatolia and delineates the network's dry backbone, which has the lowest annual totals. The trimmed singleton (Rize) is displayed separately for cartographic clarity, but it belongs to the trimmed set rather than to a structural cluster. Some resulting assignments may appear climatologically counterintuitive at first glance; for instance, stations such as Bitlis and Bingöl appear in the same annual-magnitude cluster as Antalya and Muğla, and Trabzon clusters with inland stations such as Tunceli and Mardin. These groupings arise because our distance metric is based solely on standardized annual totals: when we inspect monthly climatologies for these stations (not shown), they exhibit clearly different seasonal regimes, yet similar long-term annual sums. In other words, the trimmed k-means solution primarily captures a first-order gradient in the magnitude of annual precipitation. A finer separation of regimes based on intra-annual distribution would require extending the panel-clustering framework to seasonal or monthly vectors and/or multivariate hydroclimate variables, which we view as a natural direction for future work.

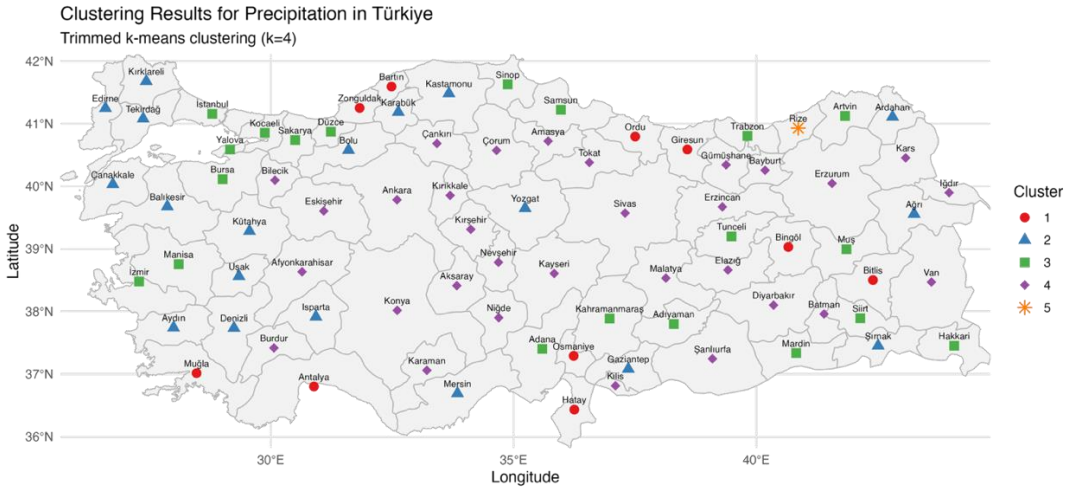

Figure 4. Station-level annual precipitation clusters over Türkiye from trimmed k-means

Figure 5 presents the corresponding ordination: the four groups are tighter, less elongated, and show reduced overlap; centroids shift away from the Rize-dominated axis toward the cluster cores, indicating that trimming reveals substructure masked in the classical solution. The first dimension orders stations primarily by annual magnitude, while the weaker second dimension behaves as a seasonality-related gradient. Together, the map and ordination show that the robust procedure accomplishes what the classical $K = 2$ solution could not: it neutralizes the extreme station's leverage, sharpens internal cohesion, and yields four interpretable station-level precipitation regimes (plus the trimmed singleton), consistent with the evidence in Tables 3–4 and with known controls on Türkiye's precipitation climatology (coastal/orographic enhancement vs. continental drying).

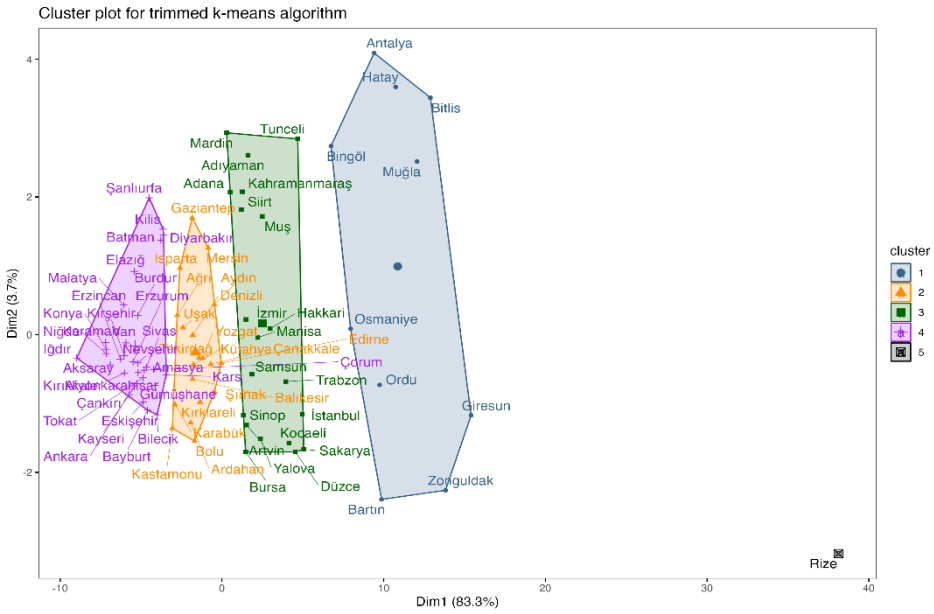

Figure 5. Scatter plots of stations clustered using the trimmed k-means algorithm

4. Conclusions

Using TSMS station- level annual precipitation series (1965–2022), we compared three classical clustering methods (k-means, PAM, and hierarchical (agglomerative)) with a robust alternative (trimmed k-means) using a seven-index validation battery (APN, AD, ADM, FOM, Connectivity, Dunn,

and Silhouette). In the classical stage, k-means with $K = 2$ achieved high Silhouette/Dunn and low Connectivity; however, the emergence of a singleton cluster (Rize) showed that an outlier with strong leverage dominated the partition and masked meaningful substructure. The trimmed k-means solution attenuated this leverage and revealed four structural precipitation regimes (C1–C4) that are more compact, less overlapping, and climatologically interpretable. Rize, sustained by persistent orographic uplift and moisture supply, appears as a trimmed-singleton feature; removal of its influence exposes the finer organization suppressed in the classical solution.

The trimmed k-means regimes (C1–C4) align with the principal controls of Türkiye's precipitation climatology. C1 unites the eastern–central Black Sea and parts of the southwestern Mediterranean orographic margins; despite contrasting seasonality (year-round wetness along the Black Sea versus a winter maximum in the Mediterranean), stations converge along the axis of annual magnitude. C4 represents the continental interior of Central/Eastern/Southeastern Anatolia—the network's dry backbone—with low annual totals and high aridity potential. C2–C3 delineate transition belts where maritime and continental influences interact; annual totals are similar, but seasonality diverges — with more pronounced summer dryness in the west and larger cold-season contributions in the east. Overall, the four regimes quantify—at the station scale—the well-known dichotomy between coastal/orographic enhancement and continental drying. At the same time, the focus of our design on annual magnitude explains why some station groupings may diverge from expectations based solely on seasonal precipitation regimes. For example, certain high-precipitation coastal stations in the eastern Black Sea (e.g., Rize and its neighbors) share similar long-term annual totals with parts of the southwestern Mediterranean coast, even though their intra-annual distributions differ markedly. Likewise, stations such as Bitlis and Bingöl can appear in the same annual-magnitude cluster as Antalya and Muğla, while Trabzon can cluster with inland stations like Tunceli and Mardin because our distance metric does not explicitly encode month-to-month seasonality. In these cases, the clustering reflects similarity in long-term annual sums rather than in the detailed timing of rainfall within the year. We therefore interpret the C1–C4 regimes as first-order annual precipitation zones, and view a full seasonal or multivariate regime classification—incorporating monthly totals, temperature, humidity, and circulation indices—as a complementary extension rather than a substitute for the present analysis.

Our results broadly accord with earlier work while highlighting methodological contrasts. The seven Ward-based regions in Ünal et al. (2003) isolate the continental interior, echoing C4; our annual robust approach, however, merges the very wet Black Sea sector with the SW Mediterranean orography into C1, offering a more parsimonious regime map. Türkeş and Tatlı (2011) identified eight precipitation zones that separate the Black Sea, Mediterranean, and continental belts; we retain the core C1–C4 contrast, while coastal micro-gradients that their spectral method resolves are absorbed into C2–C3 transitions under annual aggregation. Kartal et al. (2011) showed that Ward clustering can deviate from administrative regions; our trimmed solution addresses this issue by reducing sensitivity to outliers and yielding tighter clusters. Fuzzy approaches (Dikbaş et al., 2012; Zeybekoğlu & Ülke Keskin, 2020) have demonstrated that geographically distant stations can share a regime; our C1–C4 similarly group peers of similar annual magnitude across space (e.g., the SW Mediterranean with the wettest Eastern Black Sea). Kömüşçü et al. (2022) emphasized NAO/MO modulation across five regimes. In our ordination, Dim1 captures the magnitude, and the weaker Dim2 implies differences in seasonality; teleconnection effects are best tested at seasonal or monthly resolution.

The derived station regimes provide spatially coherent units for water-resources management (runoff, storage and conveyance sizing), agricultural planning (cropping patterns, irrigation demand, and drought-tolerant varieties), infrastructure design (stormwater, flood, and landslide protection—especially in the Eastern Black Sea), and risk management. Priorities differ by regime: C1 emphasizes flood and landslide mitigation and enhanced coastal drainage capacity; C4 emphasizes water-use-efficiency, irrigation modernization, and drought adaptation; C2–C3 call for balancing storage–demand asymmetries arising from seasonality. Hydropower operations (managing flood/drought asymmetries) and stormwater standards in rapidly urbanizing coastal regions benefit directly from this framework.

Annual aggregation can partially conflate distinct seasonal behaviors; station density is heterogeneous, and metadata issues and residual inhomogeneities may introduce bias. Moreover, by design, we work with a univariate annual panel on a core set of 81 long-record stations; this choice facilitates reproducibility but inevitably leaves out shorter or less homogeneous series that could further refine spatial gradients. In addition, some stations with similar long-term annual totals but contrasting seasonal regimes may be grouped together, as our distance metric does not explicitly encode intra-annual timing of precipitation. Future research should (i) extend panel-clustering to seasonal/monthly scales and jointly model teleconnections (e.g., NAO, MO), (ii) test fuzzy–robust hybrids and spatially constrained algorithms (neighborhood/contiguity), (iii) integrate topographic/remote-sensing covariates (elevation, coastal distance), and (iv) quantify uncertainty (resampling, cross-validation; cluster stability) and evaluate alternative distance metrics. Such advances will strengthen station-scale evidence for climate-adaptation policy and translate into more resilient, data-driven decisions.

Acknowledgements and Additional Information

The authors would like to thank the Turkish State Meteorological Service for providing the data used in this study.

Author Contributions

Author Name	Contribution Types
Çağlar Sözen	Research Design, Data Collection, Analysis, Methodology, Writing, Editing.
Hasan Bulut	Research Design, Data Collection, Analysis, Methodology, Writing, Editing.

Conflict of Interest

The authors declare that they have no conflict of interest. They also declare that there is no commercial or financial relationship that could be construed as a potential conflict of interest related to this study.

Ethics Statement

This study was conducted using publicly available meteorological data and did not involve any human participants or animal subjects. Consequently, ethical approval from an institutional review board was not required. All analyses were carried out following applicable guidelines and regulations.

Referanslar/References

Aggarwal, C. C. & Reddy, C. K. (Eds.). (2014). *Data clustering: Algorithms and applications*. Boca Raton, FL: CRC Press.

Brock, G., Pihur, V., Datta, S. & Datta, S. (2008). clValid: An R package for cluster validation. *Journal of Statistical Software*, 25(4), 1-22. https://doi.org/10.18637/jss.v025.i04

Bulut, H. (2023). *Multivariate statistical methods with R applications [R uygulamaları ile çok değişkenli istatistiksel yöntemler]* (2. bs.). Ankara: Nobel Akademik Yayıncılık.

Cuesta-Albertos, J. A., Gordaliza, A. & Matrán, C. (1997). Trimmed k-means: An attempt to robustify quantizers. *The Annals of Statistics*, 25(2), 553-576.

Demircan, M., Gürkan, H., Eskioğlu, O., Arabacı, H. & Coşkun, M. (2017). Climate change projections for Turkey: Three models and two scenarios. *Turkish Journal of Water Science and Management*, 1(1), 22-43. https://doi.org/10.31807/tjwsm.297183

Deniz, Z. A., Gönençgil, B., & Mestav, B. (2016). Türkiye ekstrem sıcaklıklarının kümeleme analizine göre değerlendirilmesi. *TÜCAUM Uluslararası Coğrafya Sempozyumu Bildiriler Kitabı* içinde (ss. 220-229). Ankara: TÜCAUM.

Dikbas, F., Firat, M., Koc, A. C., & Gungor, M. (2012). Classification of precipitation series using fuzzy cluster method. *International Journal of Climatology*, 32(10), 1596-1603. https://doi.org/10.1002/joc.2350

García-Escudero, L. A., Gordaliza, A., Matrán, C. & Mayo-Iscar, A. (2008). A general trimming approach to robust cluster analysis. *The Annals of Statistics*, 36(3), 1324-1345. https://doi.org/10.1214/07-AOS515

Hubert, L. & Arabie, P. (1985). Comparing partitions. *Journal of Classification*, 2(1), 193-218. https://doi.org/10.1007/BF01908075

Intergovernmental Panel on Climate Change (IPCC). (2021).*Climate change 2021: The physical science basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (V. Masson-Delmotte, P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, ... & B. Zhou, Eds.). Cambridge University Press.

Iyigun, C., Türkeş, M., Batmaz, İ., Yozgatligil, C., Purutçuoğlu, V., Koç, E. K. & Öztürk, M. Z. (2013). Clustering current climate regions of Turkey by using a multivariate statistical method. *Theoretical and Applied Climatology*, 114(1), 95-106. https://doi.org/10.1007/s00704-012-0823-7

Jain, A. K., Murty, M. N. & Flynn, P. J. (1999). Data clustering: A review. *ACM Computing Surveys (CSUR)*, 31(3), 264-323. https://doi.org/10.1145/331499.331504

Kartal, E., İyigün, C., Fahmi, F. M., Yozgatlıgil, C., Purutçuoğlu, V., Batmaz, İ., ... & Türkeş, M. (2011). Türkiye iklim bölgelerinin hiyerarşik kümeleme yöntemi ile belirlenmesi. *İstatistik Araştırma Dergisi*, 8(1), 13-25. https://izlik.org/JA75EU78ZY

Kaufman, L., & Rousseeuw, P. J. (1990). *Finding groups in data: An introduction to cluster analysis*. New York, NY: John Wiley & Sons.

Kömüşcü, A. Ü., Turgu, E. & DeLiberty, T. (2022). Dynamics of precipitation regions of Turkey: A clustering approach by K-means methodology in respect of climate variability. *Journal of Water and Climate Change*, 13(10),3578-3606. https://doi.org/10.2166/wcc.2022.186

MacQueen, J. (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* (Vol. 1, pp. 281–297). Berkeley, CA: University of California Press.

Maechler, M. (2018). *cluster: Cluster analysis basics and extensions* (R package version 2.0-7–1).

Mahmud, S., Sumana, F. M., Mohsin, M., & Khan, M. H. R. (2022). Redefining homogeneous climate regions in Bangladesh using multivariate clustering approaches. *Natural Hazards*, 111(2), 1863-1884. https://doi.org/10.21203/rs.3.rs-633865/v1

Rencher, A. C. (2002). *Methods of multivariate analysis* (2nd ed.). New York, NY: John Wiley & Sons.

Taşoğlu, E., Öztürk, M. Z. & Yazıcı, Ö. (2024). High resolution Köppen–Geiger climate zones of Türkiye. *International Journal of Climatology*, 44(14), 5248–5265. https://doi.org/10.1002/joc.8635

Team, R. C. (2023). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Tekin, M. K., Tatli, H., & Koç, T. (2021). Climate classification in Turkey: A case study evaluating Holdridge life zones. *Theoretical & Applied Climatology*, 144. https://doi.org/10.1007/s00704-021-03565-5

Trenberth, K. E. (2011).Changes in precipitation with climate change. *Climate Research*, 47(1-2), 123-138. https://doi.org/10.3354/cr00953

Türkeş, M. (1996). Spatial and temporal analysis of annual rainfall variations in Turkey. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 16(9), 1057-1076. https://doi.org/10.1002/(SICI)1097-0088(199609)16:9 <1057::AID-JOC75>3.0.CO;2-D

Türkeş, M. & Tatlı, H. (2011). Use of the spectral clustering to determine coherent precipitation regions in Turkey for the period 1929–2007. *International Journal of Climatology*, 31(14), 2055-2067. https://doi.org/10.1002/joc.2212

Turkish State Meteorological Service (2023). *Annual precipitation data for 81 provinces in Turkey (1965–2022)* [Data set]. Ankara: Turkish State Meteorological Service.

Unal, Y., Kindap, T. & Karaca, M. (2003). Redefining the climate zones of Turkey using cluster analysis. *International Journal of Climatology*, 23(9), 1045-1055. https://doi.org/10.1002/joc.910

Wilks, D. S. (2011). *Statistical methods in the atmospheric sciences*. Oxford: Academic Press.

Zeybekoğlu, U. & Ülke Keskin, A. (2020). Defining rainfall intensity clusters in Turkey by using the fuzzy c-means algorithm. *Geofizika*, 37(2), 181–195. https://doi.org/10.15233/gfz.2020.37.8