DOI: 10.24850/j-tyca-14-06-02

Articles

Prioritization of watersheds for soil and water conservation based on GIS, PCA and WSA techniques

Priorización de cuencas hidrográficas para la conservación del suelo y el agua basado en las técnicas GIS, PCA y WSA

Maiquel López-Silva[1], ORCID: https://orcid.org/0000-0002-0946-6160

Dayma Carmenates-Hernández[2], ORCID: https://orcid.org/0000-0001-5482-7562

[1]Engineering Faculty, Catholic University Sedes Sapientiae, Peru, mlopezs@ucss.edu.pe

[2]Engineering Faculty, Catholic University Sedes Sapientiae, Peru, dcarmenates@ucss.edu.pe

Corresponding author: Maiquel López-Silva, mlopezs@ucss.edu.pe

Abstract

Soil and water conservation is a priority in the watersheds of arid and semi-arid regions for the proper planning and integrated management of water resources. The objective of this work is prioritizing 91 watersheds in 14 regions of Peru with different geographical, hydrological and geological conditions, through the integration of Geographic Information Systems (GIS), Principal Component Analysis (PCA) and the Weighted Sum Approach (WSA). In addition, homogeneous regions were identified with hierarchical Cluster Analysis in R and Ward's method. The result showed the existence of 19.49 % of total area in high and very high priority category in two regions. The results of the Cluster Analysis showed that 35 % of the watersheds have homogeneous zones within the geometric and shape factor, while 65 % within the drainage and relief factor. In general, GIS, PCA and SWA methods are an efficient tool, which allows decision-making authorities for better planning and conservation of natural resources in hydrographic watersheds.

Keywords: Erosion, morphometric, drainage, cluster analysis, factor analysis.

Resumen

La conservación del suelo y el agua es una prioridad en las cuencas hidrográficas de las regiones áridas y semiáridas para la adecuada planificación y gestión integrada de los recursos hídricos. El objetivo de este trabajo fue priorizar 91 cuencas hidrográficas en 14 regiones del Perú con diferentes condiciones geográficas, hidrológicas y geológicas, mediante la integración de sistemas de información geográfica (SIG), análisis de componentes principales (PCA) y el análisis de suma ponderada (WSA). Se identificaron regiones homogéneas con el análisis de conglomerados jerárquico en el método de R y Ward. El resultado mostró la existencia de un 19.49 % del área total en categoría de alta y muy alta prioridad en dos regiones. Los resultados del análisis de conglomerados mostraron que el 35 % de las cuencas hidrográficas tiene zonas homogéneas dentro del factor geométrico y de forma, mientras que el 65 % dentro del factor de drenaje y relieve. En general, los métodos SIG, PCA y SWA son herramientas eficientes, que permiten a las autoridades encargadas la toma de decisiones para una mejor planificación y conservación de los recursos naturales en las cuencas hidrográficas.

Palabras clave: erosión, morfometría, drenaje, análisis de conglomerados, análisis factorial.

Received: 08/11/2021

Accepted: 12/04/2022

Published online: 07/07/2022

Introduction

Peru, globally, is one among the countries most affected by variability and global climate change (Vega, Lavado, & Felipe, 2018) impacting on the spatial distribution of water resources in Peru's Andes, a fundamental source of water supply for the population, agriculture and energy production (Wongchuig, Mello, & Chou, 2018). In this context, there is a requirement for sustainable watershed management (Bhattacharya, Chatterjee, & Das, 2019) ranging from an efficient morphometric characterization (Ghosh & Gope, 2021) with new scientific tools to enable adaptation to variability and global climate change with attention on prediction until the solution of problems (Jacobs & Brian, 2020).

The Principal Component Analysis (PCA) method allows within the various morphometric variables to explain the most significant part of the variance of the data (Helness, Damman, Sivertsen, & Ugarelli, 2019) to identify cause - effect relationships; zones with homogeneous climate or similar hydrological characteristics. This analytical approach groups similar data while identifying relationships between variables (Gorgoglione, Gioia, & Iacobellis, 2019) precisely to identify the most dominant ones in hydrological processes (Balbín *et al*., 2020). Subyani, Qari and Matsah (2012) performed the PCA and Cluster Analysis (CA) but only to ten watersheds of different sizes in the western region of Saudi Arabia with 18 morphometric variables. Yunus, Oguchi and Hayakawa (2014) presented the geomorphological quantification of mountainous terrain, but limited to 36 watersheds in the Western Arabian Peninsula.

Research such as that of Gajbhiye and Sharma (2017) is based on the prioritization of watersheds based on the PCA and the Composite Factor (CF), this includes the calculation of the simple arithmetic average of the final priority rankings, but they have the disadvantage that they assume the same degree of importance to all morphometric parameters (Aher, Adinarayana, & Gorantiwar, 2014; Bharath *et al*., 2021). The aforementioned aspect is contradictory because each watershed has its own characteristics of shape, relief and hydrographic network. Recent research by Malik *et al*. (2019), and Setiawan and Nandini (2021) have shown strategic optimism between PCA and WSA techniques for watershed prioritization. The WSA method is known because it offers consistency in land surface analysis that contributes to watershed prioritization. According to Singh and Singh (2018) emphasize that WSA is a traditional watershed prioritization method compared to the rest and provides dynamic and efficient results. New integration methods are essential that allow better management of hydrographic watersheds in Peru, due to its climatic diversity and fragile ecosystems. Therefore, the objective of the research is to prioritize watersheds for soil and water conservation based on GIS, PCA and WSA techniques in different regions of Peru.

Materials and methods

Fourteen regions with different hydrographic, hydrological, geological and ecological conditions were selected as the study area for the 91 watersheds. These watersheds are located in two of the country's largest watersheds, the Pacific and the Atlantic. The digitized watersheds were processed in the GIS of the ArcMap 10.5 software, and based on the formulas in Table 1, the morphometric parameters of the watersheds can be calculated. The characterization of the watersheds will be carried out in qualitative and quantitative terms that are grouped into surface, shape, drainage and relief parameters.

Table 1. Formulas of the morphometric parameters of the watersheds.

Parameter	Equation	Units	References
watersheds amplitude	$W = AL_c^{-1}$	km	Horton (1932)
Slope of the watersheds	$S_c = e\sum l_i A^{-1}$	%	Zavoianu (1985)
Gravelius compactness coefficient	$K_c = 0.282PA^{-0.5}$	-	Zavoianu (1985)
Form factor	$F_f = AL_c$	-	Horton (1932)
Shape index	$S_w = L_c^2 A^{-1}$	%	Horton (1945)
Elongation ratio	$R_e = 1.1284A^{0.5}L_c^{-1}$	-	Schumm (1956)
Circularity ratio	$R_c = 4\pi AP^{-2}$	-	Miller, Ritter and Kochel (1990)
Relief ratio	$R_a = HL_c^{-1}$	-	Schumm (1956)
Mean stream slope	$S_r = \Delta HL_r^{-1}$	%	Zavoianu (1985)
Drainage density	$D_d = L_u A^{-1}$	km/km²	Horton (1945)
Mean length of overland flow	$L_f = 0.5D_d^{-1}$	km	Horton (1945)
Constant channel maintenance	$C = 1D_d^{-1}$	km/km²	Schumm (1956)
Roughness coefficient	$C_r = D_d S_r$	-	Horton (1945)

Principal component analysis

It is a difficult and complicated process to classify potential areas with more benefits from water resources in 91 watersheds. The aim of principal component analysis (PCA) is to explain the variance-covariance structure in multiple data sets using a few linear combinations of the original variables, according to Kottegoda and Rosso (2008). The PCA technique is designed to transform p correlated *X* variables, which are known, into an equal number of uncorrelated *Z* indices. PCA is a technique for summarizing a complex set of data, distinguishing quantitative dependent and independent variables, as well as identifying closely related variables. Matrix notation in its general form is represented using Equation (1). In which *Z* and *X* are n×p matrices and A is a p×p matrix of coefficients:

$$Z_j = X_{aj} \; to \; j = 1,2, \dots p \qquad (1)$$

Where Z_j is an n×1 (column) vector and a_j is a p×1 (column) vector of coefficients., j of 1, 2, ... *p*.

Considering that each column is an observed variable in rigor of n answers to the same question, one can study the variance and covariance between different variables, represented by a square *C* matrix of order p×p symmetric (Yang *et al*., 2020). The variance-covariance matrix *X* is represented using equations (2) and (3). In practice is estimated by the sample covariance matrix *C*:

Tecnología y ciencias del agua, ISSN 2007-2422, 14(6), 40-75. DOI: 10.24850/j-tyca-14-06-02

$$C = \frac{X^T X}{(n-1)} = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & & c_{2p} \\ \vdots & & \ddots & \vdots \\ c_{p1} & c_{p2} & \cdots & c_{pp} \end{pmatrix} \tag{2}$$

$$c_{ij} = \frac{1}{n-1} \sum_{k=1}^{n} X_{ij} X_{ki} \tag{3}$$

Where *C* becomes the sample correlation matrix; X^T is the transpose of the standardized *X* variables, and c_{ij} is the value of each element of the covariance.

In the PCA, the variance of the Principal Component (PC) is maximized, which is mainly related to the eigenvalues of the sample covariance or correlation matrix (Malik *et al*., 2019; Kottegoda & Rosso 2008). The correlation r between the X_{ij} of the PC is obtained by the ratio of the covariance of the *X* matrix and the variance of Z_j of the PCs using Equation (4):

$$r(x_i; z_j) = \frac{Cov(x_i\,; z_j)}{\sqrt{Var(z_j)}} \tag{4}$$

Where *r* is the correlation coefficient; Cov (x_i, z_j) is the covariance of *X* variables in the *i*th row and *j*th row of the PCs; and, Var (z_j) is the variance of *j*th PCs. Value of r greater than 0.90 indicates strong correlation, *r* between 0.75 and 0.90 indicates good correlation, and *r* between 0.60 and 0.75 indicates moderate correlation (Sharma, Gajbhiye, & Tignath, 2015).

Factor analysis

Factor analysis (FA) is a procedure similar to principal components analysis to identify physically significant variables and reduce morphometric parameters. The analysis via this technique produces easily interpretable results. FA techniques are well described in the literature (Finch, 2020). The correlation between the variables is related to the factor loadings, considering correspondence between the factor score matrix and the Z_s in the PCs, the orthogonal in the rotation has an impact on the PCs and the corresponding factors are equivalent.

Therefore, a comparison with the unrotated components is possible, while the rotation should be such that meaningful physical interpretations are possible from the resulting components obtained (Jolliffe, 2002). Thus, the rotated components are given by using Equation (5):

$$Z_r = XS^{-1}RPD^{-0.5} \quad (5)$$

Where Z_r is the rotated components; $D^{-0.5}$ is a diagonal matrix in which the nonzero elements are the square roots of the eigenvalues of the correlation matrix S expressed by Equation (2), and RP the orthogonal rotation.

Tecnología y ciencias del agua, ISSN 2007-2422, 14(6), 40-75. DOI: 10.24850/j-tyca-14-06-02

Cluster analysis

The hierarchical Cluster Analysis (CA) technique was applied to classify the homogeneous groups called clusters. The method applied for data agglomeration was in R mode and the Ward's criterion (Ward, 1963), one of the most widely used in hydrology and meteorology (Cupak, Wałęga, & Bogusław, 2017), with the objective of achieving similarity between watersheds based on morphometric parameters with a focus on water resource and soil.

Weighted sum approach

WSA was applied to the most significant morphometric variables that are identified with PCA, WSMP or morphometric parameter weight was used in cross-correlation analysis (Siddiqui, Said, & Shakeel, 2020) whereas, Composite Factor (CF) was evaluated for final priority ranking and related category, which are expressed according to Singh and Singh (2018), Malik *et al*. (2019), and Setiawan and Nandini (2021) as:

$$W_{SMP} = \frac{Sum\ of\ correlation\ coefficient}{Total\ of\ correlation} \quad (6)$$

$$CF = PR_{SMP}W_{SMP} \quad (7)$$

Where *CF* is the composite factor; PR_{SMP} is the preliminary ranking of the most significant morphometric parameter of the PCA, and W_{SMP} is the Weight of the significant morphometric parameter.

Based on the *CF* value, the watershed was assigned a priority ranking. The lowest *CF* value belongs to priority ranking one, the second lowest value to priority ranking belongs to priority ranking two and so on for all watersheds. The study used the SPSS 26.0 statistical package and STATISTICA 12 as the computational tool for the PCA.

Results and discussion

The results of the statistical analysis in Table 2 of the morphometric shape parameters such as compactness coefficient, shape factor, elongation ratio and circularity factor indicate the existence of elongated and circular watersheds for their mean values between 0.31 to 1.42, with round oval watersheds prevailing according to the criteria by Horton (1932) and Schumm (1956). In the geological aspect, the circularity factor ranged from 0.24 to 0.73, the high results are indicative of rock structures that are impermeable and non-homogeneous and that control drainage.

Table 2. Statistical analysis of morphometric parameters.

Parameters	Min	Max	Mean	Sd	Cv
Z_{max}	124.45	5 500.00	3 698.89	1 724.57	46.62
Z_{min}	4.89	4 013.00	1 516.20	1 183.44	78.05
H	45.56	5 429.55	2 182.67	1 395.39	63.93
A	45.90	1 496.85	459.69	372.18	80.96
P	33.02	258.12	101.92	52.75	51.75
L_c	20.40	268.38	85.02	12.60	35.98
L_r	49.53	1467.02	449.55	364.45	81.07
S_c	6.90	50.99	23.34	13.70	58.69
W	2.22	21.89	11.30	5.57	49.26
K_c	1.17	2.06	1.42	0.18	12.49
F_f	0.10	0.46	0.31	0.09	28.32
S_w	2.17	9.90	3.68	1.74	47.47
R_e	0.36	0.77	0.62	0.01	15.81
R_c	0.24	0.73	0.51	0.11	21.09
R_a	0.82	189.18	66.99	44.38	66.25
S_r	2.36	37.31	11.63	8.90	76.53
D_d	0.98	3.05	1.88	0.39	20.81
L_f	0.16	0.51	0.28	0.07	23.78
C	0.33	1.02	0.56	0.13	23.46
C_r	1.49	143.78	23.91	25.89	108.28

The rugged relief is evidenced by average slopes of 23.34 %, classified as mountainous watersheds prone to dominant environmental processes such as deterioration of soil fertility, erosion and sediment transport in watercourses according to Kumar, Das and Das (2020). According to Amare, Kassie and Sulla (2020) significant variations in relief influence the distribution of rainfall and other climatic factors. The drainage density in the hydrographic watersheds ranges from 0.98 km/km^2 to 3.05 km/km^2 for an average of 1.88 km/km^2. According to the categorization of Horton (1945), the watersheds have a medium drainage trend and consequently a medium permeability and infiltration of water; as well as, moderate tendencies to avenues that induce floods. Results that are in correspondence with those obtained by Mahala (2019), but in the Kosi River watersheds of North India for the mountainous flat tropical environment.

In general, these are watersheds with a greater probability of using water resources spatially and temporally; therefore, there is a high need for hydraulic works to mitigate the problems of flooding and frequent landslides in the watersheds of Peru's territory. Table 3 and Table 4 show the results of the Preliminary Ranking (PR_{smp}) based on morphometric parameters by region.

Tecnología y Ciencias del Agua

OPEN ACCESS

Check for updates

Table 3. Preliminary priority ranking based on watersheds.

Parameters	Amazonas	Ancash	Apurimac	Cajamarca	Cusco	Huancavelica	Huanuco	Junin	La Libertad	Lima	Loreto	Pasco	San Martin	Ucayali
Z_{max}	2	13	7	9	11	8	14	6	4	10	12	1	3	5
Z_{min}	3	9	11	13	14	10	4	7	2	8	12	1	5	6
H	2	13	1	7	5	9	14	6	10	12	11	4	3	8
A	2	7	12	9	11	5	8	13	3	1	14	4	6	10
P	2	6	13	11	9	5	8	14	3	1	12	4	7	10
L_c	2	6	12	9	11	5	8	13	3	1	14	4	7	10
L_r	2	8	13	7	12	5	4	14	3	1	9	4	6	6
W	1	5	12	8	14	3	11	13	4	2	9	6	7	10
S_c	11	9	8	2	7	12	14	13	3	1	5	4	6	10
K_c	4	4	12	9	3	8	5	10	2	7	13	1	6	11
F_f	3	5	11	6	13	2	14	7	8	1	4	12	10	9
S_w	14	11	4	7	1	13	3	6	8	12	10	2	9	5
R_e	3	4	11	6	13	2	13	7	8	1	5	14	9	10
R_c	11	10	2	7	12	6	9	3	13	5	1	14	8	4
R_a	8	12	1	6	3	9	14	2	11	13	10	7	4	5
D_d	9	11	13	3	14	12	1	4	6	5	8	10	2	7
L_f	6	3	4	12	1	2	14	11	9	8	7	5	13	10
C	6	3	4	12	1	2	14	11	9	8	7	5	13	10
C_r	14	9	8	7	11	13	3	5	12	6	4	10	1	2

Table 4. Preliminary priority ranking based on regions.

Regions	**CP**	**PR_{SMP}**
Amazonas	5.53	2
Ancash	7.79	8
Apurimac	8.37	10
Cajamarca	7.89	9
Cusco	8.74	12
Huancavelica	6.89	5
Huanuco	9.21	14
Junin	8.68	11
La Libertad	6.37	4
Lima	5.42	1
Loreto	8.79	13
Pasco	5.89	3
San Martin	6.58	6
Ucayali	7.78	7

The PCA was started by obtaining the matrix of the simple linear Pearson correlation coefficient (r), there is a strong correlation ($r > 0.9$) between the morphometric parameters of 1) A, P; A, L_c; A, L_r and A, W; 2) P, L_c; P, L_r and P, W; 3) L_c, L_r and L_c, W; 4) L_r and W; 5) K_c and R_c; 6)

F_f, S_w and F_f, R_e; 7) S_w; R_e; 8) D_d, L_f and D_d, C; and 9) L_f and C_c. While, there is a good correlation ($0.75 < r < 0.9$) between 1) H_m, R_a and 2) S_r, C_r. In addition, there are moderately acceptable Pearson correlation coefficients ($0.60 < r < 0.75$) at 1) $Z_{\max}$, $Z_{\min}$; $Z_{\max}$, R_a; 2) P, K_c; P, R_c; 3) W, F_f; W, S_w and W, R_e. The parameters of S_c and S_r did not show a relationship with the rest to the other variables that explain their effect, with except for C_r. Similar results achieved Yunus *et al*. (2014) in watersheds of the Western Arabian Peninsula, and Mokarram and Sathyamoorthy (2015) in hydrographic watersheds of Iran fundamentally in the parameters of A, L_c, R_e, F_f, R_c and D_d with strong correlation. While, with a certain approximation by those of Malik *et al*. (2019) strongly correlated to L_f, F_f, R_c, K_c, R_e and differs with D_d in Indian watersheds. Based on the significant correlations, no drainage correlations are observed with the surface variables, which are in agreement with those reached by Sharma *et al*. (2015). Other authors, such as Gede, Anwar and Lasminto (2017), demonstrated that A and L_r influence the maximum flow and the unit hydrograph at its peak time with a correlation coefficient of 0.98, an aspect that could be related in the current research.

The application of the FA to the twenty morphometric parameters in the first iteration, six PCs were defined with eigenvalues greater than one and that represent around 88.26 % of the total variance. However, the morphometric parameters with weak loads (S_c and S_r) were extracted to improve the FA. Under these conditions, the same factor load was reached with 6 PC that explain 94.58 % of the total variance. The Varimax rotation showed improvement in the loads, Table 6, which is because the matrix of the rotated PC was considered.

Tecnología y ciencias del agua, ISSN 2007-2422, 14(6), 40-75. DOI: 10.24850/j-tyca-14-06-02

In Table 5, the factor load of PC1 constitutes 26.56 % of the total variance in the rotated matrix, the second PC2 17.87 % and the third PC3 explains 16.82 % of the total variance, having these first three components a strong charge of 60.95 %; while, from PC4 to PC6, factor loadings only explain 33.63 % of the total variance.

Table 5. Factor loading matrix of reduced morphometric variables (Rotated Varimax).

Parameters	**Principal components**					
	PC1	**PC2**	**PC3**	**PC4**	**PC5**	**PC6**
Z_{max}	0.15	0.09	-0.10	0.75	-0.03	0.63
Z_{min}	-0.01	0.04	-0.13	-0.05	-0.07	0.98
H	0.18	0.08	-0.01	0.97	0.02	-0.05
A	0.93	0.26	-0.03	0.01	0.21	0.01
P	0.83	0.32	-0.02	0.02	0.43	-0.01
L_c	0.94	0.18	-0.03	0.01	0.21	0.01
L_r	0.93	0.27	-0.04	0.01	0.20	-0.01
W	0.83	0.52	0.01	0.02	0.17	0.01
K_c	0.32	0.01	-0.05	-0.01	0.93	-0.06
F_f	0.28	0.94	0.06	0.03	0.01	0.02
S_w	-0.26	-0.93	-0.05	-0.06	-0.01	-0.05
R_e	0.29	0.95	0.06	0.03	0.01	0.02
R_c	-0.28	0.01	0.04	0.07	-0.92	0.03
R_a	-0.28	-0.02	0.03	0.93	-0.09	-0.04
D_d	0.02	-0.08	-0.97	-0.01	0.08	0.07
L_f	0.01	0.04	0.99	-0.03	-0.01	-0.05
C	-0.01	0.03	0.98	-0.03	-0.01	-0.05
C_r	-0.50	0.04	-0.15	0.34	0.28	-0.14

Tecnología y ciencias del agua, ISSN 2007-2422, 14(6), 40-75. DOI: 10.24850/j-tyca-14-06-02

The factor load of PC1 in Table 5 is strongly correlated with the factor loadings of A, P, L_c, L_r and W, which could be called geometric factors that explain 26.56 % of the total variance of the 18 parameters analyzed. The second PC2 has a strong correlation with the variables of F_f, S_w and R_e, which is why it is called the form factor of the hydrographic watersheds that explains 17.87 % of the total variance; while, the PC3 load factor, its strong correlation is with the D_d, L_f and C called drainage factor that explains 16.52% of the total variance. The fourth factor load of PC4 has a strong correlation with the variables of $Z_{\max}$ and H_m, which is why it is called the relief factor of the watersheds that explains 13.82 % of the total variance. The first four components account for 74.77 % of the total variance, which explains the most important morphometric parameters and the most significant physical factors. Being very useful, the PCA to define the most relevant variables. Yunus *et al*. (2014) achieved results similar to the current ones, in 36 watersheds in which three CPs explained 73 % of the total variance that strongly reflected the dimensions and surface of the watershed, as well as the drainage texture. In the same sense, Sharma *et al*. (2015) three PCs defined the drainage, slope and shape; later with similar results it was by Malik *et al*. (2019) with strong correlation in shape and drainage parameters.

Table 6 shows that there are variables with a strong correlation in each PC. These parameters are considered the most significant for applying the WSA model and in turn the prioritization of the watersheds. In the final prioritization of the watersheds by region, the composite value was calculated based on the preliminary ranking and weight of the six

Tecnología y ciencias del agua, ISSN 2007-2422, 14(6), 40-75. DOI: 10.24850/j-tyca-14-06-02

variables L_c, R_e, L_f, H, K_c and Z_{min} with the cross-correlation analysis Table 6. The composite value was determined using Equation (8).

$$CF = 0.26L_c + 0.23R_e + 0.10L_f + 0.16H + 0.18K_c + 0.08Z_{min} \quad (8)$$

Table 6. The cross-correlation between the important parameters.

Parameter	L_c	R_e	L_f	H	K_c	Z_{min}
L_c	1.00	0.43	-0.01	0.17	0.48	0.02
R_e	0.43	1.00	0.10	0.16	0.11	0.04
L_f	-0.01	0.10	1.00	-0.03	-0.07	-0.17
H	0.17	0.16	-0.03	1.00	0.09	-0.11
K_c	0.48	0.11	-0.07	0.09	1.00	-0.13
Z_{min}	0.02	0.04	-0.17	-0.11	-0.13	1.00
Sum	2.08	1.84	0.82	1.28	1.47	0.63
Sum total	8.12	8.12	8.12	8.12	8.12	8.12
WSA	0.26	0.23	0.10	0.16	0.18	0.08

Prioritization of watershed using PCA-WSA

The final priority ranking of the watersheds was according to the lowest CF value as shown in Table 7 and Table 8, where five priority category rankings are established. The Amazonas region with very high priority category of ≤ 3.10 and area of 11.12 %; the Lima region with high priority category of 3.11 ≤ 5.10 but with the minimum affected area of 8.37 %; the regions of La Libertad, Huancavelica, Pasco, Ancash, San Martin with medium priority category 5.11 ≤ 7.10 and area of 29.71 %; the regions of Apurimac, Junin, Ucayali, Huanuco, Loreto with very low priority category and the highest percentage of 39.35 % in the least affected.

Table 7. Final priority ranking of watersheds based on CF value in regions.

Regions	CF	Final priority ranking
Amazonas	3.07	1
Ancash	6.22	6
Apurimac	9.15	10
Cajamarca	8.62	9
Cusco	8.28	8
Huancavelica	5.58	4
Huanuco	9.83	13
Junin	9.33	11
La Libertad	5.57	3
Lima	5.07	2
Loreto	10.45	14
Pasco	5.59	5
San Martin	7.09	7
Ucayali	9.55	12

Table 8. The priority category for watersheds of the regions.

No.	Priority level	Priority category	Regions	Area (%)
1	≤3.10	Very high priority	Amazonas	11.12
2	3.11 to 5.10	High priority	Lima	8.37
3	5.11 to 7.10	Medium priority	La Libertad, Huancavelica, Pasco, Ancash, San Martin	29.71
4	7.11 to 9.10	Low priority	Cusco, Cajamarca	11.45
5	> 9.10	Very Low priority	Apurimac, Junin, Ucayali, Huanuco, Loreto	39.35

The priority ranking map for the watersheds by regions of Peru is shown in Figure 1. In the research the highest category level represents the highest degree of erosion potential and potential areas for the application of soil and water conservation, as well as a need for immediate reforestation. In this sense, the Lima and Amazonas regions are in the high and very high priority category with a totally vulnerable area of 19.49 %. According to Puno and Puno, (2019), and Setiawan and Nandini, (2021), structural soil and water conservation techniques should be applied in the most vulnerable watersheds to minimize the susceptibility of the watersheds.

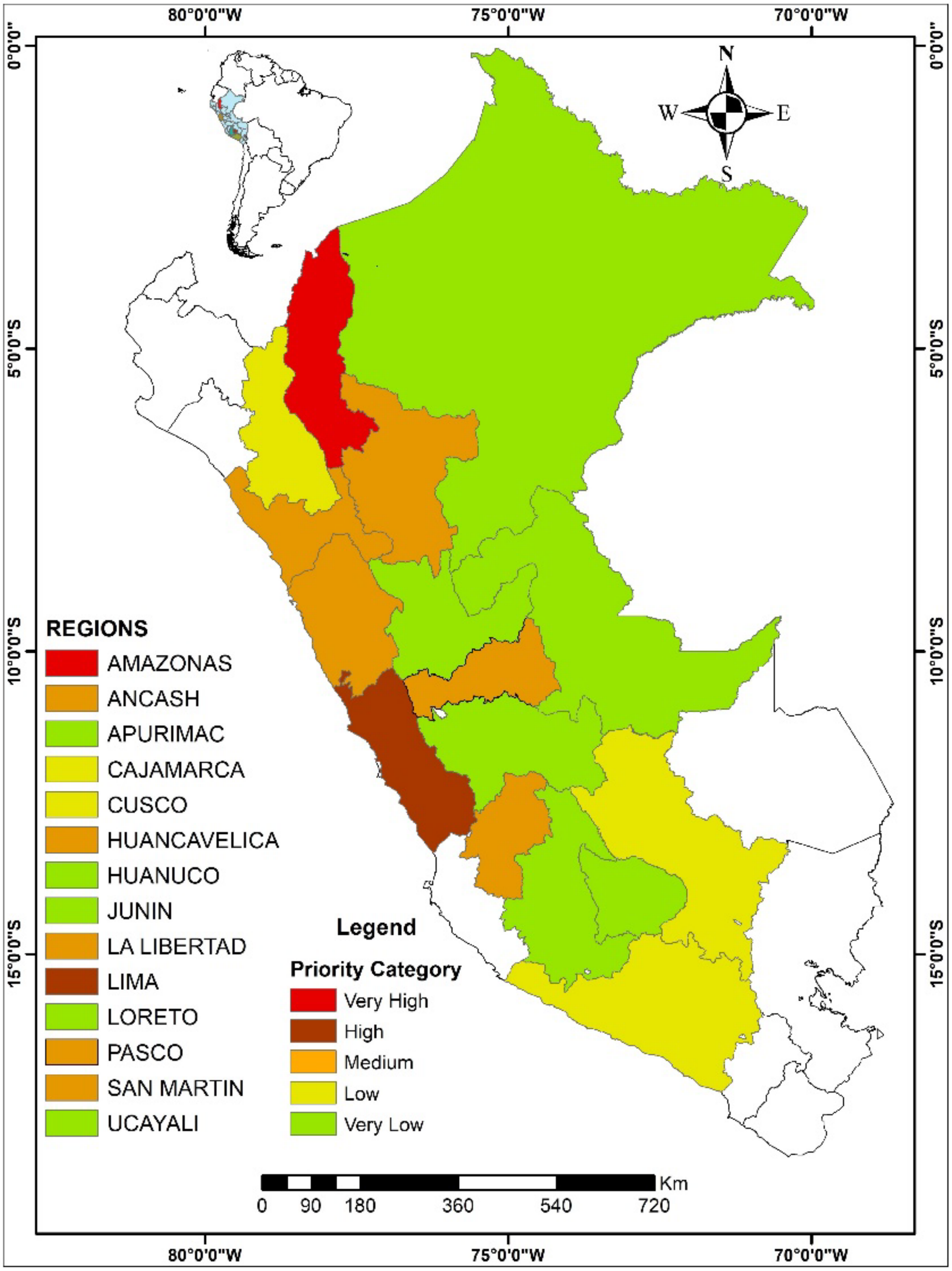

Figure 1. Priority ranking map for watersheds of regions in Peru.

Tecnología y ciencias del agua, ISSN 2007-2422, 14(6), 40-75. DOI: 10.24850/j-tyca-14-06-02

The medium and low priority category is indicative of a moderate degree of erosion and partial existence of vegetation cover, in seven regions of Peru: La Libertad, Huancavelica, Pasco, Ancash, San Martin, Cusco and Cajamarca, representing an area of 41.16 %. Reforestation programs and better sustainable use of soil and water resources should be implemented in these watersheds.

In the very low priority category, it is indicated that the watersheds have sustainable morphometric characteristics such as the regions of Apurimac, Junin, Ucayali, Huanuco and Loreto, in the latter the protection of the vegetation cover of the soil should be maintained.

Homogeneous hydrological characteristics

The Figure 2 shows the dendrogram obtained from the cluster analysis in R mode. All parameters of the watershed are presented with the exclusion of light loads (S_c and S_r). It is observed that there are two main clusters (A and B) and each one is subdivided into two main clusters. The cluster A is subdivided into cluster A(I) with two watersheds parameters (F_f and R_e), and cluster A (II) includes six parameters (A, L_r, L_c, P, W and K_c); while, cluster B is subdivided into cluster B (III) that includes two parameters (L_f and C) and cluster B(IV) that contains eight parameters ($Z_{\max}$, $Z_{\min}$, H_m, R_c, R_a, S_w, D_d and C_r), the latter, with the lowest correlation and highest distance. In this sense, cluster A presents better correlation and less distance below 0.2 with respect to cluster B.

Tecnología y ciencias del agua, ISSN 2007-2422, 14(6), 40-75. DOI: 10.24850/j-tyca-14-06-02

According to Eltahan, Elhamid and Abdelaziz (2021), although FA is not a grouping method, there are similarities between main cluster A (I and II) and CP (1 and 2), as well as B (I and II) and CP (3, 4, 5 and 6).

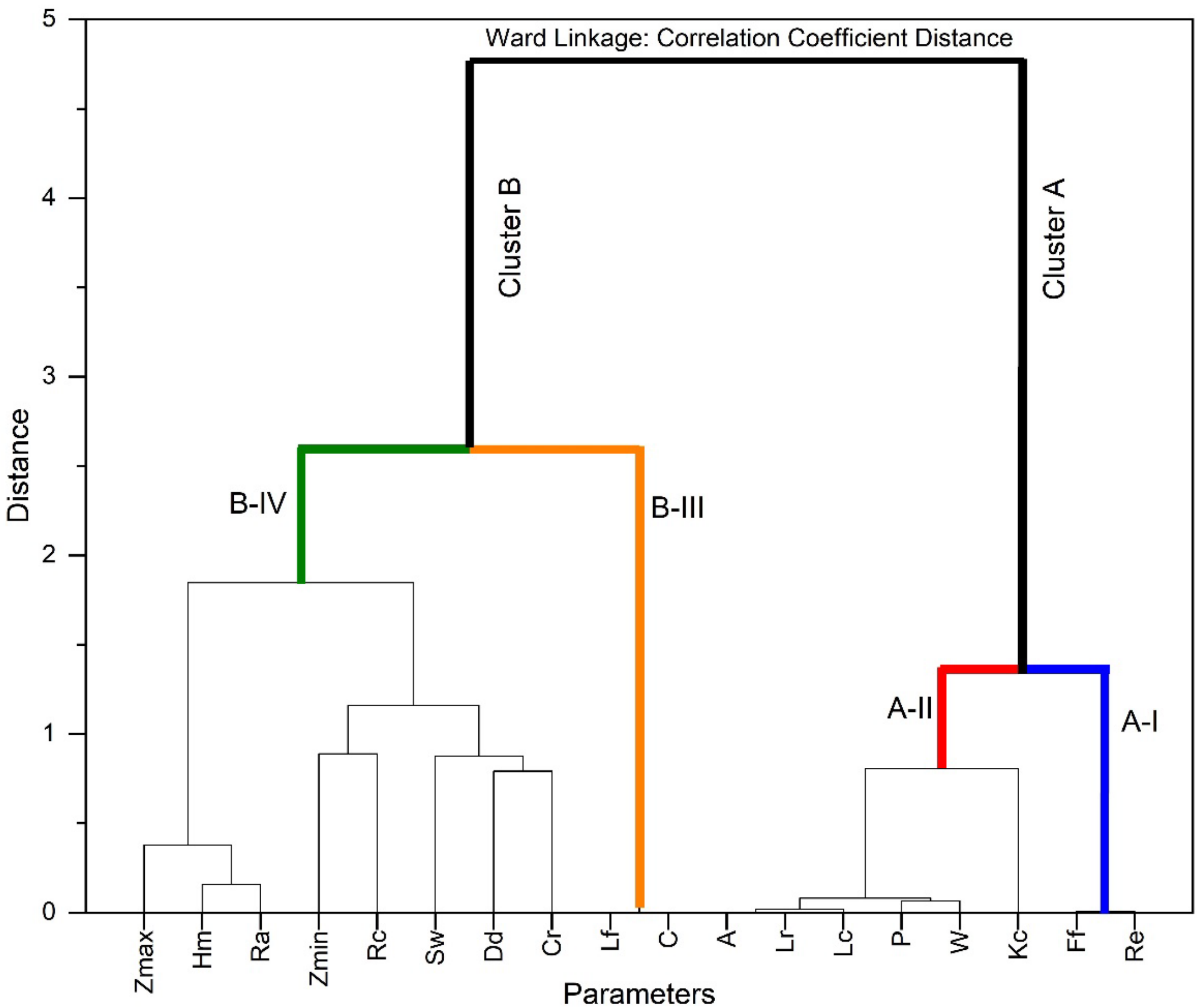

Figure 2. Dendrogram of cluster analysis results in R-mode.

Figure 3 shows the result of the dendrogram obtained by the CA for the 91 hydrographic watersheds in Peru, numbered in the abscissa axis. It is observed that there are two main clusters C and D. Cluster C is subdivided into C-I and C-II with 19 and 13 watersheds respectively. The cluster D also subdivided to cluster D-III and D-IV in watersheds 11 and 48 respectively. From the CA of Figure 2 and Figure 3, it can be seen that 35 % of the hydrographic watersheds have homogeneous areas in geometric and shape factor, of which 72 % belong to the Atlantic slope. However, 65 % are less homogeneous based on the drainage and relief factor, which implies regions with 25 % that have constant flows in the rivers as a result of the high tendency to run-off in the watersheds located in the regions of Loreto, San Martin and Cajamarca.

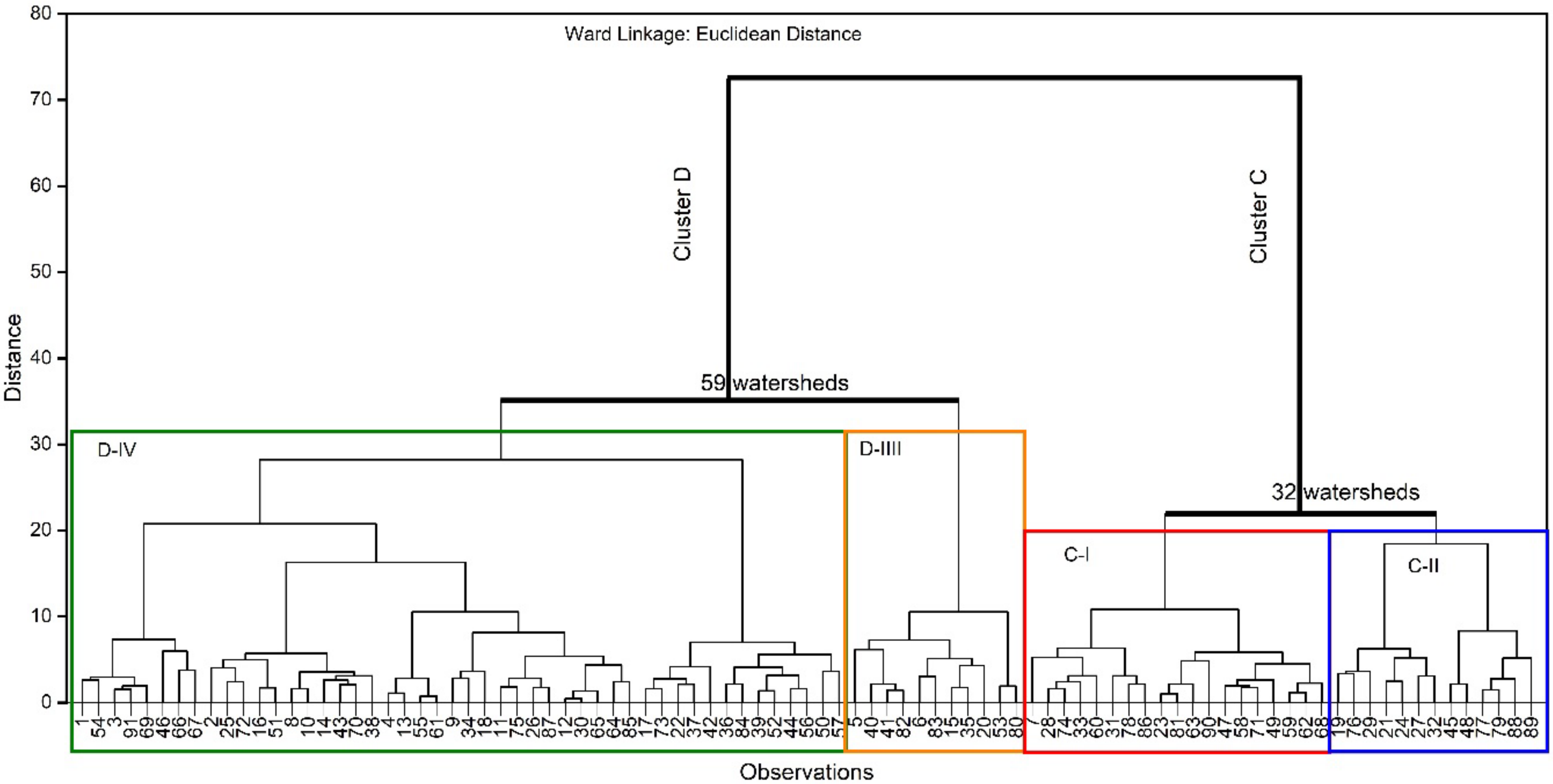

Figure 3. Dendrogram obtained by CA for the 91 drainage watersheds using Ward Linkage.

The Drainage and relief are the main cause of greater dissimilarity between the studied watersheds. This anomaly is due to the fact that the watersheds located on the Atlantic slope and on the Amazon, plains have a higher drainage density than the Pacific slope. However, the northern area of Peru due to the El Niño Southern Oscillation (ENSO) phenomenon is favoured by the intense rains, giving rise to greater floods in the rivers, in relation to the southern area where they are scarcer. In previous studies by Subyani *et al*. (2012) and Eltahan *et al*. (2021) have obtained in the CA a similarity of the hydrographic watersheds in the geometric variables.

Conclusions

The prioritization of watersheds is an elemental factor for sustainable development and adequate management of water resources. In this research, GIS, PCA and WSA techniques were used to characterize the morphometric parameters of watersheds with a focus on soil and water conservation. The study revealed that the watersheds analyzed in the Lima and Amazon regions are the most vulnerable to erosion and present potential areas for the application of soil and water conservation. These results contribute to decision-making and priority recommendations for integrated planning and management of Peru's watersheds.

The integration of the GIS, PCA and WSA methods was successful. The WSA allowed defining the weights for the most relevant parameters. Likewise, with the PCA it was possible to determine the largest number of significant parameters such as: L_c, R_e, L_f, H, K_c and $Z_{\min}$.

The homogeneous hydrological characteristics by the dendrogram by Ward's method explained that only 35 % of the hydrographic watersheds are homogeneous based on geometric parameters and in shape, with dissimilarities in sources of water resources. The study provides information and useful tools for the development of integrated river watersheds management projects at the national level. Also, it provides a series of important data for the implementation of hydraulic works for the protection and control of streams in specific regions.

Abbreviations

% = Percentage

ΔH = Mean stream height

A = Watersheds area

C = Constant channel maintenance

CA = Cluster analysis

Cov = Covariance of principal component

CF = Composite factor

Cv = Coefficient of variation

PC = Principal component

PR = Preliminary ranking

C_r = Roughness coefficient

$D^{-0,5}$ = Matriz diagonal

D_d = Drainage density

e = Equidistance between contour lines

ENSO = El Niño Southern Oscillation

FA = Factorial analysis

F_f = Form Factor

GIS = Geographic Information Systems

H = Altimetric amplitude

$\mathbf{K_c}$ = Gravelius compactness coefficient

km = Kilometer

$\mathbf{Km^2}$ = Kilometer squared

L = Watersheds length

$\mathbf{L_f}$ = Mean length of overland flow

$\mathbf{l_i}$ = Contour length

$\mathbf{L_r}$ = Stream length

Max = Maximum

Min = Minimum

P = Watersheds perimeter

PCA = Principal component analysis

r = Correlation coefficient

$\mathbf{R^2}$ = Determination coefficient

$\mathbf{R_a}$ = Relief ratio

$\mathbf{R_c}$ = Circularity ratio

$\mathbf{R_e}$ = Elongation ratio

RP = Rotación ortogonal

S = Variance

$\mathbf{S^2}$ = Covariance

$\mathbf{S_c}$ = Slope of the watersheds

$\mathbf{S_d}$ = Standard deviation

$\mathbf{S_r}$ = Mean stream slope

$\mathbf{S_w}$ = Shape index

W = Watersheds Amplitude

WSA = Weighted Sum Approach

X = Matrix

$\mathbf{X^T}$ = Transpose variables

Z = Vector

$\mathbf{Z_{max}}$ = Maximum altitude

$\mathbf{Z_{min}}$ = Minimum altitude

$\mathbf{Z_r}$ = Rotated components

References

Aher, P., Adinarayana, J., & Gorantiwar, S. (2014). Quantification of morphometric characterization and prioritization for management planning in semi-arid tropics of India. A remote sensing and GIS approach. *Journal of Hydrology*, 511, 850-860. DOI: 10.1016/j.jhydrol.2014.02.028

Amare, D. G., Kassie, A. E., & Sulla, G. G. (2020). Prioritization of watershed using morphometric analysis in Kulfo watershed in Ethiopia. *Journal of Ecology and Natural Resources*, 4, 1-7. DOI: 10.23880/jenr-16000194

Balbín, S. B., Gastmans, D., Vásquez, V. K., Vituri, S. L., Vinícius, S., & Kirchheim, R. (2020). Hydrological responses in equatorial watersheds indicated by Principal Components Analysis (PCA) – study case in Atrato River Basin (Colombia). *Brazilian Journal of Water Resources*, 25(24), 1-8. DOI: 10.1590/2318-0331.252020190165

Bharath, A., Kiran, K. K., Ramesh, M., Manjunatha, M., Ranjitha, T., & Preethi, S. (2021). Drainage morphometry based sub-watershed prioritization of Kalinadi basin using geospatial technology. *Environmental Challenges*, 5, 1-10. DOI: 10.1016/j.envc.2021.100277

Bhattacharya, R. K., Chatterjee, N. D., & Das, K. (2019). Multi-criteria-based sub-basin prioritization and its risk assessment of erosion susceptibility in Kansai-Kumari catchment area. *Applied Water Science*, 9, 76. DOI: 10.1007/s13201-019-0954-4

Cupak, A., Wałęga, A., & Bogusław, M. (2017). Cluster analysis in determination of hydrologically homogeneous regions with low flow. *Acta Scientiarum Polonorum - Formatio Circumiectus*, 16 (1), 53-63. DOI: 10.15576/ASP.FC/2017.16.1.53

Eltahan, H. A., Elhamid, A. M., & Abdelaziz, S. M. (2021). Multivariate statistical analysis of geomorphological parameters for Sinai Peninsula. *Alexandria Engineering Journal*, 60, 1389-1402. DOI: 10.1016/j.aej.2020.10.059

Finch, W. (2020). Introduction to factor analysis. In: *Exploratory factor analysis* (pp. 1-12). Thousand Oaks, USA: SAGE Publications, Inc. DOI: 10.4135/9781544339900

Gajbhiye, M. S., & Sharma, S. K. (2017). Prioritization of watershed through morphometric parameters: A PCA-based approach. *Applied Water Science*, 7, 1505–1519. DOI: 10.1007/s13201-015-0332-9

Gede, T. I., Anwar N., & Lasminto, U. (2017). Analysis of main morphometry characteristic of watershed and it's effect to the hydrograph parameters I. *IPTEK The Journal for Technology and Science*, 28, 30-36. DOI: 10.12962/j20882033.v28i1.2220

Ghosh, M., & Gope, D. (2021). Hydro-morphometric characterization and prioritization of sub-watersheds for land and water resource management using fuzzy analytical hierarchical process (FAHP): A case study of upper Rihand watershed of Chhattisgarh State. *Applied Water Science*, 11, 1-20. DOI: 10.1007/s13201-020-01340-x

Gorgoglione, A., Gioia, A., & Iacobellis, V. A. (2019). A Framework for assessing modeling performance and effects of rainfall-catchment-drainage characteristics on nutrient urban runoff in poorly gauged watersheds. *Sustainability*, 11, 1-16. DOI: 10.3390/su11184933

Helness, H., Damman, S., Sivertsen, E., & Ugarelli, R. (2019). Principal component analysis for decision support in integrated water management. *Water Supply*, 19, 2256-2262. DOI: 10.2166/ws.2019.106

Horton, R. E. (1932). Drainage basin characteristics. *Transactions, American Geophysical Union*, 13, 350-361. DOI: 10.1029/TR013i001p00350

Horton, R. E. (1945). Erosional development of streams and their drainage basins: Hydrological approach to quantitative morphology. *Bulletin of the Geological Society of America*, 56, 275-370. DOI: 10.1177/030913339501900406

Jacobs, L. K., & Brian, S. R. (2020). The next generation of climate services. *Climate Services*, 20, 1-7. DOI: 10.1016/j.cliser.2020.100199

Jolliffe, I. T. (2002). Principal component analysis and factor analysis. In: Principal component analysis (pp. 150-156). Springer Series in Statistics. New York, USA: Springer. DOI: 10.1007/0-387-22440-8_7

Kottegoda, N. T., & Rosso R. (2008). *Applied statistics for civil and environmental engineers* (2nd ed.) (pp. 326-394). Oxford, UK: Blackwell Publishing Ltd.

Kumar, B. R., Das, C. N., & Das, K. (2020). Sub-basin prioritization for assessment of soil erosion susceptibility in Kangsabati, a plateau basin: A comparison between MCDM and SWAT models. *Science of the Total Environment*, 734, 1-21. DOI: 10.1016/j.scitotenv.2020.139474

Mahala, A. (2019). The significance of morphometric analysis to understand the hydrological and morphological characteristics. *Applied Water Science*, 11, 1-16. DOI: 10.1007/s13201-019-1118-2

Malik, A., Kumar, A., Kushwaha, D., Kisi, O., Salih, S., AlAnsari, N., & Yaseen, Z. (2019). The implementation of a hybrid model for hilly sub-watershed prioritization using morphometric variables. *Water*, 11, 1-20. DOI: 10.3390/w11061138

Miller, J. R., Ritter, D. F., & Kochel, R. C. (1990). Morphometric assessment of lithologic controls on drainage basin evolution in the Crawford upland. *American Journal of Science*, 290, 569-599. DOI: 10.2475/ajs.290.5.569

Mokarram, M., & Sathyamoorthy, D. (2015). Morphometric analysis of hydrological behavior of north fars watershed. *European Journal of Geography*, 6, 88-106.

Puno, G. R, & Puno, R. C. C. (2019). Watershed conservation prioritization using geomorphometric and land use-land cover parameters. *Global Journal of Environmental Science and Management*, 5(3), 279-94. DOI: 10.22034/gjesm.2019.03.02

Schumm, S. A. (1956). Evolution of drainage systems and slopes in Badlands at Perth Amboy. New Jersey. *Bulletin of the Geological Society of America*, 67, 597-646. DOI: 10.1130/0016-7606

Setiawan, O., & Nandini, R. (2021). Sub-watershed prioritization inferred from geomorphometric and landuse/landcover datasets in Sari Watershed, Sumbawa Island, Indonesia. *IOP Conference Series: Earth and Environmental Science*, 747 (012004).

Sharma, S. K., Gajbhiye, S., & Tignath, S. (2015). Application of principal component analysis in grouping geomorphic parameters of a watershed for hydrological modeling. *Applied Water Science*, 5, 89-96. DOI: 10.1007/s13201-014-0170-1

Siddiqui, R., Said, S., & Shakeel, M. (2020). Nagmati River sub-watershed prioritization using PCA, integrated PCWS, and AHP: A case study. *Natural Resources Research*, 26, 2411-2430. DOI: 10.1007/s11053-020-09622-6

Singh, O., & Singh, J. (2018). Soil erosion susceptibility assessment of the lower Himachal Himalayan watershed. *Journal of the Geological Society*, 92, 157-165. DOI: 10.1007/s12594-018-0975-x

Subyani, A. M., Qari, M. H, & Matsah, M. I. (2012). Digital elevation model and multivariate statistical analysis of morphometric parameters of some wadis. *Arabian Journal of Geosciences,* 5, 147-157. DOI: 10.1007/s12517-010-0149-7

Vega, J. F., Lavado, C. W., & Felipe, O. O. (2018). Assessing hydrological changes in a regulated river system over the last 90 years in Rimac Basin (Peru). *Theoretical and Applied Climatology*, 132, 347-362. DOI: 101007/s00704-017-2084

Ward, J. H. (1963). Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association*, 58, 236-244. DOI: 10.2307/2282967

Wongchuig, S. C., Mello, C. R., & Chou, S. C. (2018). Projections of the impacts of climate change on the water deficit and on the precipitation erosive indexes in Mantaro River Basin. *Journal of Mountain Science*, 15, 264-279. DOI: 10.1007/s11629-017-4418-8

Yang, W., Zhao, Y., Wang, D., Wu, H., Lin, A., & He, L. (2020). Using principal components analysis and IDW interpolation to determine spatial and temporal changes of surface water quality of Xin'anjiang River in Huangshan, China. *International Journal of Environmental Research and Public Health*, 17(8), 1-14. DOI: 10.3390/ijerph17082942

Yunus, A. P., Oguchi, T., & Hayakawa, Y.S. (2014). Morphometric analysis of drainage basins in the western Arabian peninsula using multivariate statistics. *International Journal of Geosciences*, 5, 527-539. DOI: 10.4236/ijg.2014.55049

Zavoianu, I. (1985). *Morphometry of drainage bassins*. Amsterdam, The Netherlands: Elsevier Science Publishers.