Journal of Multidisciplinary Engineering Science and Technology (JMEST)
ISSN: 2458-9403
Vol. 13 Issue 6, June - 2026

Comparative Assessment Of Flood Vulnerability In The Devoll, Osum, Seman, And Vjosa River Basins Based On Topographic, Vegetative, And Climatic Indicators

Prof.Dr.Valentina Sinaj
Faculty of Economics
University of Tirana
Albania
sinajv@yahoo.com

Dr.Feruze Shakaj
Faculty of Economics University of Tirana, Tirana, Albania
Dr. Alfred Kume
University of Kent
Prof.Asoc.Dr. Elira Hoxha
Faculty of Economics University of Tirana, Tirana, Albania

***Abstract*—This study presents a comparative assessment of flood vulnerability in the Devoll, Osum, Seman, and Vjosa river basins, based on topographic, hydrological, and climatic indicators. A composite flood vulnerability index (FSI) was constructed, integrating variables such as elevation, slope, precipitation, vegetation index (NDVI), and hydrological indicators.**

The results show clear differences between the basins, with the Semani basin appearing as the most sensitive system due to its low elevation and minimal slope, while the Vjosa basin exhibits more moderate sensitivity, despite higher precipitation, due to more favorable drainage conditions. Devolli and Osumi basins present similar profiles, consistent with their hydrological connection.

Principal component analysis (PCA) and clustering identified two main dimensions of variability and confirmed the differences between river basin profiles. Historical flood evidence also supports these results, highlighting the high exposure of the Vjosa and Seman basins to flood risk.

The study shows that flood vulnerability is determined by the interaction of environmental factors, and not just by rainfall, confirming the validity of the FSI index as a tool for comparative assessment in conditions characterized by a lack of detailed data.

Keywords: Flooding, vulnerability, FSI, PCA, river basins, Albania.

I.INTRODUCTION

Floods represent one of the most widespread natural hazards with a significant impact on socio-economic and environmental systems, especially in regions with a Mediterranean climate, where rainfall variability and the intensity of extreme episodes are increasing [1] [2]. In the context of climate change and anthropogenic pressures on land use, the frequency and magnitude of flood events have increased, necessitating the development of integrated approaches to risk assessment and management [3]. Albania is characterized by a complex hydrographic network, in which river basins exhibit marked heterogeneity in topography, hydrology, and climate. The Devoll, Osum, Seman, and Vjosa basins constitute key systems in this network, particularly important for their impact on inhabited areas and agricultural activities. In particular, the Seman river system represents an interconnected hydrological unit, formed by the union of the Devoll and Osum (rivers) basins, which creates an integrated configuration of water flow and accumulation processes. The hydrological interdependence between the Devoll and Osum rivers implies that these two tributaries cannot be treated as independent systems, as they share similar climatic conditions and jointly contribute to the formation of the Seman hydrological regime. This configuration has direct implications for flood dynamics, increasing the potential for flood events in the lower areas of the basin, especially in the western lowlands of Albania (World Bank, 2015).

On the other hand, the Vjosa River represents one of the last relatively natural systems in Europe, with a hydrological regime mainly influenced by precipitation and climate variability [4]. However, intense rainfall and its uneven distribution can generate flooding events in this system as well. Meanwhile, the Seman River, due to its morphological characteristics, including low elevation and minimal slope, presents a high sensitivity to flooding even in non-extreme rainfall conditions.

Floods constitute a significant risk in Albania, affecting around 2.4% of the population each year and causing damages amounting to around 3% of GDP. According to a report by the Food and Agriculture Organization of the United Nations, floods are estimated to be the cause of the greatest economic losses at 38% during the period 1990-2014. And over the last 33 years, economic losses caused by floods are estimated at 2.3 billion USD [5]. This phenomenon is linked to intense rainfall, poor water management, and land encroachment, with the most vulnerable areas being the lowlands of rivers such as Vjosa, Semani, Devolli, and Osumi. Furthermore, climate change is expected to exacerbate this risk in the future [6]. The rivers of the Balkans, including those in Albania such as the Vjosa and the Seman system (Devoll–Osum), are among the most intact in Europe, but are being seriously threatened by infrastructure development (especially hydropower plants), which causes flow fragmentation, changes in the hydrological regime, and degradation of biodiversity [4]

In this context, flood susceptibility assessment constitutes an important methodological approach, especially in the absence of a complete and standardized inventory of historical flood events. This approach is based on the integration of physical and climatic factors that influence the generation and propagation of floods, enabling the construction of spatial risk indices [7].

This study aims to assess the vulnerability to flooding in the Devoll, Osum, Seman, and Vjosa river basins through the integration of geospatial and climatic data with different temporal resolutions. Specifically, the study combines topographic variables, land use indicators, and climatic variables, aiming to identify the most vulnerable areas and analyze the relative role of factors that influence this phenomenon.

II.LITERATURE REVIEW

Studies on Albanian river systems emphasize the key role of hydromorphological, climatic, and land-use factors in determining flood vulnerability, but there is a divergence in how these factors are weighed and interpreted. On the one hand, [8] argues that hydromorphological changes, such as bed narrowing and modifications in sediment transport, constitute critical factors that increase the risk of floods in Albanian rivers, closely linking this phenomenon to anthropogenic interventions and socio-economic changes. This approach emphasizes the importance of topographic and structural indicators in risk analyses.

Regional studies illustrate the complex factors at play. They show that coastal areas are more exposed to flood risk, particularly those located near river deltas [9]. In contrast, [10] emphasizes the role of natural river dynamics, especially in the case of the Vjosa, arguing that natural processes of flooding and sediment transport are not only a hazard, but also an essential element for the ecological functioning of the river system. This perspective suggests that flood control interventions may have negative consequences on ecological integrity, raising a conceptual tension between risk management and the conservation of natural processes.

Another critical dimension is added by [11], who argues that the intensive development of hydropower plants in the Balkan region has led to river fragmentation and significant changes in the hydrological regime. According to them, these interventions reduce the natural capacity of rivers to absorb floods, increasing the vulnerability of downstream areas. This position supports the analysis of [8] on anthropogenic impact, but contrasts with the ecological approach of [10], which emphasizes the need to preserve free-flowing rivers as reference systems.

On a broader level, the report of the [12] offers an integrative perspective, emphasizing that flood risk in Albania is the result of a complex interaction between climatic factors (intensive precipitation and climate variability), topographical characteristics, and institutional weaknesses in water management. Unlike previous studies that focus on specific processes, this approach emphasizes the institutional and managerial dimensions, extending the analysis beyond physical factors.

In this context, the literature shows a division between approaches that focus on:
(i) physical and hydromorphological factors [8] ,

(ii) ecological functions and natural dynamics of rivers [10],
(iii) anthropogenic impacts and management policies [11] [12]. However, a full integration of these dimensions in comparative flood vulnerability analyses, especially at the basin level, is lacking. This gap in the literature justifies the use of an integrated approach that combines topographic, vegetative, and climatic indicators to more comprehensively assess flood vulnerability in the Devoll, Osum, Seman, and Vjosa basins.

The existing literature on Albanian river systems has focused mainly on isolated aspects of flood risk, such as hydromorphological factors, anthropogenic impacts, and ecological dynamics of rivers. Although previous studies have highlighted the importance of topographic and climatic factors, comparative analyses at the basin level that simultaneously integrate topographic, climatic, hydrological, and vegetative indicators within a unified geospatial framework are lacking. Furthermore, the use of multidimensional statistical methods, such as Principal Component Analysis (PCA) and clustering, remains limited in studies on flood vulnerability in Albania. In this context, this study aims to address these gaps through an integrated comparative analysis of the Devoll, Osum, Seman, and Vjosa river basins, using multi-source data processed on the Google Earth Engine platform.

RESEARCH QUESTIONS AND HYPOTHESES

Based on the gaps identified in the literature and the increasing flood risk in Albanian river systems, this study aims to analyze the impact of topographical, climatic, and vegetative factors on flood vulnerability in the Devoll, Osum, Seman, and Vjosa river basins.

The study is guided by the following research questions:

RQ1: How do topographic, climatic, and vegetative factors influence flood vulnerability in selected river basins in Albania?

RQ2: Do the Devoll, Osum, Seman, and Vjosa basins present statistically different flood vulnerability profiles?

RQ3: Which set of factors – topographic or climatic – plays the dominant role in shaping flood vulnerability in the studied basins?

To answer these questions, the following hypotheses are formulated:

H1: River basins characterized by lower elevation and smaller slope present higher vulnerability to floods.

H2: Topographic factors have a greater impact on flood vulnerability than average annual precipitation.

H3: The Seman basin presents a statistically different flood vulnerability profile compared to other basins due to its morphological characteristics.

H4: The integration of topographic, climatic, and vegetation indicators within the Flood Sensitivity Index (FSI) provides an effective approach for comparative analysis of flood vulnerability in conditions of a lack of detailed data.

III.METHODOLOGY

1. Study approach

This study adopts an integrated geospatial and statistical approach to assess flood vulnerability in the Devoll, Osum, Seman, and Vjosa river basins. In the absence of a standardized inventory of flood events, the study focuses on identifying areas with high flood susceptibility based on the physical and climatic characteristics of the territory.

The methodological approach relies on combining data with different spatial and temporal resolutions, integrating static, annual, and dynamic variables into a unified analytical framework [7].

STUDY AREA

The study focuses on the Vjosa River basin and the Seman River system, which includes the main tributaries Devoll and Osumi. These basins present similar hydro-climatic and geomorphological conditions, creating a solid basis for integrated flood risk analysis. The Devoll and Osumi join to form the Seman River, creating an interconnected hydrological system, where processes are interdependent and analyzed jointly. In this context, the Vjosa River represents a relatively natural system, with limited anthropogenic interference, serving as a comparative case for assessing flood vulnerability.

DATA SOURCE

The data used in this study were obtained and processed primarily on the Google Earth Engine cloud platform, which provides access to a wide range of high-resolution geospatial data with global coverage. This platform enables advanced spatial analysis and efficient processing of large amounts of data [13]

Through Google Earth Engine, data from various satellite sources and digital terrain models have been selected and processed, including:

- elevation models (SRTM),
- hydrological data (HydroSHEDS),
- river networks (HydroRIVERS),
- vegetation index (Sentinel-2),
- and climate data on precipitation (CHIRPS).

This approach enables the integration of multi-source and multi-temporal data into a single analytical framework, increasing the reliability and reproducibility of results.

Estimation of the Flood Vulnerability Index

A composite index has been estimated to assess flood vulnerability. (Flood Susceptibility Index – FSI), which combines the main factors affecting flood risk.

The index is formulated as: $FSI = \sum_{i=1}^{n} w_i \cdot X_i$

ku: X_i are the standardized variables, and w_i are the relative weights of each factor.

The weights are determined based on the literature and the relative importance of the factors in hydrological processes. [7].

To assess the existence of long-term trends in annual precipitation, the non-parametric Mann–Kendall test was applied, which is widely used in the analysis of climatic and hydrological series to identify monotonic trends without requiring assumptions about the data distribution [14] [15]. The test was applied to each river basin, while Kendall's τ coefficient and p-value were used to assess the direction and statistical significance of the trend.

IV.DATA ANALYSIS

The data analysis focused on identifying differences between four selected river systems, namely Devoll, Osumi, Seman, and Vjosa, based on static, annual, and daily indicators extracted from Google Earth Engine. Special attention was paid to topographic and climatic variables, with the aim of identifying factors related to flood vulnerability.

DESCRIPTIVE ANALYSIS OF STATIC VARIABLES

Static variables represent the basic physical characteristics of the basins and constitute an essential component in assessing the predisposition to floods. The data analysis results that the Seman River presents significantly lower values of elevation and slope compared to other rivers. This characteristic is particularly important from a hydrological point of view, since territories with low elevation and small slope tend to favor the accumulation and retention of water on the surface.

In contrast, the Vjosa presents higher values of elevation and slope, suggesting a more dynamic morphology and a greater potential for faster water flow. However, this does not exclude the possibility of flooding, but indicates that the flood mechanisms in this basin may be different from those of the Seman.

An important element in the interpretation of the data is the fact that Devolli and Osumi present very similar values in some of the static indicators. This similarity should not be interpreted as analytical inaccuracy, but as a consequence of their hydrological connection, since these two rivers constitute the main branches that join to form the Seman. For this reason, their similarity in physical and climatic parameters is methodologically justifiable.

To assess whether there are statistically significant differences between river basins, parametric and non-parametric statistical tests were applied. Analysis of Variance (ANOVA) test was used to compare the means of the main variables between the four basins [16], while the Kruskal–Wallis test was used as a non-parametric alternative in cases where a normal distribution of data could not be assumed [17]. For pairwise comparisons between basins, the Wilcoxon Rank-Sum test was applied, which allows the identification of statistically significant differences between two independent groups [18]. The level of statistical significance was set at $\alpha = 0.05$, a threshold widely used in empirical statistical analyses [19].

Annual precipitation analysis

Table 1: Descriptive statistics of annual precipitation (2000–2023)

The river	Average (mm)	Standard Deviation	Minimum (mm)	Maximum (mm)
Vjosa	1263.54	241.02	851.29	1823.86
Devoll	1033.49	192.46	717.87	1412.31
Osum	1033.49	192.46	717.87	1412.31
Seman	950.21	186.28	699.99	1444.64

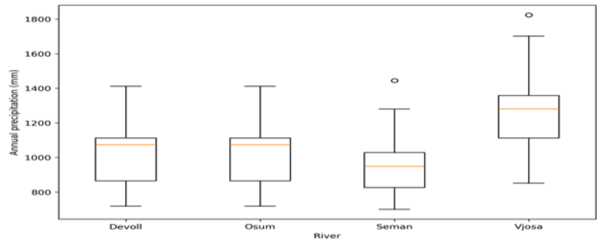


Figure 1: Annual precipitation distribution for each river basin during the period 2000–2023

As shown in Table 1, the Vjosa River is characterized by the highest level of average annual precipitation (1263.54 mm), as well as by the most pronounced variability, reflected in the highest standard deviation (241.02 mm). In contrast, the Seman River presents the lowest average precipitation value (950.21 mm), with a relatively more stable distribution.

Figure 1 illustrates the distribution of annual precipitation for each river basin during the period 2000–2023. As can be seen, the Vjosa exhibits not only the highest median precipitation but also a wider range of values, indicating greater interannual variability.

Devolli and Osumi present identical values in all statistical indicators, which reflects their hydrological connection as constituent branches of the Seman system and the similar climatic conditions that characterize them.

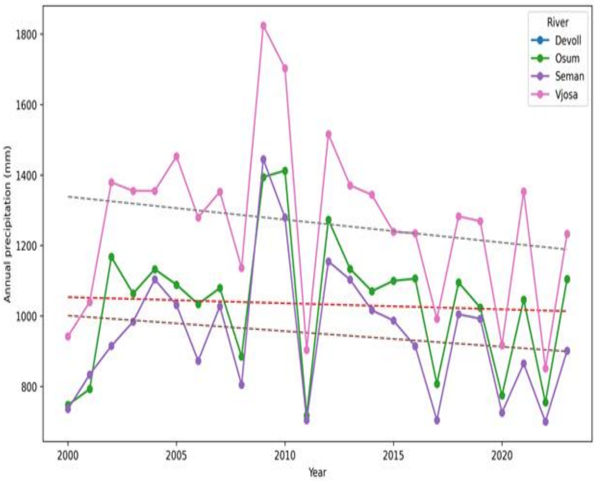


Figure 2: Total annual precipitation by river basin (2000-2023)

The analysis of annual precipitation for the period 2000–2023 reveals a high interannual variability for all basins. This variability indicates that the precipitation regime in these areas is not stable, but is characterized by significant fluctuations from one year to the next. This is an important finding, as floods are often related not only to the average amount of precipitation but also to its concentration in certain periods.

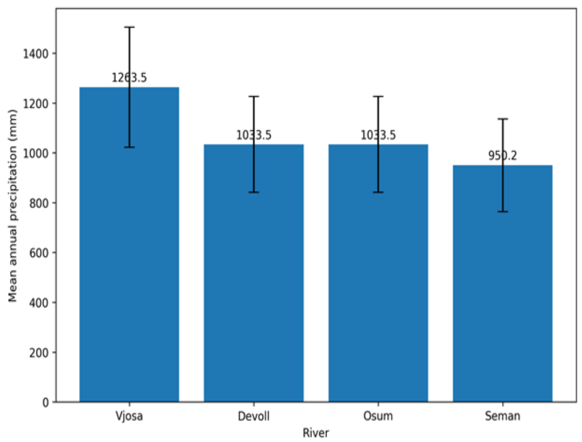


Figure 3: Annual average rainfall by river basin (2000-2023)

From the comparison of annual averages, it results that Vjosa has the highest level of precipitation, with an average of about 1263.5 mm per year. Devolli and Osumi present the same average, about 1033.5 mm, while Semani has the lowest average, about 950.2 mm per year. This result shows that Vjosa is the basin with the largest pluviometric supply, while Semani presents a more moderate situation from the point of view of precipitation.

However, the interpretation of flood risk cannot be based solely on average rainfall. In the case of Seman, although average rainfall is lower, the combination with the morphological characteristics of the basin, such as low elevation and very small slope, significantly increases the predisposition to floods. This indicates that flood vulnerability is the result of the interaction of climatic and topographic factors, and not just a single factor.

The analysis of linear trends in precipitation shows that for all four rivers, the trends are slightly negative, but very weak. For the Vjosa, the annual trend is estimated to be around -6.5 mm per year, for the Seman around -4.4 mm per year, while for the Devoll and Osumi, around -1.7 mm per year. However, these trends are not statistically strong and do not support the conclusion that there is a clear long-term trend in the analyzed dataset. This suggests that the interannual variations are dominated more by natural fluctuations than by a stable linear direction.

Precipitation seasonality analysis

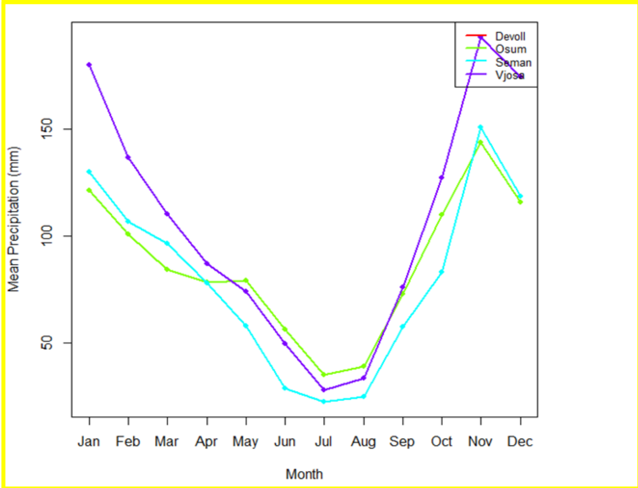


Figura 4. Seasonal distribution of average monthly precipitation by river basins

The seasonality analysis of precipitation showed a clear seasonal distribution in all river basins. The average monthly precipitation reaches the highest values during the months of November–January and the lowest during July–August. The Vjosa basin presents the highest values of precipitation in almost all months of the year, reflecting the wetter climatic character of this system. This seasonal behavior suggests that the potential risk for flooding is higher during the autumn and winter periods, when intense precipitation coincides with soil saturation and increased surface runoff.

Daily rainfall analysis

Daily data for January 2023 provide a more detailed perspective on the short-term behavior of precipitation. Analysis of these data shows that precipitation is not uniformly distributed throughout the month, but appears in the form of episodes concentrated on a few days with higher intensity. This is particularly important in flood studies, as floods are often caused by short but intense events, or by the accumulation of precipitation over a few consecutive days.

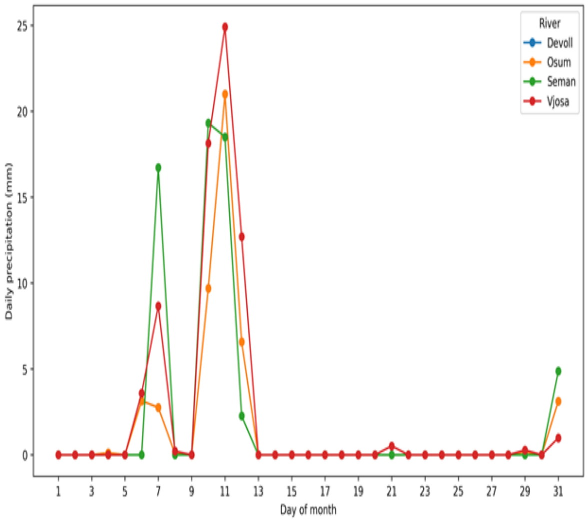


Figure 5: Daily precipitation by river basin (January 2023)

In this context, the use of variables such as 1-day, 3-day, and 7-day rainfall constitutes an appropriate approach to capture the antecedent conditions that may trigger floods. Although at this stage the study does not have a binary flood/no flood variable, the analysis of daily rainfall clearly shows that river basins are exposed to strong short-term fluctuations, which can constitute important driving factors for hydrological risk.

The Vjosa presents higher peaks even on a daily scale, reflecting its wetter regime. On the other hand, the Semani, even with more moderate daily values, remains important for flood analysis due to its topographic characteristics, which favor the spread and stability of water on the surface.

Basins comparison

The combined analysis of static and dynamic indicators shows that each river presents a unique profile of vulnerability to floods.

The Vjosa is characterized by higher precipitation and more pronounced climatic variability, suggesting a greater exposure to intense hydrometeorological episodes. However, its steeper morphology may favor faster water discharge compared to lower systems.

The Vjosa is characterized by higher precipitation and more pronounced climatic variability, suggesting a greater exposure to intense hydrometeorological episodes. However, its steeper morphology may favor faster water discharge compared to lower systems.

Devolli and Osumi exhibit similar characteristics, which is consistent with the fact that they are integral parts of the Seman system. For this reason, their interpretation should be made not only as two separate units, but also as components of a common hydrological system.

The analysis of index

To assess flood susceptibility, a composite index was constructed based on min-max standardization of physical and climatic variables. Variables that increase risk were standardized in the positive direction, while those with a protective effect were standardized in the inverse direction. Each indicator was then multiplied by a relative weight and aggregated (merged) into a final Flood Susceptibility Index. Index values were classified into five categories, from very low susceptibility to very high susceptibility.

Table 2 : Proposed weights

Variables	Impact	Weight
Elevation	the lower, the higher the risk	0.2
Slope	the lower, the higher the risk	0.2
Mean annual precipitation	the higher, the higher the risk	0.2
Water occurrence	the higher, the higher the risk	0.15
NDVI	the lower, the higher the risk	0.1
b1 / flow-related variable	the higher, the higher the risk	0.15

The weights were determined based on existing literature on flood vulnerability modeling, particularly in studies of [7], where topographic and climatic factors are identified as the main determinants of flood generation and propagation. Given that the purpose of the study is comparative assessment between basins and not the construction of a validated predictive model, a weighting scheme based on expertise and evidence from the literature was used

Table 3: Standardized indicator values and Flood Susceptibility Index (FSI) scores across the studied river basins

The river	**Elevation_n**	**Slope_n**	**Precipitation_n**	**Occurrence_n**	**NDVI_n**	**b1_n**	**FSI**
Devoll	0.046	0.297	0.266	0.072	0.565	0.927	**0.328**
Osum	0.046	0.297	0.266	0.072	0.565	0.927	**0.328**
Seman	1	1	0	1	1	0	**0.65**
Vjosa	0	0	1	0	0	1	**0.35**

Flood sensitivity index analysis

The Flood Sensitivity Index (FSI) was calculated by combining standardized topographic, hydrological, and climatic variables using the min-max method. The results showed that the Semani presented the highest index value (FSI = 0.650), classifying it as the system with the highest flood sensitivity. The Vjosa was ranked second (FSI = 0.350), while the Devolli and Osumi presented identical values (FSI = 0.328), reflecting their hydrological connection as constituent branches (components)of the Semani system.

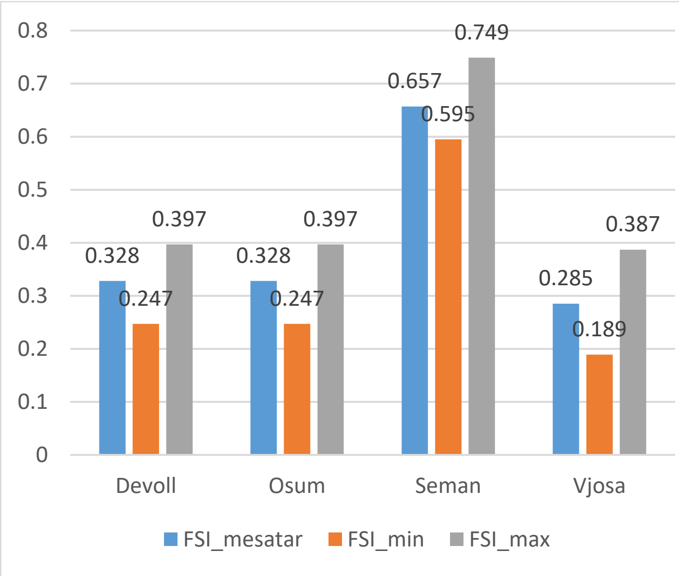


Figura 4: FSI classification

To analyze the temporal evolution of flood vulnerability, the FSI index was calculated annually for the study period, integrating static variables with annual dynamic variables. This approach allowed the construction of time series of the index for each basin and the assessment of vulnerability trends over time.

Figure 6: Annual trend of Flood Susceptibility Index (FSI, 2000-2023)

To analyze the temporal evolution of flood vulnerability, the FSI index was calculated for each year of the period 2000–2023 for each river system. The results show that the Semani maintains the highest index values consistently throughout the analyzed period, reflecting higher structural vulnerability to floods. The Devolli and Osumi present identical index series, which is consistent with their hydrological connection as constituent branches (components) of the Semani system. The Vjosa presents interannual fluctuations, but with lower average levels of FSI compared to the Semani. Although the Vjosa presents the highest precipitation, the flood vulnerability index (FSI) results are higher for the Semani, highlighting the dominant role of topographic factors in shaping the risk.

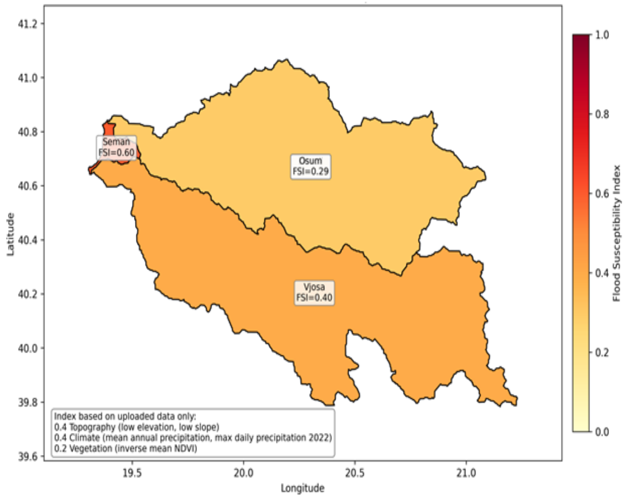


Figure 7: Spatial distribution of the flood susceptibility index (FSI) for the Devoll, Osum, Seman, and Vjosa river basins in Albania.

The figure presents the spatial distribution of the flood susceptibility index (FSI) for the studied river basins. As can be seen, the Seman River is characterized by the highest value of the index, reflecting greater vulnerability to floods, while the Devoll and Osum rivers present lower and similar values. The Vjosa River exhibits an intermediate level of vulnerability, reflecting the interaction between climatic and topographic conditions.

Sensitivity and robustness analysis of the fsi index

To assess the robustness of the FSI index to changes in the weights used, a sensitivity analysis based on Monte Carlo simulation was applied. The weights of each variable were randomly changed within the range of ±20% of their initial value and then renormalized so that their sum remained equal to 1. In total, 5000 alternative combinations of weights were generated.

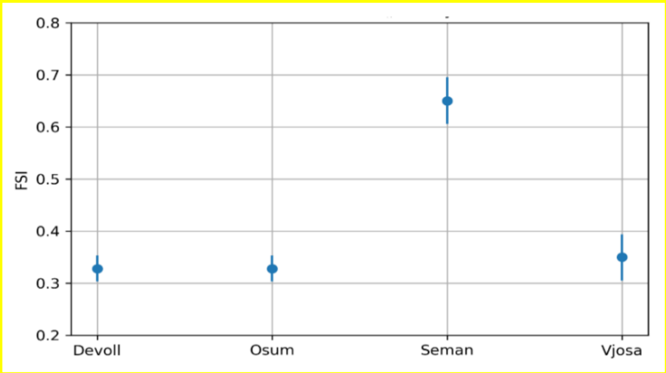


Figure 8: Monte Carlo sensitivity analysis of the Flood Susceptibility Index (FSI) across the studied river basins. Error bars represent the 95% confidence intervals obtained from 5000 Monte Carlo simulations

The results showed that the 95% intervals of the FSI were 0.303–0.353 for Devolli and Osumi, 0.606–0.695 for Seman, and 0.305–0.394 for Vjosa. In all simulations conducted, Seman maintained the highest index value, while the relative ranking of the basins remained unchanged. These results indicate that the main conclusions of the study do not critically depend on a single weighting scheme and that the FSI index presents a satisfactory level of robustness for comparative analyses of flood vulnerability.

In accordance with the recommendations of [20] Sensitivity analysis provides an important mechanism for assessing the impact of methodological uncertainty on the results of composite indices. In the case of this study, the stability of the basin rankings suggests that the results are robust to moderate changes in weights.

Comparing fsi with benchmark machine learning models

To provide an indicative comparison with benchmark models, extreme daily precipitation events were used as a proxy indicator of potentially favorable flood conditions. In the absence of a standardized flood inventory, days with precipitation above the 95th percentile were classified as extreme events, an approach often used in hydrological studies when direct flood data are lacking [21] [2]. The performance of the FSI index was compared with several machine learning models, including Logistic Regression [22]. Random Forest [23] , Support Vector Machine – SVM [24], and Gradient Boosting [25]. Performance evaluation was performed through ROC-AUC, Precision, Recall and F1-score indicators, which are widely used in machine learning and classification studies [26] [27]. The use of statistical and machine learning models in the analysis of hydrometeorological phenomena has expanded significantly in recent years. In the Albanian context, [28] compared statistical models (ARIMA, ETS) and deep learning models (LSTM, DFNN) for rainfall prediction, highlighting the importance of data-driven methods in the analysis of climate and hydrological processes. This supports the inclusion of benchmark machine learning models in the current analysis as a point of comparison to the FSI index.

Table 4 . Indicative comparison of the FSI index with benchmark machine learning models [1]

Models	ROC-AUC	Precision	Recall	F1-score
FSI	0.632	0.081	0.714	0.145
Logistic Regression	0.424	0.06	0.571	0.108
Random Forest	0.368	0	0	0
SVM	0.276	0.057	1	0.109
Gradient Boosting	0.425	0.07	0.571	0.125

The results showed that FSI achieved the highest ROC-AUC value (0.632), outperforming the benchmark models used in this indicative analysis. However, the relatively low Precision and F1-score values indicate that this comparison should not be interpreted as full predictive validation, but as an orienting test in the absence of real flood data. These results suggest that FSI offers a relatively robust and interpretable approach for cross-basin comparison under current data conditions. At the same time, the results support the need to integrate historical flood inventories and more advanced machine learning methods in future studies, in line with contemporary trends in flood vulnerability modeling [7] [29]. This recommendation is in line with recent studies conducted in Albania, which show that machine learning and deep learning models can significantly improve the analysis and prediction of climate phenomena, including precipitation and extreme hydrological events [28].

Vegetation index analysis

Figure 8 presents the evolution of the vegetation index (NDVI) and average annual rainfall in the Osum, Seman, and Vjosa river basins during the period 2000–2023.

[1] Note: The classification of extreme cases was based on daily precipitation above the 95th percentile. Because the analysis uses extreme precipitation as a proxy indicator and not a true flood inventory, the results should be interpreted as indicative benchmark comparisons and not as definitive predictive validation.

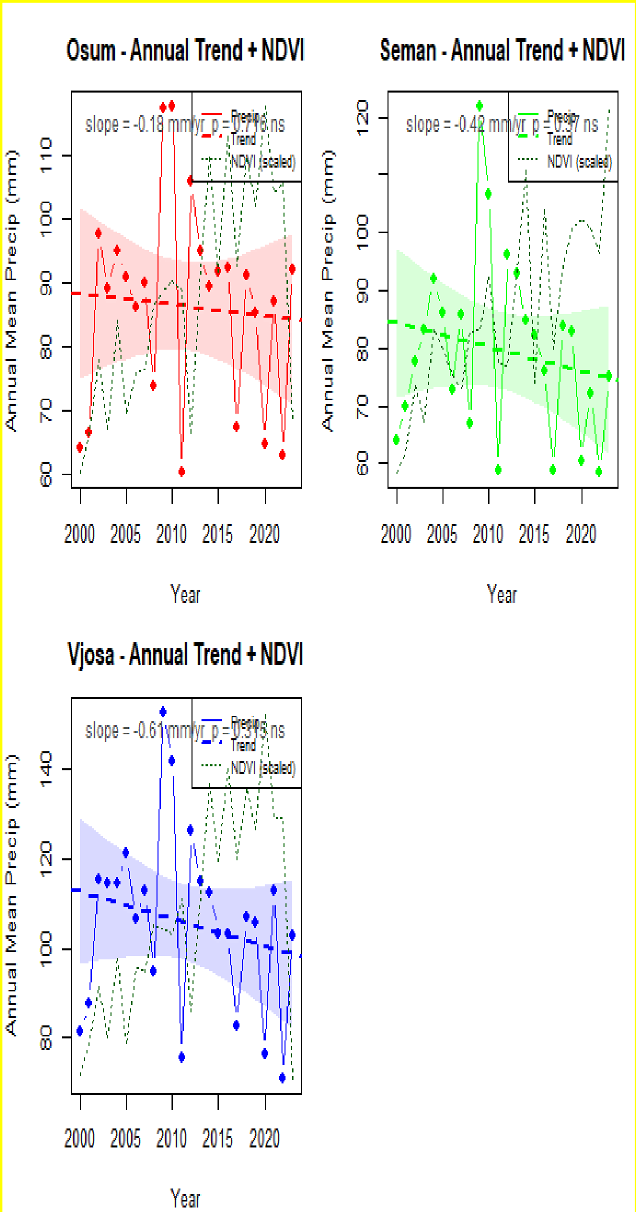


Figure 9: Annual trend and NDVI by river

In general, all three basins exhibit relatively stable NDVI values with moderate interannual fluctuations. The Vjosa basin presents the highest NDVI values for most of the analyzed period, reflecting the presence of a more developed vegetative cover and more favorable ecological conditions. On the other hand, the Seman is characterized by higher interannual variability, while the Osumi presents a more stable behavior over time.

To assess the role of vegetation in explaining precipitation changes, the relationship between NDVI and precipitation was analyzed for each river basin.

Table 5. Results of the analysis of the relationship between NDVI and precipitation by river basin

Basin	Correlation (r)	p-value	Coefficient NDVI (β)	p-value (β)	R^2 base	R^2 with NDVI
Osum	-0.079	0.715	-0.0022	0.838	0.350	0.350
Seman	0.038	0.861	0.0237	0.056	0.469	0.476
Vjosa	-0.222	0.297	-0.0092	0.545	0.481	0.482

Note: r represents the Pearson correlation coefficient between NDVI and precipitation, while β represents the coefficient of NDVI in the extended regression model. Basic R^2 refers to the model without NDVI, while R^2 with NDVI refers to the model that also includes the vegetation index.

As shown in Table 5, the relationship between NDVI and precipitation is weak and statistically insignificant in all studied basins ($p > 0.05$). Although the Seman basin shows a relatively higher NDVI impact (β = 0.0237), the model's explanatory power improves only minimally, increasing the coefficient of determination from 0.469 to 0.476. Similarly, for Osumi and Vjosa, the inclusion of NDVI does not bring a significant improvement in model performance.

These results suggest that changes in vegetation cover have a limited impact on the explanation of basin-scale precipitation variation. Consequently, topographic and climatic factors appear to play a more important role in shaping flood vulnerability than changes in vegetation index, partially supporting hypothesis H2 of this study.

Statistical tests of differences between basins

To assess whether river basins present statistically significant differences in the main variables, ANOVA and Kruskal–Wallis tests were applied for comparisons between the four basins. Also, for pairwise comparisons between basins, the Wilcoxon test was used for annual precipitation and NDVI series.

Table 6. Statistical tests for differences between basins

Variable	Test	Statistics	p-value	Interpretation
Annual rainfall	ANOVA	F = 10.446	< 0.001	Significant differences
Annual rainfall	Kruskal–Wallis	H = 21.947	< 0.001	Significant differences
Annual NDVI	ANOVA	F = 7.611	< 0.001	Significant differences
Annual NDVI	Kruskal–Wallis	H = 17.750	< 0.001	Significant differences
Daily rainfall	ANOVA	F = 0.172	0.915	Not significant
Daily rainfall	Kruskal–Wallis	H = 1.386	0.709	Not significant

The results showed statistically significant differences between basins for annual precipitation, both according to ANOVA ($F = 10.446$, $p < 0.001$) and according to the Kruskal–Wallis test ($H = 21.947$, p <

0.001). Statistically significant differences were also evidenced for annual NDVI, where ANOVA resulted in F = 7.611, p < 0.001, while Kruskal–Wallis resulted in H = 17.750, p < 0.001. These results confirm that the basins do not have the same climatic and vegetation profiles.

Wilcoxon tests showed significant differences between most pairs of basins. Devolli and Osumi did not show differences, as their values are identical in the dataset, consistent with their hydrological connection. In contrast, Semani and Vjosa were significantly differentiated from the other basins for annual precipitation and NDVI. For daily precipitation, the tests did not show statistically significant differences between basins (ANOVA: p = 0.915; Kruskal–Wallis: p = 0.709), suggesting that daily variation is more unstable and less differentiated at the basin level.

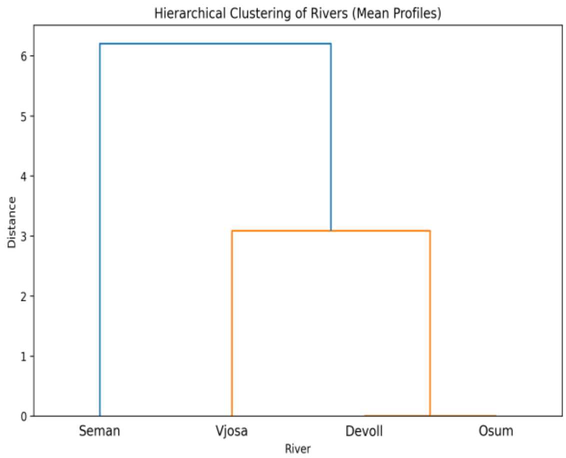
Multivariate analysis (pca)

To analyze the multidimensional structure of the data, Principal Component Analysis (PCA) was applied, combined with clustering methods. PCA is a widely used statistical technique for reducing the dimensionality of data and identifying dominant/main/major factors that explain the variability between observations [30] [31]. Similarly, clustering methods are used to group observations based on their similarity, revealing hidden structures in the data [32].

This combined approach enables the reduction of data complexity and, at the same time, the identification of relationships between topographic, hydrological, and climatic variables, which are essential for flood vulnerability analysis.

To avoid visual overload and maintain clarity of presentation, the results are presented through two main figures: a PCA plot and a dendrogram. These graphical representations synthesize the structure of the data and highlight relationships between river basins, providing an integrated interpretation of the factors influencing flood vulnerability.

Figure 10: Hierarchical Cluster of Rivers



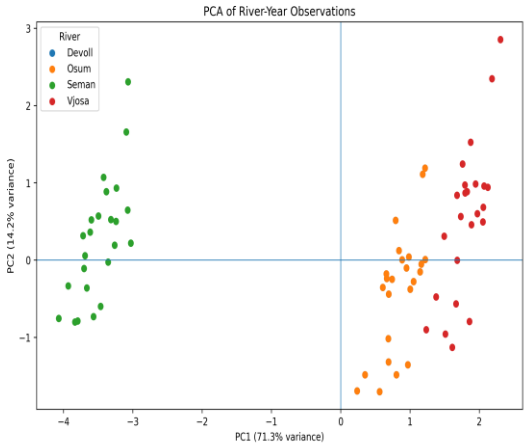


Figure 11: PCA of River-Year observations

The results showed that the first two components of the PCA explain 85.5% of the total variance, where PC1 is mainly related to topographic and hydrological indicators, while PC2 is dominated by precipitation. The clustering analysis revealed a distinct profile for the Seman River, while the Devolli and Osumi presented high similarity, in accordance with their hydrological connection. The Vjosa resulted in a distinct profile, but closer to the Devolli and Osumi than to the Seman River.

DISCUSSION

The results show that flood vulnerability in Albanian river basins is determined by the interaction of topographic, hydrological, and climatic factors. The FSI index identified the Seman basin as the most flood-sensitive basin, while the Devolli and Osumi basins presented similar profiles, consistent with their hydrological connection. This result supports hypothesis H1, according to which basins characterized by lower elevation and smaller slope present higher flood vulnerability. Although Vjosa recorded the highest level of average annual precipitation, its morphological characteristics, especially the higher elevation and slope, favor more efficient drainage. In contrast, Seman is characterized by very low elevation and slope, which favors water accumulation and increases the potential for flooding. These results support the findings of [7] and [8] according to which topographic factors play a crucial role in shaping flood vulnerability.

The seasonality analysis revealed a clear Mediterranean rainfall regime, with maxima during the period November–January and minima during July–August. This pattern suggests that the potential flood risk is highest during autumn and winter, when intense rainfall is combined with soil saturation and increased surface runoff. On the other hand, the rainfall trend analysis and the Mann–Kendall test did not identify statistically significant trends ($p > 0.05$), indicating that the interannual variability is dominated by natural climatic fluctuations and not by a long-term stable change. This result is consistent with the assessments of (IPCC, 2021), according to which the Mediterranean region is characterized by increased climate variability and extreme events rather than by clear linear rainfall trends.

The results of the NDVI analysis showed that the relationship between vegetation cover and precipitation is relatively weak and statistically insignificant. The inclusion of NDVI in the analysis did not significantly change the relative ranking of basins according to flood vulnerability, suggesting that topographic and hydrological factors exert a greater influence than vegetative factors at the scale of analysis used in this study. At the same time, the fact that Vjosa presents the highest precipitation, but not the highest level of flood vulnerability, partially supports hypothesis H2, according to which topographic factors have a greater impact on flood vulnerability than average annual precipitation. This result is in line with studies [7], which emphasize the importance of the physical characteristics of basins in the generation and distribution of floods.

PCA and clustering analysis confirmed the existence of distinct vulnerability profiles between basins, highlighting in particular the position of Seman as the most exposed system to risk. Devolli and Osumi were grouped together, reflecting their similar physical and climatic characteristics, while Vjosa presented a distinct profile associated with higher rainfall and more favorable drainage conditions. These results support hypothesis H3, according to which Seman presents a statistically different flood vulnerability profile compared to other basins due to its morphological characteristics.

The consistency of the results with historical flood evidence in the Seman and Vjosa rivers increases the reliability of the approach used. Furthermore, the ability of the FSI index to identify clear differences between basins supports hypothesis H4, according to which the integration of topographic, climatic, and vegetation indicators provides an effective approach for the comparative analysis of flood vulnerability. However, the study is based on aggregated data at the basin level and does not include a standardized inventory of flood events, which limits the possibility of direct validation of the index. For this reason, the results should be interpreted as a relative assessment of vulnerability and not as a predictive model of flood risk. Despite this limitation, the FSI proves to be a useful tool for identifying relative differences between basins and for supporting decision-making in conditions of a lack of detailed hydrological and historical data.

LIMITATIONS

This study presents several limitations that should be considered when interpreting the results. First, the analysis was conducted at the basin level and is based on aggregated data, which does not allow the identification of spatial variability within each basin. Second, the lack of a standardized inventory of historical flood events limits the possibility of directly validating the FSI index and using more advanced performance assessment methods. Also, the study focuses mainly on physical and climatic factors and does not include socio-economic variables, which may affect exposure and vulnerability to floods. Despite these limitations, the approach used provides a useful basis for the comparative assessment of flood vulnerability in conditions of a lack of detailed data.

CONCLUSIONS

This study shows that the vulnerability to floods in the Albanian river basins is determined by the interaction of topographic, hydrological, and climatic factors. The results identify the Seman basin as the most vulnerable system, due to the combination of low elevation and minimal slope, which favors water accumulation.

In contrast, the Vjosa basin, although characterized by higher precipitation, presents a more moderate sensitivity due to more favorable drainage conditions, while the Devolli and Osumi basins exhibit similar profiles, consistent with their hydrological connection as part of the same river system.

The results of PCA and clustering analysis reinforce these findings, highlighting clear divisions between river profiles and supporting the validity of the FSI index as a suitable tool for the comparative assessment of flood vulnerability.

References

1. European Environment Agency (EEA). (2020). *Flood risks and environmental vulnerability in Europe*. Copenhagen: EEA.
2. Intergovernmental Panel on Climate Change (IPCC). (2021). *Climate Change 2021: The Physical Science Basis*. Cambridge University Press.
3. United Nations Office for Disaster Risk Reduction (UNDRR). (2021). *Global Assessment Report on Disaster Risk Reduction 2021*. Geneva: UNDRR.
4. Schwarz, U., & Hudek, H. (2016). *The Balkan Rivers: Save the Blue Heart of Europe*. Riverwatch. Available at: https://www.balkanrivers.net
5. United Nations (UN). (2018). *Comprehensive Analysis of Disaster Risk Reduction and Management Systems for Agriculture in Albania*
6. World Bank. (2015). *Albania: Flood Risk Management and Adaptation Strategy*. Washington, DC: World Bank.
7. Tehrany, M. S., Pradhan, B., Jebur, M. N., & Shafri, H. Z. M. (2015). Flood susceptibility mapping using a novel ensemble weights-of-evidence and support vector machine models in GIS. *Journal of Hydrology, 512*, 332–343. https://doi.org/10.1016/j.jhydrol.2014.03.008
8. Cekrezi, B. (2024). *Hydro-morphology, channel change and sediment transport dynamics of major Albanian rivers* (Doctoral dissertation). University of Trento, Italy.
9. Rentschler, J., & Salhab, M. (2020). *People in Harm's Way: Flood Exposure and Poverty in 189 Countries*. World Bank Policy Research Working Paper No. 9447. https://doi.org/10.1596/1813-9450-9447
10. Schiemer, F., Drescher, A., Hauer, C., & Schwarz, U. (2018). *The Vjosa River corridor: A riverine ecosystem of European significance*. *Acta ZooBot Austria, 155*, 1–40.
11. Hudek, H., & Schwarz, U. (2021). *Balkan rivers are endangered by construction of new hydropower plants*. *Danube News, 43*, 4–7. International Association for Danube Research (IAD).
12. World Bank. (2023). *Albania: Water Security Assessment and Action Planning*. Washington, DC: World Bank.
13. Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment, 202*, 18–27. https://doi.org/10.1016/j.rse.2017.06.031
14. Mann, H. B. (1945). Nonparametric tests against trend. *Econometrica, 13*(3), 245–259. https://doi.org/10.2307/1907187
15. Kendall, M. G. (1975). *Rank Correlation Methods* (4th ed.). London: Charles Griffin.
16. Fisher, R. A. (1925). *Statistical Methods for Research Workers*. Edinburgh: Oliver and Boyd.
17. Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association, 47*(260), 583–621. https://doi.org/10.2307/2280779
18. Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin, 1*(6), 80–83. https://doi.org/10.2307/3001968
19. Field, A. (2018). *Discovering Statistics Using IBM SPSS Statistics* (5th ed.). Sage Publications.
20. Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., & Tarantola, S. (2008). *Global Sensitivity Analysis: The Primer*. John Wiley & Sons. https://doi.org/10.1002/9780470725184
21. Kundzewicz, Z. W., Kanae, S., Seneviratne, S. I., Handmer, J., Nicholls, N., Peduzzi, P., et al. (2014). Flood risk and climate change: Global and regional perspectives. *Hydrological Sciences Journal, 59*(1), 1–28. https://doi.org/10.1080/02626667.2013.857411
22. Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied Logistic Regression* (3rd ed.). Wiley. https://doi.org/10.1002/9781118548387
23. Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32. https://doi.org/10.1023/A:1010933404324
24. Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning, 20*(3), 273–297. https://doi.org/10.1007/BF00994018
25. Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics, 29*(5), 1189–1232. https://doi.org/10.1214/aos/1013203451
26. Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters, 27*(8), 861–874. https://doi.org/10.1016/j.patrec.2005.10.010
27. Curma, M., & Sinaj, V. (2026). Automating the classification of economic activities in official statistics: A comparative study of neural networks and transformers. *WSEAS Transactions on Computer Research, 14*, 45–55. https://doi.org/10.37394/232018.2026.14.4
28. Xhabafti, M., Vika, B., & Sinaj, V. (2024). *A comparative study of statistical and deep learning model-based weather prediction in Albania*. *WSEAS Transactions on Computer Research, 12*, 151–160. https://doi.org/10.37394/232018.2024.12.15
29. Mosavi, A., Ozturk, P., & Chau, K. W. (2018). Flood prediction using machine learning models: Literature review. *Water, 10*(11),

1536. https://doi.org/10.3390/w10111536

30. Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. *Philosophical Magazine, 2*(11), 559–572. https://doi.org/10.1080/14786440109462720
31. Jolliffe, I. T. (2002). *Principal Component Analysis* (2nd ed.). Springer. https://doi.org/10.1007/b98835
32. Everitt, B. S., Landau, S., Leese, M., & Stahl, D. (2011). *Cluster Analysis* (5th ed.). Wiley.
33. Zaimi, K., & Oruçi, J. (2020). Flood forecasting in the western lowland of Albania with application of hydrological modeling. *Journal of International Environmental Application & Science, 15*(4), 216–223.