Artificial Intelligence and GIS-Based Identification of Environmental Degradation Zones in an Andean Urban River Corridor: The Case of the Chili River

Valkiria Ibarcena

Universidad Politécnica de Valencia, Valencia, Spain.

Abstract

Urban growth in recent years has transformed the ecological and functional conditions of river landscapes. In response to these changes, this study examines the spatial distribution of environmental degradation in the Chili River landscape (Arequipa, Peru) using GIS-based multicriteria analysis and unsupervised machine learning. Ecological condition, environmental pressure, and territorial functionality were synthesized into three normalized indices, which were used as input data for K-means/Iso Cluster classification. Solutions ranging from two to six clusters were evaluated using WCSS, the Silhouette Coefficient, and the Davies–Bouldin Index. Four clusters were selected because they preserved the main environmental gradients while providing a simpler territorial interpretation. High degradation covered 46.58% of the classified area and very high degradation 21.89%, together accounting for 68.47% of the analyzed territory; low degradation represented 12.72%. Cluster stability was also assessed by independently varying the normalized indices by $\pm 10\%$, resulting in a mean Adjusted Rand Index of 0.904. The classification reveals a marked contrast between highly degraded areas and sectors that retain greater ecological and functional capacity. These spatial differences provide a basis for locating critical areas and prioritizing conservation and restoration measures across the Chili River landscape.

Keywords

Artificial intelligence, Environmental degradation, K-means, GIS, Urban fluvial landscape, Spatial clustering

1. Introduction

Urban pressure on river landscapes leads to the fragmentation of ecological networks, thereby weakening ecosystem functions. These changes are particularly relevant in Arequipa, where the Chili River crosses a landscape in which urban development, agricultural land, and ecologically important areas coexist. Previous research has documented a reduction in riparian connectivity associated with the conversion of agricultural land and grasslands into built-up surfaces [1]. Changes in land use and land cover have also affected annual water yield, surface runoff, and urban drainage along the river [2]. Environmental degradation in this setting therefore involves several interacting processes that cannot be adequately represented through a single territorial variable.

Urban rivers function as socio-ecological systems in which hydrological, ecological, and human processes operate across different spatial scales [3]. Their assessment consequently requires information on both environmental conditions and the pressures produced by territorial transformation. GIS and remote sensing provide the spatial basis for examining these relationships, particularly for mapping land-use change and identifying areas affected by environmental transformation [4].

Machine-learning techniques add another analytical possibility because they can identify patterns within multivariable spatial datasets that may not be evident from individual thematic layers [5, 6, 7, 8]. Thus, unsupervised clustering is particularly relevant when predefined degradation classes are not available [9]. Despite its growing application in spatial and environmental analysis, its use in Andean river landscapes remains limited, particularly when ecological condition, anthropogenic pressure, and ecosystem functionality are considered simultaneously.

IWSAI 2026: 1st International Workshop on Systems, Applications, Artificial Intelligence and Innovation, September 16–18, 2026, Piura, Peru
valkiria.ibarcena@gmail.com (V. Ibarcena)
0000-0003-4985-0228 (V. Ibarcena)

Therefore, this study addresses the following research question: To what extent can ecological condition, environmental pressure, and territorial functionality, integrated through GIS-based multicriteria spatial analysis and unsupervised machine learning, identify spatially differentiated environmental degradation patterns in the Chili River landscape? To address this question, three indices were developed: the Ecological Index (IE), Environmental Pressure Index (IPA), and Territorial Functionality Index (IFT). After normalization, these indices were used as input data for K-means/Iso Cluster classification. Alternative clustering solutions were quantitatively compared, and sensitivity analysis was conducted to assess the stability of the resulting spatial partition.

This study makes both a methodological and territorial contribution. Rather than deriving degradation zones from cartographic overlay alone, it uses unsupervised clustering to identify recurring combinations of IE, IPA, and IFT. Applied to the Chili River landscape, this procedure provides a spatial basis for distinguishing critical areas from sectors that retain greater ecological and functional capacity, with potential applications in conservation, restoration, and territorial management.

2. Study Area

The case study was conducted in Arequipa, Peru, covering an area of 318,475 ha. The analyzed territory includes urban and peri-urban sectors, agricultural land, and areas of ecological importance, resulting in marked differences in land use and environmental conditions. This delimitation corresponds to the spatial extent used to calculate the IE, IPA, and IFT. It is therefore broader than the morphological urban corridor of the Chili River and should not be interpreted as its physical extent.

This broader delimitation was necessary because several variables included in the analysis extend beyond the immediate riverbanks. The Chili River remains the main fluvial focus due to its hydrological, ecological, and landscape role within Arequipa. Consequently, the environmental model was calculated across the entire analytical domain, while the maps use enlarged views of the river corridor to show

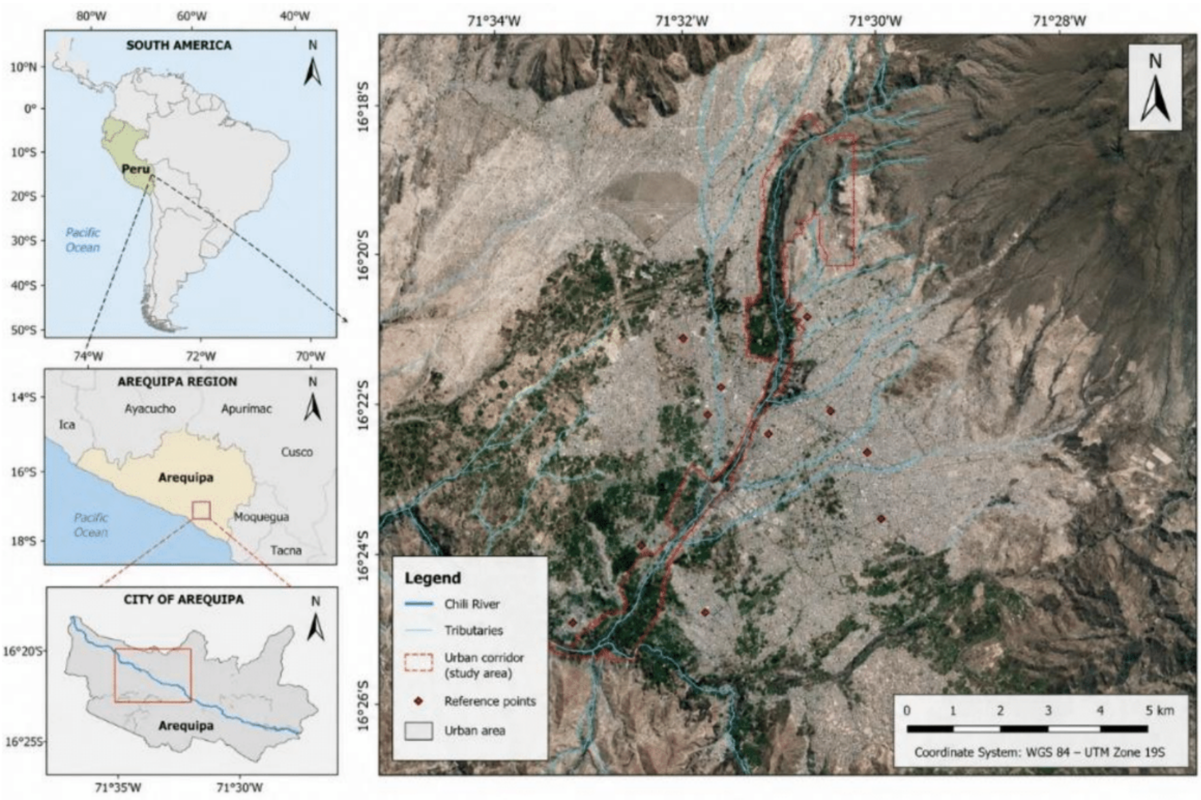

Figure 1: Location of the Chili River urban corridor in Arequipa, Peru

degradation patterns in greater detail within the immediate fluvial landscape.

3. Methodology

A methodology combining GIS-based multicriteria analysis and unsupervised clustering was applied to map environmental degradation across the entire study area. Figure 2 summarizes the workflow from the original thematic layers to the final degradation map.

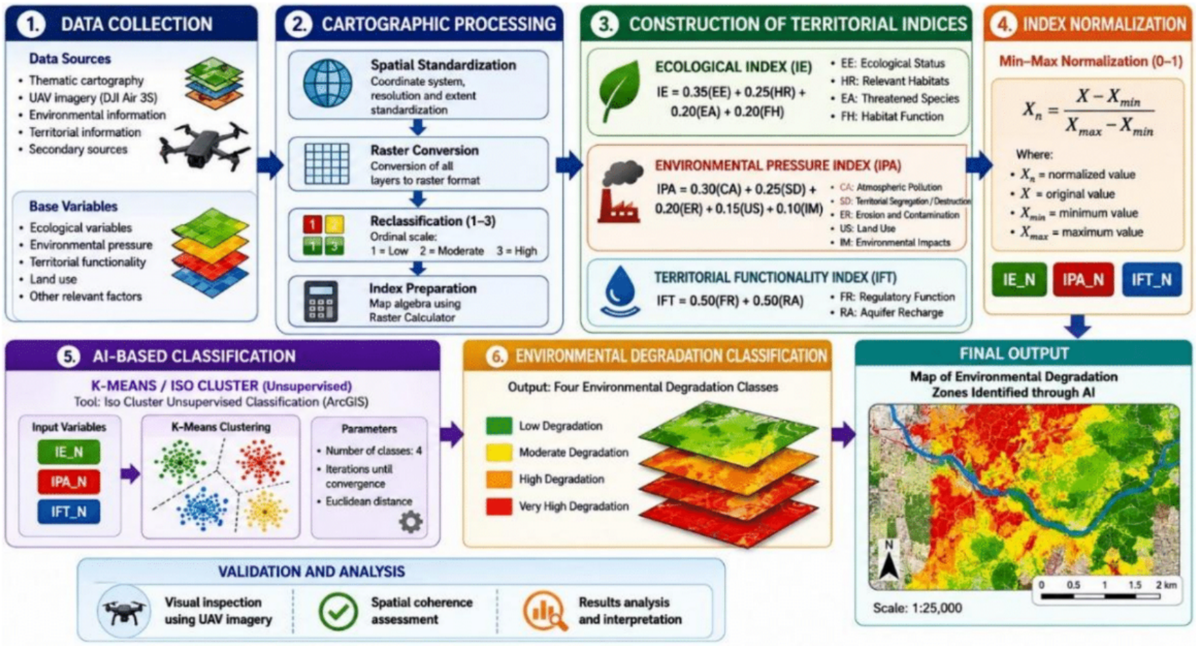

Figure 2: Methodological framework for environmental degradation zoning in the Chili River urban fluvial landscape

3.1. Cartographic Processing and Data Preparation

Before index calculation, all layers were projected to WGS 84 / UTM Zone 19S (EPSG:32719) and converted to a common raster structure with a 30 × 30 m cell size. A shared spatial extent and raster alignment were used to ensure cell-to-cell correspondence among datasets.

The variables were then reclassified into comparable ordinal categories according to their environmental meaning and grouped into three analytical dimensions: the Ecological Index (IE), representing ecological condition and conservation relevance; the Environmental Pressure Index (IPA), representing anthropogenic and environmental pressures; and the Territorial Functionality Index (IFT), representing regulatory and hydrological functions. Variable weights were assigned within each index and sum to 1.00 independently, as detailed in Table 1. No additional weighting was applied among IE, IPA, and IFT before the clustering stage.

Variable weights were defined through expert judgment within a GIS-based Multi-Criteria Decision Analysis (GIS-MCDA) framework. Following established spatial decision-making approaches [10, 11] and applications to ecological assessment [12], higher weights were assigned to variables representing primary ecological or environmental processes, whereas complementary variables received lower weights according to their relative contribution to each index.

3.2. Construction of Territorial Indices

The reclassified variables were combined through weighted linear aggregation to obtain three territorial indices. Each index was calculated independently using the weights reported in Table 1.

Table 1
Environmental variables, analytical dimensions, and weights used to construct territorial indices

Index	Variable	Weight
Ecological Index (IE)	Ecological status	0.35
	Relevant habitats	0.25
	Threatened species	0.20
	Habitat function	0.20
	Total	1.00
Environmental Pressure Index (IPA)	Atmospheric pollution	0.30
	Segregation and destruction	0.25
	Erosion	0.20
	Land use	0.15
	Environmental impacts and risks	0.10
	Total	1.00
Territorial Functionality Index (IFT)	Regulatory function	0.50
	Aquifer recharge	0.50
	Total	1.00

1. Ecological Index (IE). The IE combines ecological status, relevant habitats, threatened species, and habitat function:

$$IE = 0.35(EE) + 0.25(HR) + 0.20(EA) + 0.20(FH) \tag{1}$$

   where ES represents ecological status, RH relevant habitats, TS threatened species, and HF habitat function.
2. Environmental Pressure Index (IPA). The IPA integrates the main pressures associated with atmospheric pollution, territorial fragmentation, erosion, land use, and environmental impacts:

$$IPA = 0.30(CA) + 0.25(SD) + 0.20(ER) + 0.15(US) + 0.10(IM) \tag{2}$$

   where AP represents atmospheric pollution, SD segregation and destruction, ER erosion, LU land use, and IR environmental impacts and risks.
3. Territorial Functionality Index (IFT). The IFT combines regulatory function and aquifer recharge:

$$IFT = 0.50(FR) + 0.50(RA) \tag{3}$$

   where RF represents regulatory function and AR aquifer recharge.

3.3. Index Normalization

To ensure comparability among the three territorial indices, IE, IPA, and IFT were normalized to a common $0 - 1$ scale using min–max normalization:

$$X_n = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{4}$$

where is the normalized value, is the original index value, and and are the minimum and maximum values of the corresponding index.

This procedure generated IE_N, IPA_N, and IFT_N, which were subsequently used as three equally represented input dimensions in the K-means/Iso Cluster analysis. No additional weighting was applied among the normalized indices.

3.4. Classification through Artificial Intelligence

Environmental degradation patterns were identified through unsupervised classification using the K-means/Iso Cluster algorithm implemented in ArcGIS. The normalized indices IE_N, IPA_N, and IFT_N were then used as input data, allowing raster cells with similar ecological, environmental pressure, and territorial functionality profiles to be grouped without predefined class labels.

The resulting clusters were interpreted according to their centroid profiles and spatial characteristics and were subsequently organized into four operational degradation levels: low, moderate, high, and very high. Information obtained using the DJI Air 3S UAV was used only as complementary evidence for territorial interpretation and was not employed to calculate a confusion matrix or an external accuracy coefficient.

3.5. Cluster-number selection and validation

Candidate solutions from $k = 2$ to $k = 6$ were compared using three internal criteria: the within-cluster sum of squares (WCSS), the mean Silhouette Coefficient, and the Davies–Bouldin Index (DBI). The WCSS curve was examined for diminishing reductions as “k” increased, while higher Silhouette values and lower Davies–Bouldin values were interpreted as indicating better cluster separation and compactness.

As the $k = 4$ and $k = 6$ partitions showed better performance, they were further compared using cluster-membership correspondence and centroid profiles based on IE_N, IPA_N, and IFT_N. The retained solution was selected by considering internal clustering performance together with parsimony and territorial interpretability.

3.6. Sensitivity Analysis

Classification robustness was examined by independently varying each normalized index by $\pm 10\%$, while keeping the other two indices unchanged. This generated six alternative scenarios, each of which was reclassified using the same $k = 4$ configuration as the baseline model. Agreement between each classification and the baseline solution was quantified using the Adjusted Rand Index (ARI), where values closer to 1 indicate greater similarity between cluster partitions.

This analysis was used to determine whether moderate changes in the contribution of each environmental dimension substantially altered the spatial clustering pattern. Therefore, it assesses the stability of the classification rather than the optimization of the original variable weights.

4. Results

4.1. Ecological Index (IE)

The IE identified areas with comparatively better ecological conditions across the analyzed Chili River landscape. The results showed that areas with better ecological conditions were located mainly in peripheral sectors and areas associated with natural corridors, characterized by the presence of relevant habitats and significant ecological functions. In contrast, urbanized areas and sectors under anthropogenic pressure exhibited lower index values, associated with ecological fragmentation and territorial transformation processes (Figure 3).

4.2. Environmental Pressure Index (IPA)

The Environmental Pressure Index (IPA) showed a clear spatial contrast across the study area (Figure 4). Higher IPA values were concentrated mainly in consolidated urban sectors and areas undergoing greater territorial transformation, where atmospheric pollution, land-use pressure, fragmentation, erosion, and other environmental impacts overlapped. Lower values occurred in sectors with less intensive urban occupation and lower combined anthropogenic pressures.

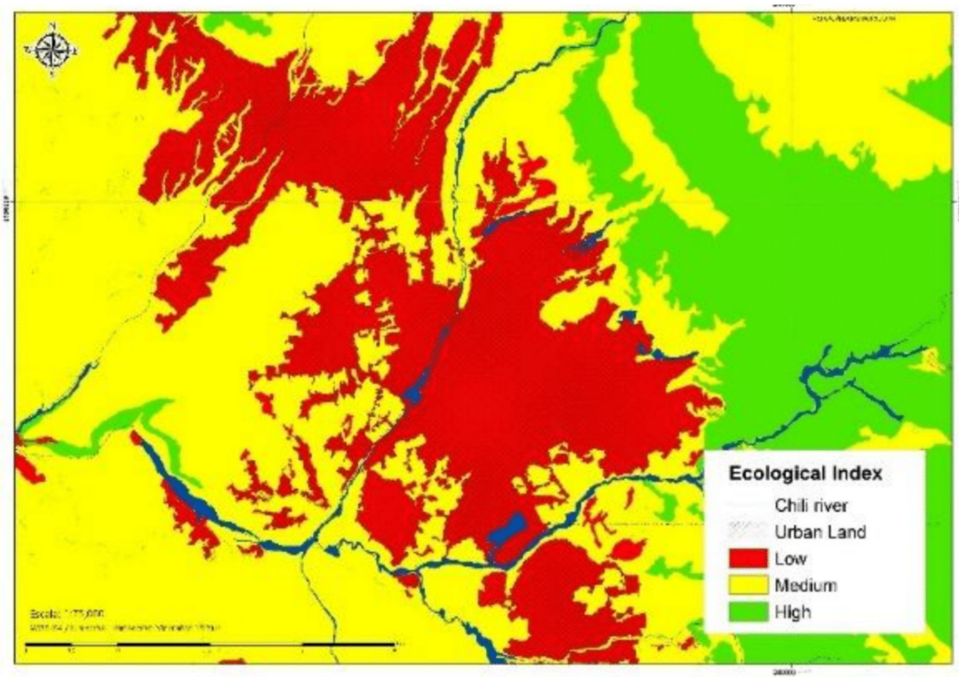

Figure 3: Ecological Index

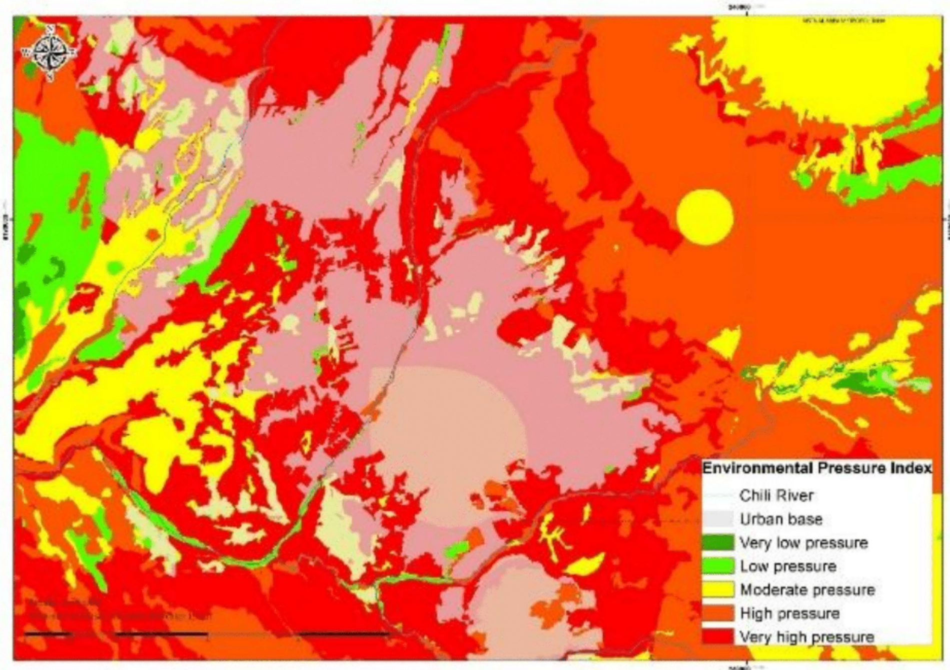

Figure 4: Environmental Pressure Index

4.3. Territorial Functionality Index (IFT)

Higher IFT values were recorded (Figure 5) in areas where regulatory functions and aquifer recharge were more strongly represented, whereas lower values were identified in urbanized and territorially transformed sectors. Across the Chili River landscape, these conditions formed a discontinuous spatial pattern, providing the functional dimension that was subsequently incorporated into the multivariable clustering analysis.

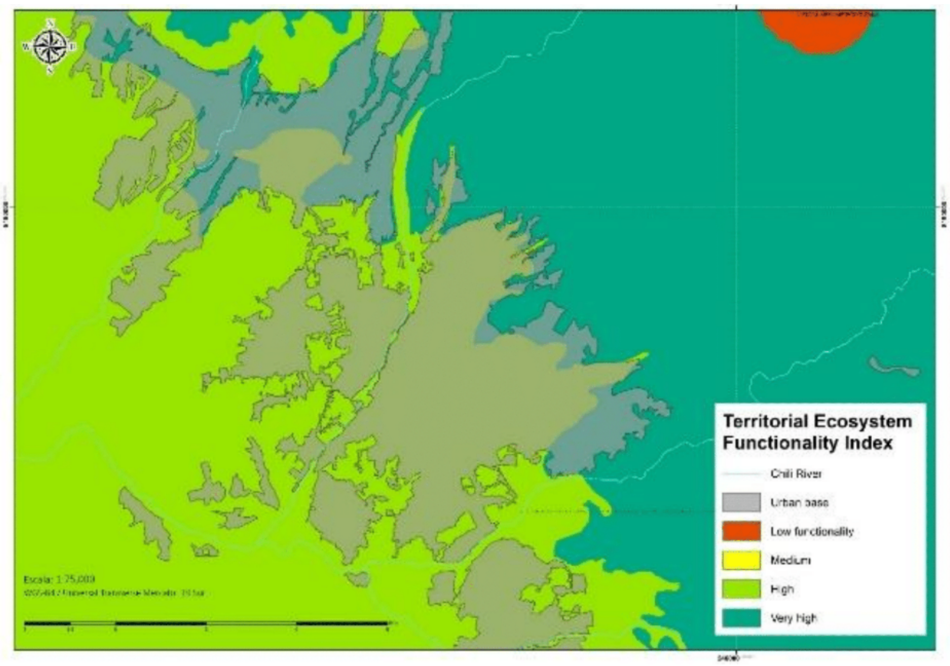

Figure 5: Territorial Ecosystem Functionality Index

4.4. Environmental Degradation Classification through AI

The comparison of alternative clustering solutions did not identify a single value of k that performed best across all criteria (Table 2). WCSS decreased progressively from $k = 2$ to $k = 6$, with smaller marginal reductions after $k = 4$. The four-cluster solution obtained a Silhouette Coefficient of 0.373 and a Davies–Bouldin Index of 0.979, whereas $k = 6$ showed better internal metrics (Silhouette = 0.411; DBI = 0.897).

Table 2
Quantitative evaluation of alternative K-means clustering solutions

Number of clusters (k)	Silhouette coefficient ↑	Davies–Bouldin index ↓	WCSS
2	0.363	1.127	945.761
3	0.410	0.984	670.941
4	0.373	0.979	506.551
5	0.361	1.096	385.662
6	0.411	0.897	296.944

Note: *Higher Silhouette and lower Davies–Bouldin values indicate better clustering performance. WCSS was interpreted based on its marginal reduction as increased.*

Comparison of the $k = 4$ and $k = 6$ partitions showed that the additional clusters mainly subdivided existing environmental configurations. One K4 cluster retained 100% correspondence with a K6 group, while two others showed dominant correspondences of 77.07% and 71.76%. The remaining K4 cluster was distributed among three K6 groups (40.85%, 35.74%, and 23.20%). Differences among these subdivisions were driven mainly by ecological conditions and environmental pressure, whereas territorial functionality varied less. On this basis, $k = 4$ was retained as the more parsimonious representation of the main environmental gradients and was subsequently interpreted as low, moderate, high, and very high degradation.

Sensitivity testing indicated that the four-cluster partition remained relatively stable under $\pm 10\%$

perturbations of the normalized indices, with a mean ARI of 0.904. The highest agreement with the baseline occurred for IFT -10% (ARI = 0.997), IPA -10% (ARI = 0.990), and IE $\pm10\%$ (ARI = 0.989), while IE -10% produced the largest change (ARI = 0.713). Across the six scenarios, Silhouette values ranged from 0.390 to 0.430 and DBI from 0.858 to 0.907.

High (46.58%) and very high (21.89%) degradation occurred mainly in consolidated and highly transformed urban sectors, where greater environmental pressure coincided with poorer ecological conditions and lower territorial functionality (Figure 6). Low degradation (12.72%) was more frequent in sectors that retained comparatively favorable ecological and functional conditions, while moderate degradation represented intermediate territorial configurations.

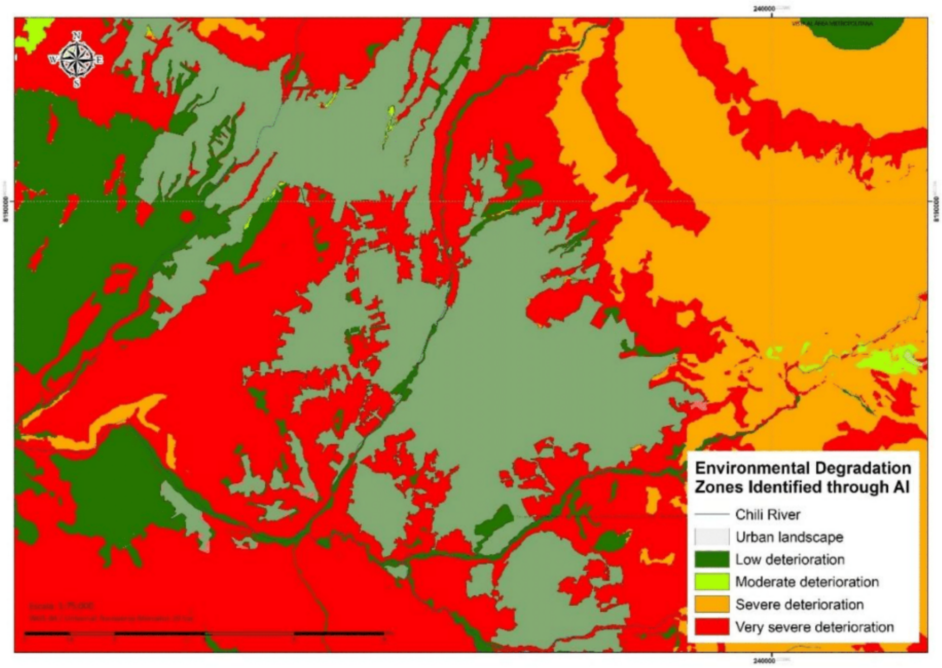

Figure 6: Environmental Degradation Zones Identified through AI

5. Discussion

The degradation identified in this study reflects the influence of anthropogenic pressure, ecological conditions, and territorial functionality across the Chili River landscape. High and very high degradation levels were concentrated mainly in urbanized and highly transformed sectors, whereas lower degradation was observed in areas with comparatively better ecological and functional conditions. This pattern is consistent with [1], which documented losses in river connectivity associated with the conversion of agricultural land into urban land. Similarly, [2] reported changes in water yield and runoff linked to land-use transformation.

The analyzed territory presents a mosaic of degraded areas. This heterogeneity is consistent with the interpretation of urban rivers as socio-ecological systems shaped by interactions between environmental and human processes [3]. In practical terms, the spatial differentiation obtained in this study can help distinguish sectors where restoration and pressure-reduction measures may be more urgent from those where existing ecological and regulatory functions warrant conservation.

From a methodological perspective, the comparison between k=4 and k=6, together with their centroid profiles, showed that cluster selection should not rely on a single statistical criterion when territorial interpretation is also an objective of the classification. This is consistent with [4], which supports the integration of GIS-based multicriteria analysis with clustering as a complement to conventional spatial assessment approaches, as well as with the increasing use of machine learning in environmental and territorial analysis [5, 6, 7, 8].

Likewise, the sensitivity analysis provides additional evidence of model stability. A mean ARI of 0.904 across the $\pm 10\%$ scenarios indicates that the main partition was largely preserved under moderate perturbations of the three normalized indices. However, stability was not identical across all scenarios: the lower agreement obtained when IE was reduced indicates greater sensitivity of the partition to this change. Therefore, these findings should be interpreted as evidence of classification stability rather than confirmation that the original weighting scheme is uniquely optimal. This distinction is important because the weights used to construct IE, IPA, and IFT were established through expert-supported GIS-MCDA criteria rather than independently calibrated against an empirical dataset.

One limitation of this study is the 30 m spatial resolution, together with the availability and scale of the thematic layers, which may constrain the representation of localized environmental processes. Cluster selection also required a balance between internal statistical performance and territorial interpretability, despite the use of WCSS, Silhouette, and DBI. In addition, the sensitivity analysis assessed the response of the classification to perturbations of IE_N, IPA_N, and IFT_N but did not independently validate the weights assigned to the variables within each index. Finally, no independent reference dataset containing predefined degradation classes was available; consequently, land-use information and UAV observations were used for contextual interpretation rather than external accuracy assessment. Future applications could address these limitations by incorporating field-based reference data, alternative or data-driven weighting schemes, and multitemporal UAV imagery.

6. Conclusion

The analysis identified marked spatial differentiation in environmental degradation across the Chili River landscape. The most degraded areas were associated with greater environmental pressure and comparatively lower ecological condition and territorial functionality, whereas areas with lower degradation retained more favorable ecological and regulatory conditions.

The integration of the normalized indices using the K-means/Iso Cluster method provided a multivariable representation of these territorial differences. The four-cluster solution was selected after comparison with alternative partitions because it preserved the main environmental gradients while providing a more parsimonious territorial classification. Its stability under $\pm 10\%$ perturbations of the normalized indices was relatively high (mean ARI = 0.904), although this result assesses classification stability rather than the optimality of the original variable weights.

Finally, the proposed methodology, which combines GIS-based Multi-Criteria Decision Analysis (MCDA) with unsupervised clustering, can serve as a transferable analytical framework rather than a fixed classification model. Its application to other river landscapes would require adapting the variables and weights to local conditions, harmonizing spatial datasets, and independently evaluating the appropriate number of clusters. In the Chili River landscape, the resulting classification provides a spatial basis for locating critical sectors and supporting the prioritization of conservation, restoration, and environmental management measures.

Declaration on Generative AI

During the preparation of this manuscript, the author used ChatGPT (OpenAI) to assist with English translation, language editing, and improvement of the clarity and organization of selected sections of the text. The research design, data processing, GIS analyses, construction of environmental indices, clustering and sensitivity analyses, interpretation of results, and scientific conclusions were developed and verified by the author. All AI-assisted content was subsequently reviewed and edited by the author, who takes full responsibility for the final content of the manuscript.

References

[1] V. Ibarcena, Impacts of urban sprawl on the ecological connectivity of landscapes in the chili river, arequipa - peru, 2023.

[2] L. Carrasco-Valencia, K. Vilca-Campana, C. Iruri-Ramos, B. Cárdenas-Pillco, A. Ollero, A. Chanove-Manrique, Effect of lulc changes on annual water yield in the urban section of the chili river, arequipa, using the invest model, Water 16 (2024) 664.

[3] A. Zingraff-Hamed, A. Serra-Llobet, G. M. Kondolf, The social, economic, and ecological drivers of planning and management of urban river parks, Frontiers in Sustainable Cities 4 (2022) 907044.

[4] T. Lillesand, R. W. Kiefer, J. Chipman, Remote sensing and image interpretation, John Wiley & Sons, 2015.

[5] D. B. Olawade, O. Z. Wada, A. O. Ige, B. I. Egbewole, A. Olojo, B. I. Oladapo, Artificial intelligence in environmental monitoring: Advancements, challenges, and future directions, Hygiene and Environmental Health Advances 12 (2024) 100114.

[6] X. X. Zhu, D. Tuia, L. Mou, G.-S. Xia, L. Zhang, F. Xu, F. Fraundorfer, Deep learning in remote sensing: A comprehensive review and list of resources, IEEE geoscience and remote sensing magazine 5 (2017) 8–36.

[7] S. Liu, K. Li, X. Liu, Z. Yin, Geospatial ai in earth observation, remote sensing, and giscience, 2023.

[8] F. Li, T. Yigitcanlar, M. Nepal, K. Nguyen, F. Dur, Machine learning and remote sensing integration for leveraging urban sustainability: A review and framework, Sustainable Cities and Society 96 (2023) 104653.

[9] J. A. Hartigan, M. A. Wong, Algorithm as 136: A k-means clustering algorithm, Journal of the royal statistical society. series c (applied statistics) 28 (1979) 100–108.

[10] J. Malczewski, Gis-based multicriteria decision analysis: a survey of the literature, International journal of geographical information science 20 (2006) 703–726.

[11] P. B. Keenan, P. Jankowski, Spatial decision support systems: Three decades on, Decision Support Systems 116 (2019) 64–76.

[12] B. Tang, H. Wang, J. Liu, W. Zhang, W. Zhao, D. Cheng, L. Zhang, L. Jiao, Identification of ecological restoration priority areas integrating ecological security and feasibility of restoration, Ecological Indicators 158 (2024) 111557.