

Article

High-Resolution Eutrophication Mapping Using Multispectral UAV Imagery and Unsupervised Classification: Assessment in the Almyros Stream (Crete, Greece)

Matenia Karagiannidou ¹, Christos Vasilakos ¹, Eleni Kokinou ^{2,*} and Nikos Gerarchakis ²

¹ Department of Geography, University of the Aegean, University Hill, 81100 Mytilene, Greece; matinakarag@gmail.com (M.K.); chvas@aegean.gr (C.V.)

² Department of Agriculture, Hellenic Mediterranean University, Estavromenos, 71410 Heraklion, Greece; gerarhakisnikos@yahoo.gr

* Correspondence: ekokinou@hmu.gr

Highlights

What are the main findings?

- UAV multispectral imagery and unsupervised clustering were combined to map eutrophication patterns in a small, heterogeneous stream system.
- The approach enables rapid, high-resolution identification of trophic gradients and supports targeted water-quality monitoring.

What are the implications of the main findings?

- The study demonstrates that UAV-based multispectral data and unsupervised machine learning can serve as a reliable and scalable alternative to dense in situ monitoring, especially in streams and rivers where traditional sampling is logistically constrained.
- The resulting eutrophication maps can guide optimized and representative sampling strategies, improve early detection of degraded water-quality hotspots, and support more informed management and protection of vulnerable aquatic ecosystems.

Abstract

Eutrophication is a form of pollution caused by elevated nutrient concentrations in water bodies, leading to excessive algal growth and subsequent oxygen depletion. This process poses significant risks to aquatic ecosystems and overall water quality. This study investigates the spatial distribution of eutrophication in the Almyros Stream, aiming to develop a rapid and high-resolution approach for identifying eutrophication patterns and selecting representative sampling sites. Almyros is an urban stream in the western Heraklion Basin (Crete, Greece) that is subjected to considerable pressures from agricultural, industrial, urban, and tourism-related activities. Data for this study were collected using a drone equipped with a multispectral sensor. The multispectral bands, together with remote sensing indices associated with chlorophyll presence, served as input data. Chlorophyll presence is a key indicator of phytoplankton biomass and is widely used as a proxy for nutrient enrichment and eutrophication intensity in aquatic ecosystems. The k-means clustering algorithm was then applied to classify the data and reveal the eutrophication spatial patterns of the study area. The results show that the methodology successfully identified spatial variations in eutrophication-related conditions and generated robust eutrophication pattern maps. These findings underscore the potential of integrating remote sensing and machine learning techniques for efficient monitoring and management of water bodies.



Academic Editor: Teodosio Lacava

Received: 27 December 2025

Revised: 24 January 2026

Accepted: 2 February 2026

Published: 4 February 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: UAV multispectral imagery; water quality monitoring; chlorophyll indices; k-means clustering; unsupervised classification; riverine environments; remote sensing

1. Introduction

Eutrophication is one of the most persistent and widespread environmental challenges affecting aquatic ecosystems on a global scale. It occurs when water bodies receive excessive inputs of nutrients—primarily phosphorus and nitrogen—often from agricultural runoff, urban wastewater, and industrial discharges. These nutrients stimulate rapid growth of algae and phytoplankton, accelerating the natural aging process of lakes and rivers. When algal biomass increases beyond ecological balance, it leads to oxygen depletion through enhanced respiration and microbial decomposition, thereby triggering a cascade of adverse effects such as fish kills, loss of biodiversity, and degradation of aquatic habitats [1–3].

Eutrophication consequences extend beyond ecological deterioration. It can also impair drinking-water quality through the proliferation of harmful algal blooms (HABs), which produce toxins that pose risks to human health, affecting skin, liver, and neurological functions [4–6]. Additionally, eutrophication imposes substantial economic costs, including reduced fisheries productivity, loss of recreational value, and increased expenses for water treatment and ecosystem restoration [7]. As climatic conditions warm and extreme precipitation events intensify, nutrient transport and bloom dynamics are also expected to worsen, further elevating global concern [2,8].

Given the complexity and spatial heterogeneity of eutrophication processes, effective monitoring requires tools capable of capturing both spatial and temporal variability. Traditional field-based sampling, while accurate, is often laboratory-intensive, spatially limited, and inadequate for rapid ecosystem assessments. Remote sensing has therefore emerged as a powerful complementary approach, offering the ability to detect changes in water quality over large areas with high temporal frequency. Satellite and UAV-based sensors can detect chlorophyll-a, suspended particulate matter, and other optical properties associated with eutrophication, enabling early-warning systems and long-term monitoring frameworks [9]. In eutrophic conditions, elevated nutrient inputs stimulate phytoplankton growth, leading to increased chlorophyll-a concentrations and changes in the optical properties of the water column that can be detected by remote sensing sensors.

Modern advancements in remote sensing increasingly rely on Artificial Intelligence (AI) and machine learning (ML) techniques to extract meaningful information from complex spectral datasets. Moreover, the integration of UAV-based multispectral imaging and machine learning provides a cost-effective, flexible, and near-real-time approach for assessing water quality, offering significant advantages over traditional manual monitoring and satellite remote sensing methods. UAV multispectral technology effectively captures the spectral response characteristics of specific water quality parameters, overcoming the spatial and temporal resolution limitations commonly associated with satellite data. Furthermore, ML algorithms can automatically learn complex patterns and extract meaningful features from multispectral imagery, enabling the accurate estimation of both optical and non-optical water quality parameters. For example, in the study by [6], it was found that the Random Forest (RF) model outperformed other ML algorithms in predicting water quality parameters like nitrite, nitrate, chlorophyll-a, phosphate, and suspended solids. That study also captured seasonal variations in water quality, with nutrient concentrations peaking in summer and declining in autumn, while chlorophyll-a reached its highest in autumn. According to Li et al. (2026), the k-means clustering algorithm is widely used among unsupervised learning approaches due to its computational simplicity and robustness [10].

By grouping pixels with similar spectral characteristics, k-means enables automated classification of water areas into categories reflecting varying degrees of eutrophication [9]. Over time, k-means has been proposed in several versions [11], and it has been applied in many research topics [12–14].

Despite these advantages, the use of UAV-based remote sensing and unsupervised clustering techniques such as k-means also presents certain limitations. UAV data acquisition is sensitive to weather conditions, illumination variability, and flight altitude, which may introduce noise or inconsistencies in the spectral signal. Additionally, UAV surveys typically cover limited spatial extents, requiring careful planning to ensure representativeness. Regarding k-means clustering, the method requires the a priori selection of the number of clusters, which may influence the resulting classification and introduce subjectivity. Moreover, k-means assumes spherical cluster shapes and equal variance, which may not fully capture the complexity of natural aquatic systems. Finally, the algorithm is sensitive to outliers and noise, potentially leading to the formation of artefact clusters in areas affected by shallow water, turbulence, or mixed surface conditions.

The aim of this study is to (a) analyze high-resolution multispectral data collected by an unmanned aerial vehicle (UAV), (b) compute remote sensing indices sensitive to chlorophyll-a, and (c) identify and classify water areas according to their relative levels of eutrophication by applying the k-means clustering algorithm. By integrating remote sensing and machine learning, this work contributes to the development of efficient, cost-effective tools for monitoring, visualizing, and managing eutrophication in aquatic ecosystems. Overall, the study demonstrates that UAV-based multispectral imaging—combined with machine-learning classification—offers a dynamic and scalable approach to environmental monitoring, particularly for regions where conventional sampling is logistically challenging or insufficient.

Research Site

The study area is Almyros, a suburban stream located in the western basin of Heraklion (Crete, Greece), approximately 8 km northwest of the city center (Figure 1). It is noted for its ecological significance and karst hydrological uniqueness [15].

The stream is about 1.8 km long, with a width ranging from 5 to 20 m [16]. The Almyros ecosystem holds significant environmental value and is included among UNESCO-protected natural sites [17]. At the stream's source, a dam has been constructed to increase aquifer pressure, prevent seawater intrusion, and improve groundwater quality. Despite these measures, the stream's water exhibits high salinity due to the mixing of seawater within the aquifer that feeds the spring. Additionally, the karstic nature of the region complicates water management, as water moves rapidly through subterranean channels with minimal natural filtration [16–18].

Although the broader Almyros area (Figure 1) does not exhibit high levels of urbanization, it is subject to considerable anthropogenic pressures [15] according to the Impervious Density Index [14]. Agricultural activities dominate the eastern part of the area near the dam and are closely linked to eutrophication phenomena observed in the stream [18]. In the northern coastal zone, land uses related to tourism and recreation have become increasingly established [18].

Overall, the Almyros Stream (Figure 1) represents a valuable yet highly vulnerable hydrological system, playing an important role both in Crete's natural ecology and in the local water supply. However, ongoing pressures from agriculture, tourism development, and water extraction underscore the need for effective management and protection strategies to ensure its long-term sustainability [18].

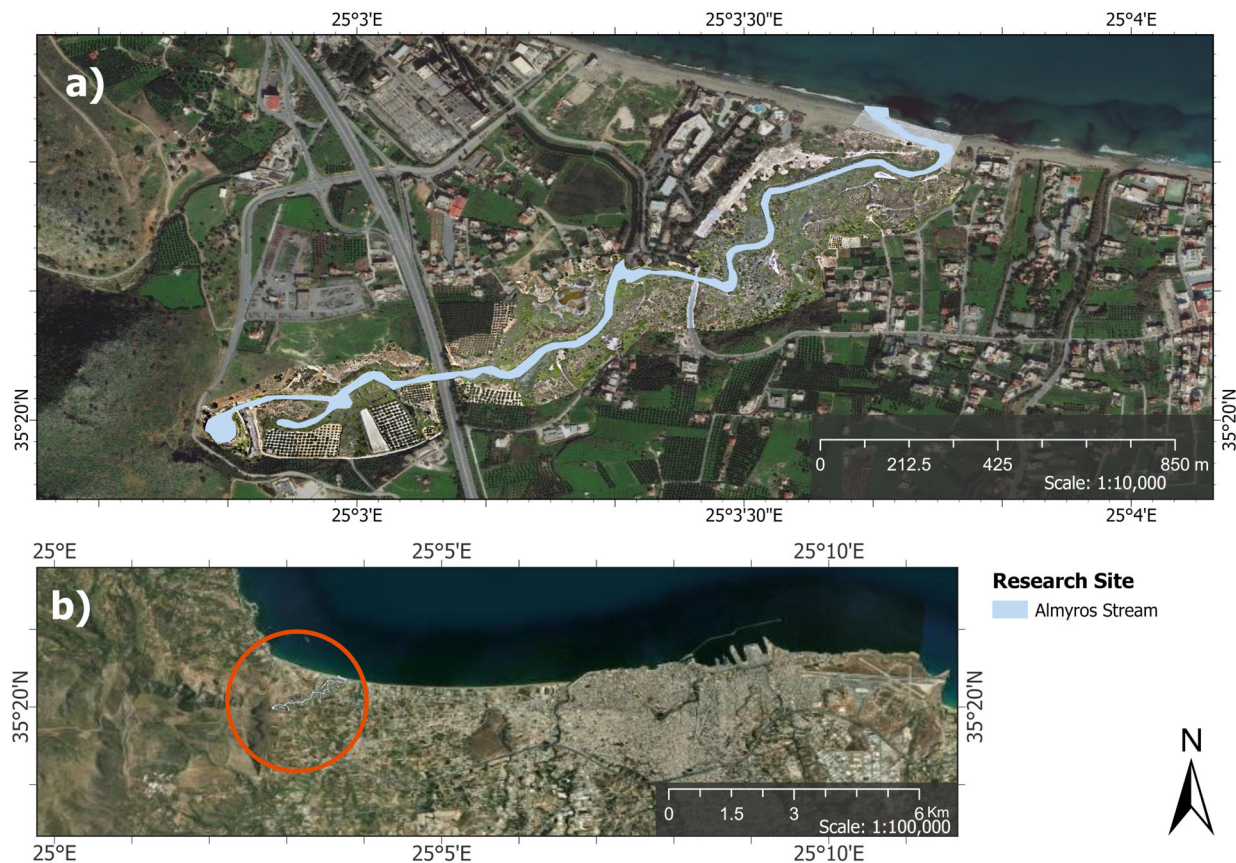


Figure 1. Location of the study area. (a) Zoomed-in view of the study area showing the Almyros Stream and the spatial extent of the analyzed stream section. (b) Regional setting of the Almyros Stream within the Heraklion Basin, Crete, Greece. The red circle indicates the location of the study area shown in (a).

2. Materials and Methods

The methodological approach adopted in this study comprises several sequential stages (Figure 2), focusing on the calculation of remote sensing indices, clustering analysis, and the environmental interpretation of the results.

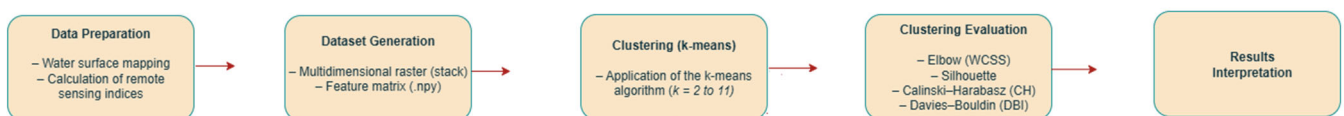


Figure 2. Flowchart of the methodological workflow applied in this study, including data preparation, dataset construction, k-means clustering, clustering evaluation, and result interpretation.

2.1. Drone Data

The dataset was collected using a DJI Mavic 3M aerial vehicle manufactured by DJI, based in Shenzhen, China. The platform is equipped with an integrated high-resolution RGB camera and multispectral sensor comprising four spectral bands (red, green, red-edge, and near-infrared [NIR]) [19], both developed and manufactured by DJI. Drone flights were conducted on 4–5 June 2024. Following data acquisition, photogrammetric processing and parameter extraction were performed, generating the initial inputs for subsequent analysis.

2.2. Data Processing

The first step in data processing was the isolation of water areas from the dataset for all multispectral images. For this process, the NDWI index was applied [20,21]:

$$\text{NDWI} = \frac{\text{Green} - \text{NIR}}{\text{Green} + \text{NIR}} \quad (1)$$

Subsequently, the generated multispectral images were used to calculate remote sensing indices. The selected indices are related to the presence of chlorophyll in the water and by extension, to eutrophication [20–23]. The indices selected are as follows [20–22,24]:

$$\text{MCI} = \text{Red Edge} - \left(\text{NIR} + (\text{NIR} + \text{Red Edge}) \frac{\text{Red Edge} - \text{Red}}{\text{NIR} - \text{Red}} \right) \quad (2)$$

$$\text{CI} = \text{Red Edge} - \left(\text{Green} + (\text{NIR} - \text{Green}) \frac{\text{Red Edge} - \text{Green}}{\text{NIR} - \text{Green}} \right) \quad (3)$$

$$\text{NDAI} = \left(\frac{\text{NIR} - \text{Red}}{\text{NIR} + \text{Red}} \right) \quad (4)$$

$$\text{NDCI} = \left(\frac{\text{Red Edge} - \text{Red}}{\text{Red Edge} + \text{Red}} \right) \quad (5)$$

The final dataset used as input for the clustering algorithm comprised four multispectral bands (red, green, red-edge, and NIR) and four eutrophication indices (NDCI, NDAI, CI and MCI).

2.3. Dataset Generation

Once all indices were computed, eight raster layers—each with identical dimensions and coordinate reference system (CRS)—were stacked to form a multidimensional dataset. Each raster represented one variable, and each pixel contained the corresponding values from all layers. A feature matrix was then generated in which each row represented a pixel, and each column represented each of the eight variables. This consolidated matrix served as the input for clustering and was stored in binary format for computational efficiency.

2.4. Clustering

Different clustering methods have been proposed and applied for unsupervised classification [25]. These methods vary in computational efficiency and sensitivity to outliers, while others provide more stable and interpretable results and improved robustness. In the present study, the k-means algorithm was chosen over other available methods. K-means offers simplicity, speed, and strong interpretability, even if it is less robust to outliers [26,27]. Furthermore, k-means is sensitive to initialization compared to alternatives like K-medoids, K-medians, and density-based methods [28]. Therefore, multiple initializations and four different methods were applied to determine the optimal number of clusters. Conclusively, even if k-means limitations are well known, it is the most used clustering algorithm [29]. In this study, where the number of samples was more than 18 million pixels, the memory required for clustering the four-dimensional feature space into k clusters is practically impossible to be allocated for algorithms such as K-medoids even when using a reduced subset of samples.

Clustering was performed in Python using the k-means algorithm. The implementation was structured to: (a) maintain spatial consistency, (b) produce georeferenced raster outputs, and (c) compute objective clustering quality metrics.

The main k-means parameters were [23,30]:

- number of clusters (k),
- randomness,
- number of initializations,
- use of the Elkan optimization algorithm.

The algorithm produced a cluster label for each pixel and a centroid table summarizing the mean values of each cluster [21].

Clustering Evaluation

To evaluate clustering quality, the following indices were computed [30–33]:

- Inertia (Within-Cluster Sum of Squares, WCSS),
- Calinski–Harabasz Index (CH),
- Davies–Bouldin Index (DBI),
- Silhouette Coefficient.

The upper limit of the number of clusters was defined as $k = 11$ for methodological and environmental reasons. First, the dataset corresponds to a relatively small and spatially constrained stream system, where a very high number of clusters would lead to excessive fragmentation and the production of classes with no ecological significance. Preliminary tests for $k > 11$ (e.g., $k = 12–15$) produced highly unstable and spatially incoherent clusters, dominated by noise rather than by meaningful spectral patterns. Second, in unsupervised classification of aquatic environments, the recommended range of k is typically limited, as the number of ecologically interpretable water-quality states is small. Relevant in remote-sensing-based water classification, it is common to evaluate k within similar ranges [11–18], since higher values rarely correspond to distinct biophysical characteristics [34–36]. Finally, the range $2 \leq k \leq 11$ ensured a balance between:

- (a) capturing meaningful spectral variability, and
- (b) avoiding over-segmentation, which would reduce interpretability and environmental relevance.

Therefore, $k = 11$ was selected as the maximum value beyond which the clustering no longer produced coherent or physically meaningful classes.

All data processing and clustering procedures were implemented in Python 3.10, using NumPy (version 1.24.4), Pandas (version 2.3.2), Scikit-Learn (version 1.3.2), Rasterio (version 1.4.3), and Matplotlib (version 3.7.5). Geospatial validation, raster inspection, and map production were carried out in ArcGIS Pro 3.5.4.

3. Results

In this section, the results obtained from the application of the k-means algorithm are presented. The primary objective of the clustering procedure was to achieve a spatial subdivision of water pixels into categories with distinct characteristics, enabling the assessment of eutrophication levels.

3.1. Cartographic Results

The cartographic analysis of the k-means clustering results highlights three key findings. First, the spatial representation of eutrophication is highly sensitive to the selected number of clusters. Second, low k values lead to excessive spatial generalization, whereas high k values result in over-fragmentation and reduced interpretability. Third, intermediate k values provide the most balanced representation of spatial variability, effectively capturing eutrophication patterns while preserving spatial coherence. The implementation of the k-means algorithm produced raster outputs in which each pixel was assigned to one of

the derived clusters. These cluster maps illustrate the spatial distribution of the identified groups across the water surface, offering an initial depiction of the spatial variability of eutrophication. The following paragraphs provide a detailed cartographic analysis of these findings, examining how different k values influence the spatial representation and interpretability of eutrophication patterns.

The cartographic outputs reveal that the level of spatial detail varies substantially with the choice of k . For lower values ($k = 2\text{--}4$), the resulting maps exhibit highly generalized spatial patterns, characterized by extensive homogeneous areas and a limited ability to distinguish finer spatial variations (Figure A1, Appendix A). At this level, the clusters represent only the most prominent spatial contrasts, failing to capture internal heterogeneity within the water body. As a result, these maps provide a useful general overview but are insufficient for detailed spatial interpretation.

At higher k values ($k = 8\text{--}11$), the clusters become increasingly fragmented, producing overly detailed maps composed of small, isolated spatial units (Figures A2–A5, Appendix A). While this high granularity allows subtle differences to be detected, it often introduces excessive complexity. Previous studies have shown that the spatial distribution of chlorophyll-*a* tends to be correlated and exhibits a clustered distribution [37,38]. In many cases, minor data fluctuations generate isolated clusters that do not correspond to meaningful environmental distinctions, indicating over-fragmentation of the dataset. Between these two extremes, intermediate values ($k = 5\text{--}7$) offer a more balanced spatial representation (Figures A6 and A7, Appendix A). Within this range, the clusters clearly reflect spatial variability while avoiding excessive fragmentation. Consequently, these k values strike a desirable balance by providing sufficient spatial detail while maintaining interpretability.

Overall, the cartographic analysis demonstrates that low k values lead to over-generalization, whereas high k values introduce unnecessary detail and fragmentation. Intermediate k values yield the most coherent and meaningful representation of spatial eutrophication patterns.

3.2. Assessment of Clustering Quality

Evaluating clustering quality is a critical component of any unsupervised learning workflow, as it supports the identification of the number of clusters that yield reliable and meaningful data partitions. In this study, the evaluation was based on a combination of graphical interpretations and quantitative metrics, with the goal of minimizing subjectivity and strengthening the robustness of the results.

Graphical representations were produced to visualize the behavior of the k -means algorithm across different values of k . Figure 3 presents the computed evaluation curves derived from the Elbow method, the Silhouette coefficient, the Calinski–Harabasz (CH) index, and the Davies–Bouldin (DBI) index.

In Figure 3, Graph 1 corresponds to the Elbow plot, which illustrates the variation in within-cluster inertia (WCSS) as the number of clusters increases [9,31,39]. As expected, inertia value decreases progressively with increasing k values. The “elbow” point indicates where further reductions in inertia become less significant, providing an indication of the optimal number of clusters.

Graph 2 in Figure 3 presents the Silhouette index, which incorporates both cluster cohesion and separation [9,32]. The highest Silhouette values appear at $k = 2$, indicating strong clustering performance but limited spatial differentiation. For $k = 5\text{--}6$, the index reaches moderate and stable values, suggesting a more balanced structure that captures spatial variability without excessive fragmentation. At higher k values, the Silhouette score declines, reflecting reduced coherence and a tendency toward over-segmentation.

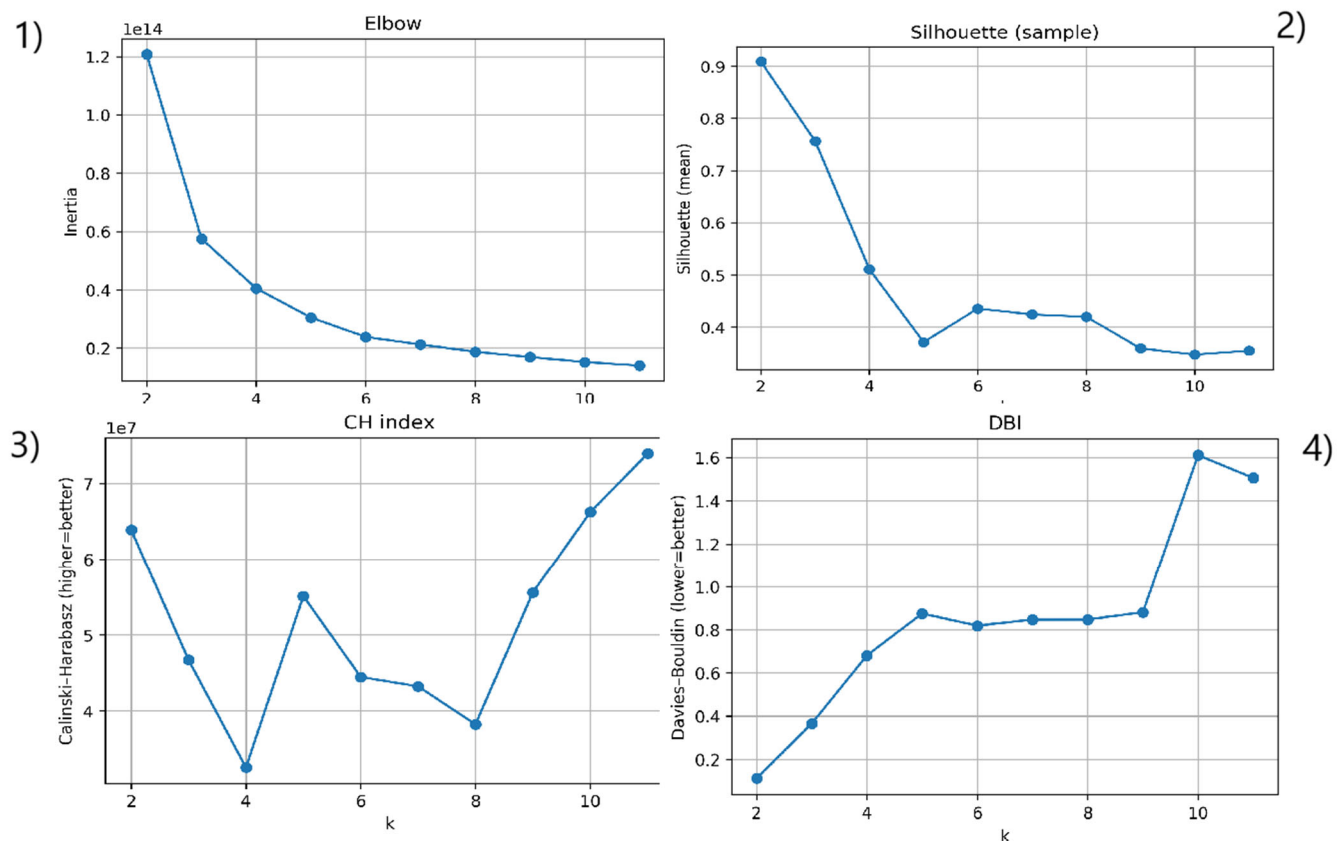


Figure 3. Evaluation of k-means clustering performance for different numbers of clusters ($k = 2–11$). (1) Elbow plot showing the variation of within-cluster sum of squares (WCSS) for different values of k , (2) Silhouette plot showing the average Silhouette coefficient for different numbers of clusters (k), (3) Calinski–Harabasz index (CH) values for different numbers of clusters (k), indicating the balance between intra-cluster cohesion and inter-cluster separation, (4) Davies–Bouldin Index (DBI) values for different numbers of clusters (k), expressing the degree of similarity between clusters.

Graph 3 in Figure 3 shows the Calinski–Harabasz (CH) index, which expresses the ratio between cluster dispersion and within-cluster compactness [31]. Elevated CH values are observed for $k > 5$, with particularly high values at $k = 10–11$, indicating improved separability but also suggesting increasingly complex and fragmented cluster structures.

Graph 4 in Figure 3 displays the Davies–Bouldin Index (DBI), which evaluates cluster similarity [9,31]. Lower DBI values, observed at $k = 2–3$, correspond to well-separated and well-defined clusters. Increasing DBI values indicate reduced cluster distinction and lower classification quality.

The combined interpretation of all evaluation metrics indicates that lower k values ($k = 2–4$) yield high cohesion and strong separation but provide insufficient spatial detail. Conversely, higher k values ($k = 8–11$) lead to overly fragmented classifications that reduce interpretability. The intermediate range ($k = 5–7$) achieves a more balanced outcome, preserving meaningful spatial differentiation while maintaining adequate cohesiveness.

Based on these evaluation indicators, the optimal number of clusters was determined to be $k = 5$, as this value provides the best compromise between cohesion and separation. The Elbow curve shows a pronounced reduction in inertia up to $k = 5$, after which the curve begins to flatten, indicating diminishing returns and an increased risk of over-fragmentation. Although the Silhouette index reaches its highest value at lower k , these correspond to overly generalized maps; at $k = 5$, Silhouette values remain acceptable, while they decline for $k > 5$. The Calinski–Harabasz index attains one of its highest values at $k = 5$, suggesting optimal separation relative to compactness. Finally, although the

Davies–Bouldin Index is lowest at lower k values, it remains comparatively low at $k = 5$ before increasing substantially at higher values ($k = 8–11$). Taken together, all metrics converge on $k = 5$ as the most appropriate solution, achieving an effective balance between cluster cohesion, separation, and spatial interpretability.

Next, the relationships among the five clusters are further examined using a dendrogram and a heatmap (Figures 4 and 5), providing additional insight into the distances among them and supporting the validity of the final classification.

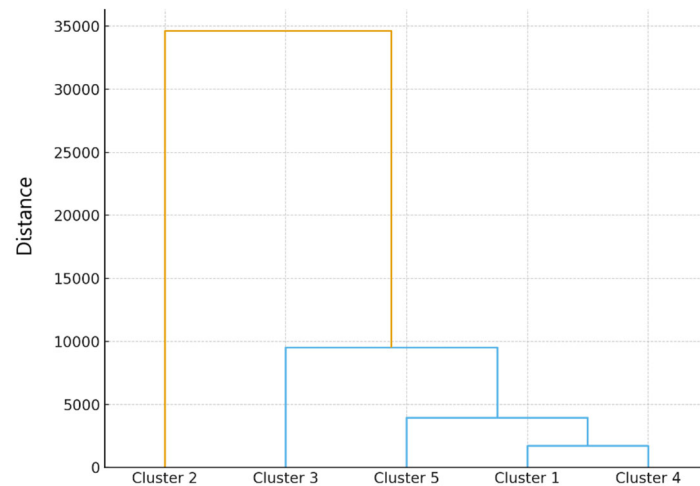


Figure 4. Dendrogram of distances among clusters for the case of $k = 5$. The diagram illustrates the relative similarity of the clusters based on their pairwise distances.

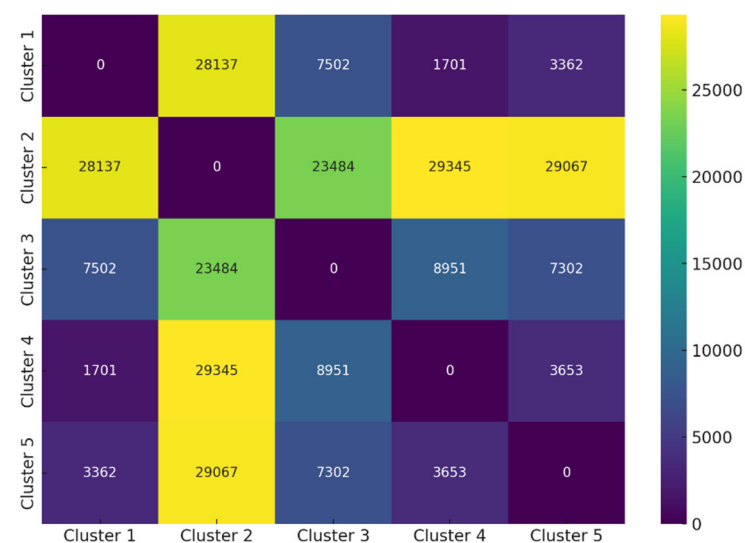


Figure 5. Heatmap of distances among clusters for the case of $k = 5$. The numerical values and color gradients indicate the degree of separation or similarity between clusters.

Figure 4 presents the dendrogram, which hierarchically visualizes the relative distances between clusters and the sequence in which they merge [40,41]. The structure of the dendrogram shows that clusters 1 and 4 exhibit the smallest distance, indicating the highest degree of similarity, whereas cluster 2 appears more distinct from the remaining clusters.

Figure 5 shows the heatmap, which provides a complementary, quantitative representation of the pairwise distances between clusters [42]. Darker colors correspond to smaller distances—such as those between clusters 1–4 and 4–5, while lighter shades indicate larger dissimilarities.

Overall, the combined evaluation of all clustering indices indicated that $k = 5$ constitutes the most suitable number of clusters, as it achieves an optimal balance between spatial detail, internal cohesion, and separation among groups. This conclusion is further supported by the spatial patterns derived from the cartographic interpretation (Figure 6).

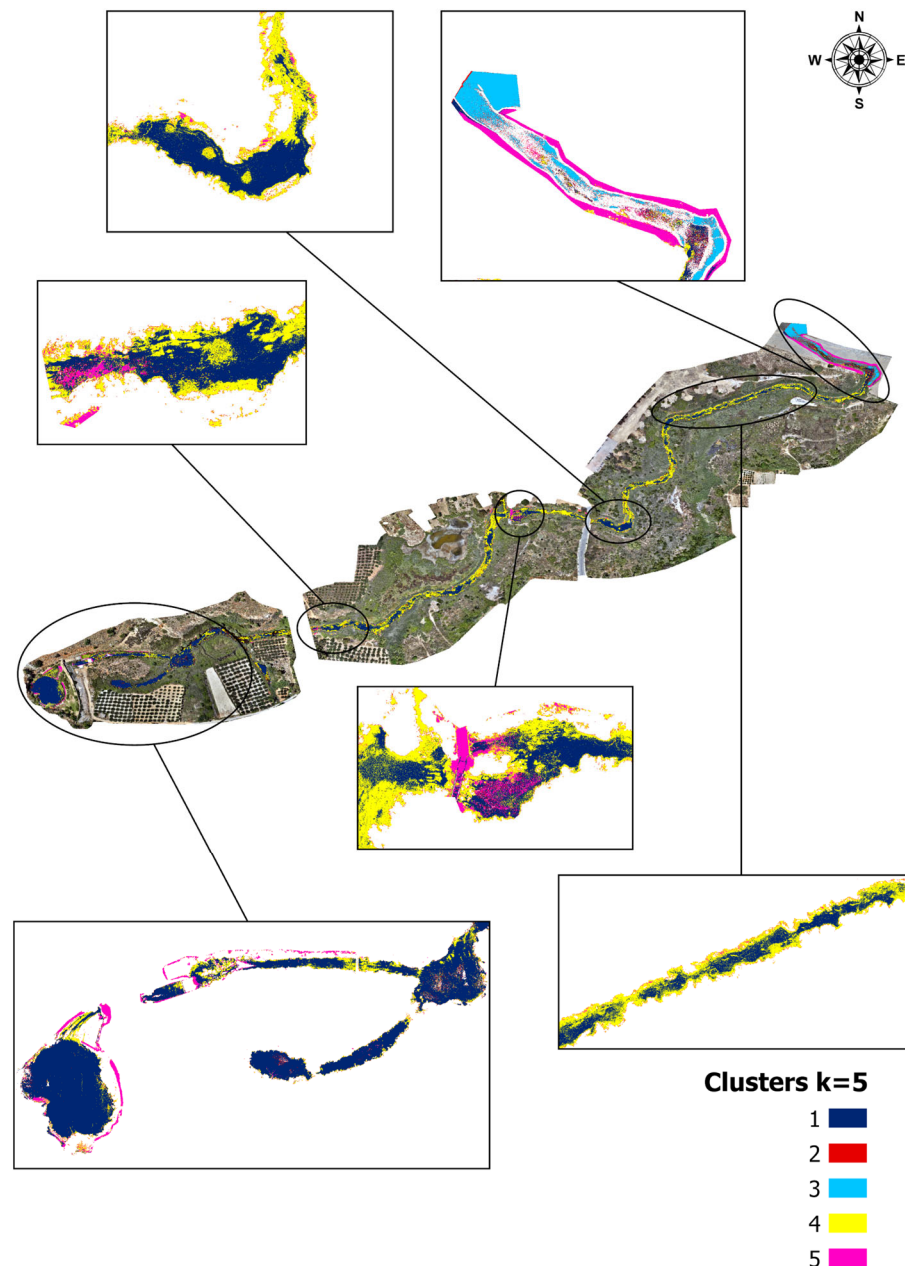


Figure 6. Cartographic representation of the k-means clustering results for $k = 5$ outputs. The black circles indicate specific locations along the stream where detailed zoom-in views were produced to examine how the clusters change spatially within the water body. These magnified areas allow a clearer interpretation of the variation in spectral characteristics and the transitions between different eutrophication levels along the stream course.

3.3. Interpretation of the Results

The application of the k-means algorithm produced five clusters, each representing a different level of eutrophication within the aquatic environment of Almyros Stream. This classification is based exclusively on remote sensing variables related to chlorophyll concentration and water optical properties, providing a comprehensive spatial depiction of the eutrophication.

As demonstrated in the preceding evaluation, $k = 5$ was identified as the most appropriate number of clusters. The following interpretation links the numerical and cartographic outputs to environmental characteristics associated with eutrophication, thereby assigning practical meaning to the clustering results and enhancing their usefulness as a monitoring tool for aquatic ecosystems.

3.3.1. Visual Representation

Cluster 5 represents a special case, as it does not correspond to an actual eutrophication level. Instead, it includes areas where the water mask failed to accurately isolate water pixels. The adjacency effect—namely, the influence of surrounding land on water pixels that leads to unreliable spectral responses—is a well-known limitation in remote sensing-based water quality studies [43,44]. As a result, this cluster contains several types of misclassified regions, including areas of high-water clarity (where shallow depths allow the streambed to be visible and the spectral signal is dominated by bottom reflectance), riparian vegetation [45], and zones of strong turbulence [46]. Such conditions commonly occur beneath bridges, near rocky formations at the dam outflow, and in locations where high flow velocities disrupt the optical signal from the water column.

In some cases, non-water surfaces were also assigned to this cluster. For example, bridge structures were misclassified due to their metal surface reflecting light in a manner distinctly different from water [38]. Visual inspection indicates that these artefacts are particularly concentrated near the stream mouth, where shallow water depth and complex lighting conditions increase the likelihood of misclassification [46].

The classification of cluster 5 as an “artefact” category is therefore supported primarily by visual assessment. Features such as shallow water, vegetated banks, turbulent flow, and artificial surfaces (e.g., bridges) cannot be reliably characterized using remote sensing indices alone. The corresponding images (Figures 7–11) illustrate these conditions and clearly demonstrate the distinct nature of this cluster.



Figure 7. General cartographic representation of the fifth cluster (artefacts). Panels (a–d) indicate the main categories of areas that were incorrectly assigned to this cluster.

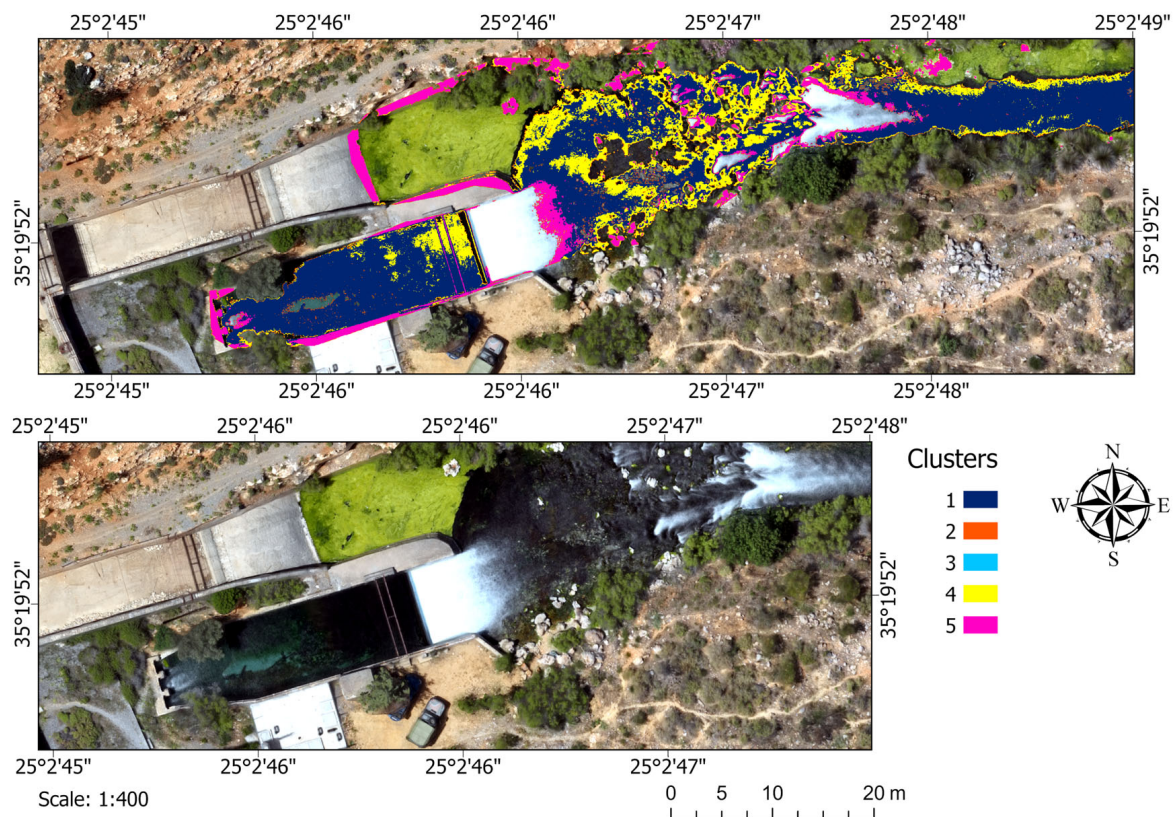


Figure 8. Example of area (Figure 7a) zones of intense water turbulence caused by the stream's flow and momentum.

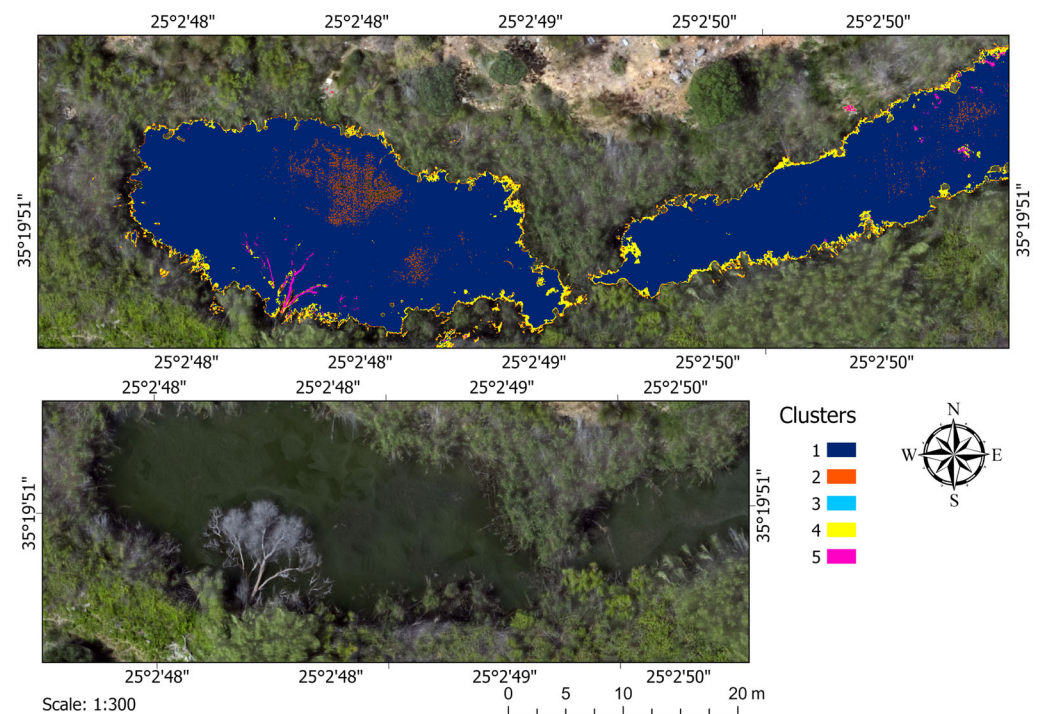


Figure 9. Example of area (Figure 7b)—riparian vegetation that was assigned to the fifth cluster due to spectral similarities.

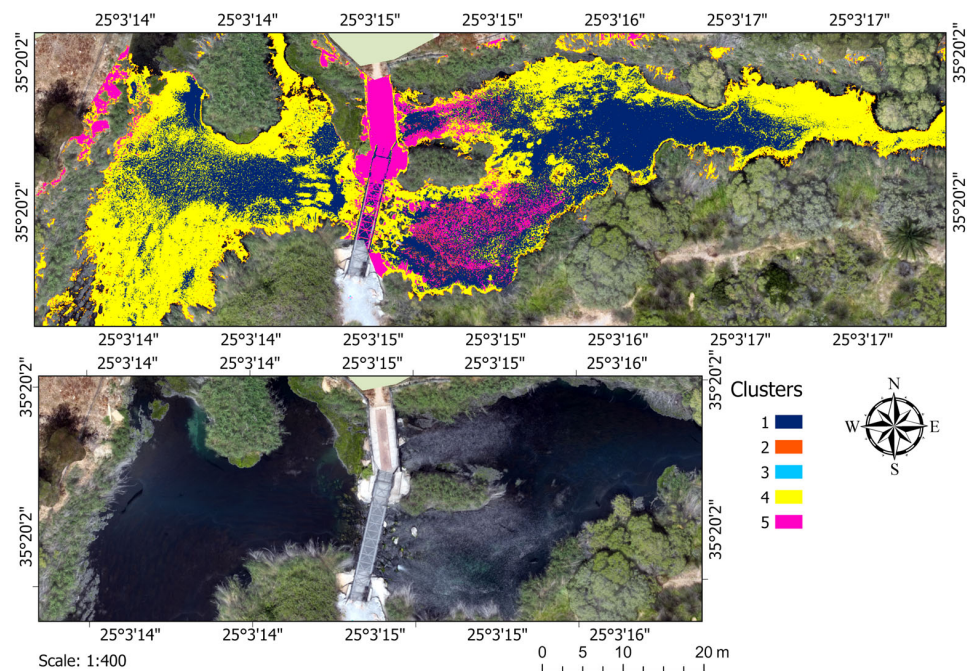


Figure 10. Example of area (Figure 7c) the bridge, where the metal structure was classified into the fifth cluster.

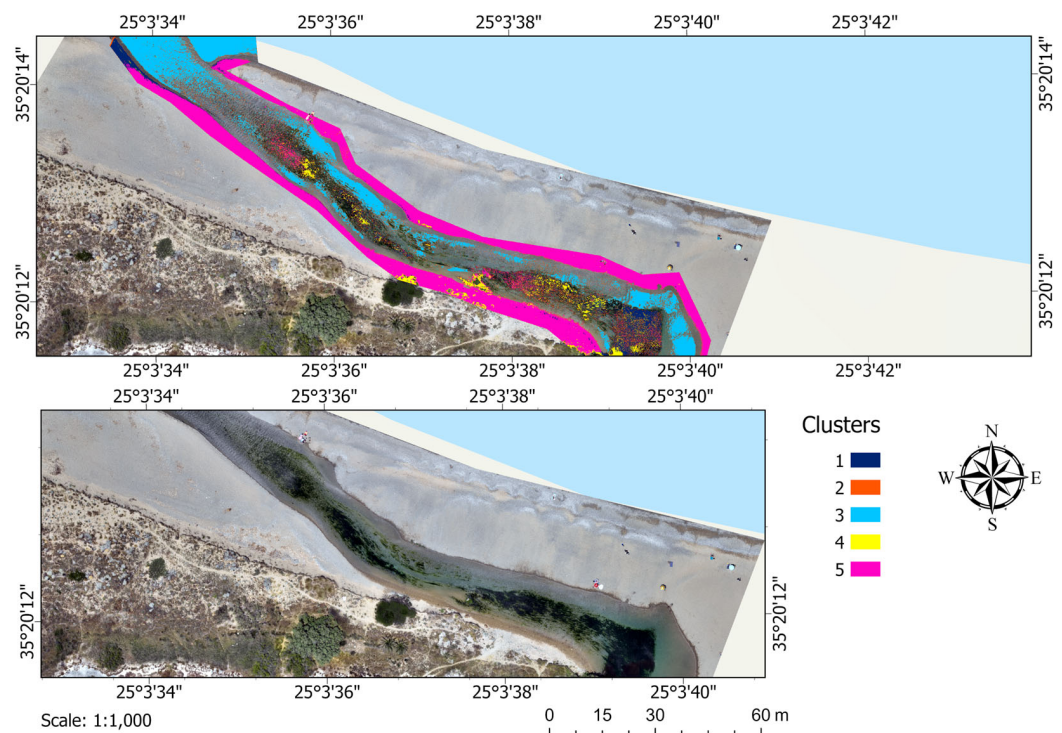


Figure 11. Example of area (Figure 7d) shallow, highly transparent water where surface reflectance leads to classification outside the actual eutrophication level.

Before proceeding with the detailed interpretation of the clusters generated by the algorithm, it is necessary to outline certain assumptions that provide a conceptual framework for interpreting the results. In areas with greater water depth and higher flow velocities, chlorophyll concentrations are generally expected to be lower, as continuous water renewal and increased turbulence inhibit phytoplankton growth. Similarly, lower chlorophyll values are anticipated in the estuarine zone, where mixing with seawater typically creates conditions that are less favorable for eutrophication [47–49].

Conversely, elevated chlorophyll concentrations are more likely to occur in shallow waters or in places with low velocity flow. Such stagnant or weakly flowing conditions restrict water renewal, creating an environment conducive to increased chlorophyll production and the development of eutrophication [50]. These assumptions provide an interpretative framework that can be meaningfully combined with the numerical and cartographic outputs of the clustering process, as demonstrated by the spatial correspondence between cluster distribution patterns, trophic scores, and the hydrological conditions observed in the results.

For clarity of interpretation, the five clusters were grouped into four eutrophication levels (1 = low, 4 = high), while Cluster 5 was designated as an artefact (Table 1).

Table 1. Classification of the clusters into eutrophication grades based on the mapping results.

Cluster	Environmental Classification
Cluster 1	3
Cluster 2	2
Cluster 3	1
Cluster 4	4
Cluster 5	Artefact

3.3.2. Chlorophyll-Based Representation

To quantitatively depict eutrophication levels, a composite index—referred to as the Trophic Score (*TS*)—was calculated for each cluster (Table 2). The *TS* was defined as the arithmetic mean of the normalized remote sensing indices, as expressed in Equation (6).

$$TS = \frac{1}{n} \sum_{i=1}^n I_i \quad (6)$$

where I_i represents the normalized value of the i -th remote-sensing index, and n is the total number of indices used. This composite formulation follows a commonly used approach in remote-sensing-based water quality studies, where normalized indices are equally weighted to derive an integrated indicator [34,51]. This procedure consisted of several steps designed to ensure an objective and balanced comparison among clusters.

In the first step, the mean value of each remote-sensing index was computed for every cluster. These indices capture the spectral signatures associated with chlorophyll presence and collectively describe the optical profile of each cluster. The resulting average values provide a concise representation of the trophic characteristics of the water pixels assigned to each group. In the subsequent step, the normalized index values were aggregated into a single composite measure by calculating their mean value. This process yielded the *TS* (Table 2), an integrated metric that expresses the trophic intensity of each cluster by equally incorporating information from all indices. This approach reduces dependence on any single index and provides a clearer, more robust assessment of eutrophication levels.

Based on the *TS* (Table 2), each cluster was assigned to one of four environmental categories, ranging from 1 (low eutrophication) to 4 (high eutrophication). Specifically, Cluster 2 exhibited near-zero index value and the lowest *TS*, placing it in Level 1. Cluster 1 showed moderately elevated index values, corresponding to Level 3, while Cluster 4 exhibited the highest index values and the highest *TS* and was therefore classified as Level 4 (Table 2).

Table 2. Environmental classification of clusters according to the Trophic Score.

Cluster	Trophic Score	Environmental Classification
Cluster 1	1.357797593	3
Cluster 2	0.002306009	1
Cluster 3	1.0999000216	2
Cluster 4	1.423457742	4
Cluster 5	Artefact	Artefact

3.3.3. Representation Based on Distance

The cluster distance analysis constituted the next step in understanding the relationships among clusters and in validating the proposed environmental classification. Distances (Figures 4 and 5) were calculated using the cluster centroids, providing a measure of similarity and dissimilarity among the groups [40,41]. These distances were not used to redefine the classification but rather to confirm the distinction among the proposed categories and to document the relative positioning of each cluster within the overall trophic scheme. In this sense, they serve as a complementary tool to the *TS*, highlighting which clusters are closely related and which are clearly differentiated, thereby offering a more comprehensive understanding of the classification structure.

The results in Table 3 indicate that cluster 1 is relatively close to cluster 4, suggesting that both represent the higher end of the eutrophication gradient, with cluster 4 corresponding to the most intense conditions. Cluster 3 occupies an intermediate position, forming a bridge between cluster 1 and cluster 2, which exhibits the lowest trophic values. This pattern identifies cluster 3 as a transitional category, connecting low to moderate eutrophication levels. By contrast, the largest distance is observed between clusters 2 and 4, confirming the strong contrast between the least and most eutrophic conditions within the stream system. Accordingly, the numerical labels 1–4 are interpreted as environmental classes consistent with the previous classification scheme (1 = low, 4 = high).

Table 3. Eutrophication level classification of clusters using inter-cluster distance.

Cluster	Environmental Classification
Cluster 1	3
Cluster 2	1
Cluster 3	2
Cluster 4	4
Cluster 5	Artefact

3.3.4. Environmental Characterization of Clusters

To convert the numerical output of the k-means algorithm into qualitative eutrophication categories, a two-step procedure was implemented: (a) an initial ranking based on the *TS*, and (b) confirmation or refinement of the classification using the distances among clusters.

(a) Initial ranking based on *TS* values

The *TS* was calculated for all clusters (excluding cluster 5, which was identified as an artefact). The four valid clusters were then ranked from lowest to highest trophic intensity according to their mean *TS* values:

- $TS_2 = 0.0023$

- $TS_3 = 1.0990$
- $TS_1 = 1.3578$
- $TS_4 = 1.4235$

This yields the following order:

$$TS_2 < TS_3 < TS_1 < TS_4$$

To derive the corresponding eutrophication classes, linear interpolation was applied to generate three classes ($n = 3$), corresponding to Low, Medium, and High eutrophication levels, as shown in Table 4. At this stage, the Medium class represents an intermediate trophic condition; its subsequent subdivision into Medium and Medium–High levels was performed later, based on inter-cluster distance analysis. These class thresholds provided the initial boundaries for assigning each cluster to one of the four eutrophication levels.

Table 4. Category thresholds (Low, Medium, High) obtained through quantile.

Category	Trophic Score Range
Low	0.0000–0.2765
Medium	0.2766–1.4070
High	≥ 1.4071

(b) Confirmation using distances

The distance-based analysis provided an additional validation of the environmental classification derived from the TS (Table 5). The results show that Cluster 2 is clearly isolated, exhibiting both very low index values and large distances from the other clusters. This confirms its classification as Low eutrophication (Table 5). At the opposite end of the gradient, Cluster 4 displays the greatest dissimilarity from the remaining clusters and is positioned at the upper extreme of the distribution (Table 5). This supports its classification as High, representing the most eutrophic conditions. Cluster 1 occupies an intermediate position, showing moderate distances from both Cluster 3 and Cluster 4 (Table 5). When combined with its TS , this pattern indicates a medium eutrophication level, with a tendency toward higher trophic intensity. Cluster 3 is between Cluster 2 and the group formed by Clusters 1 and 4 (Table 5). Although its TS places it numerically closer to lower levels, the distance analysis reveals stronger affinity with the medium-to-high portion of the gradient. This converging evidence supports its final classification as Medium, identifying it as a transitional cluster connecting lower and higher eutrophication conditions. Finally, Cluster 5, which was shown to consist primarily of artefacts (e.g., shallow water, riparian vegetation, turbulent flow, reflective surfaces), was confirmed as a non-valid category and excluded from the environmental classification (Table 5).

Table 5. Classification of clusters into trophic levels based on the Trophic Score.

Clusters	Trophic Score	Category
Cluster 1	1.3578	Medium
Cluster 2	0.0023	Low
Cluster 3	1.099	Medium
Cluster 4	1.4235	High
Cluster 5	Artefact	Artefact

In this way, the evaluation process did more than simply confirm the presence of distinct classes; it also served a corrective role—particularly for Cluster 3—ensuring a more balanced and reliable environmental classification. The combined interpretation of the *TS* and the inter-cluster distance relationships allowed the clusters to be robustly characterized. Table 6 summarizes the environmental classification derived from the integrated evaluation of cluster distances and the *TS*, consistent with the previous classification scheme.

Table 6. The final characterization of the clusters resulted from the combination of the numerical Trophic Score values and the distance analysis between the centroids.

Cluster	Environmental Classification Based on Distance	Trophic Score	Final Category
Cluster 1	3	1.357797593	Medium-High
Cluster 2	1	0.002306009	Low
Cluster 3	2	1.0999000216	Medium
Cluster 4	4	1.423457742	High
Cluster 5	Artefact	Artefact	Artefact

The percentage distribution of the clusters produced by the k-means analysis is presented in Figure 12. The largest proportion of the stream area is represented by Cluster 1 (46.85%) and Cluster 4 (35.40%), corresponding to medium and high eutrophication levels, respectively. In contrast, Clusters 2 and 3, which reflect lower-intensity conditions, occupy substantially smaller areas accounting for 5.26% and 5.22% of the stream surface. Finally, Cluster 5 (7.27%), classified as an artefact, corresponds to pixels that were not correctly identified as part of the water body (Figure 12). Figure 13 presents the spatial distribution of eutrophication classes obtained from k-means clustering. Water surfaces are assigned to four trophic levels, and Cluster 5 is excluded as an artefact class.

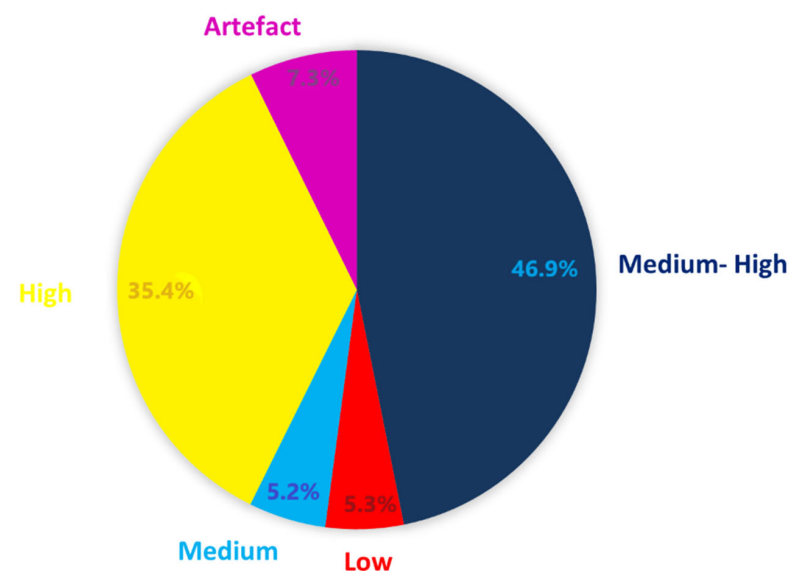


Figure 12. Spatial distribution of the clusters along the stream, as produced by the k-means clustering analysis.

To characterize the clusters based on spectral indices, the mean and standard deviation (SD) values of the four spectral indices (MCI, NDAI, NDCI and CI) were calculated for the four identified clusters (Table 7). Cluster 1 exhibits relatively high mean values across all indices (MCI = 1.03, NDAI = 1.23, NDCI = 1.17, CI = 1.99) with consistently low SDs, indicating homogeneous spectral behavior and the presence of algal biomass. Cluster 2

presents extremely low mean values across all indices (near zero) with the smallest variability, indicating the absence of chlorophyll. Cluster 3 demonstrates intermediate index values (e.g., MCI = 1.17) but relatively lower NDAI (0.60) and NDCI (0.89), accompanied by slightly higher SDs across all indices. This heterogeneous spectral composition may represent transitional or mixed zones such as shallow or turbid waters, partially vegetated surfaces, or regions with variable biomass density. Finally, cluster 4 is characterized by the highest mean values of NDAI (1.33), NDCI (1.30), and CI (2.02) and the low SDs. These suggest that cluster 4 corresponds to highly productive and spectrally stable regions. The consistency across all indices implies the presence of algal blooms, typically corresponding to areas of elevated trophic status and biological activity.

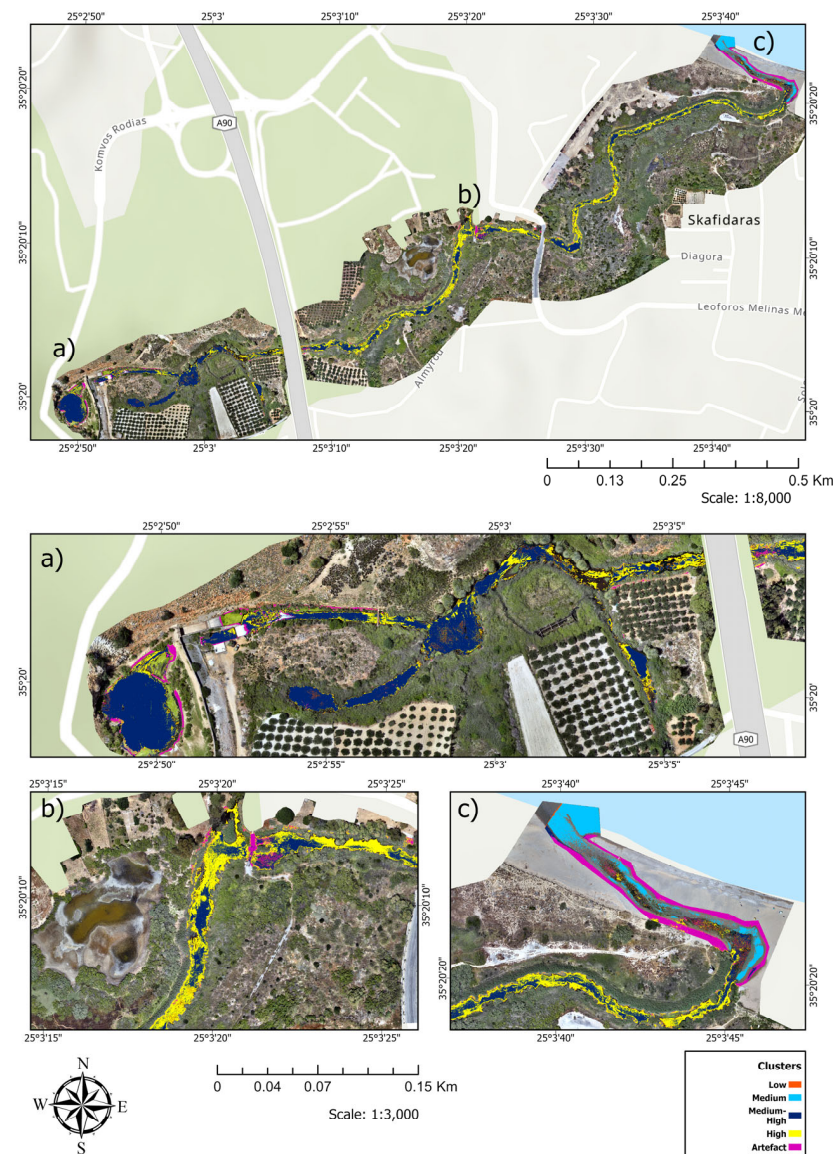


Figure 13. Final eutrophication classification map derived from the k-means clustering results. The main map presents the overall spatial distribution of the four trophic levels (Low, Medium, Medium–High, High) along the Almyros Stream, while Cluster 5 is displayed as an Artefact. Subplots (a–c) show enlarged views of selected sections of the stream, highlighted on the main map: (a) the upstream section near the dam, (b) the middle reach of the stream, and (c) the zone near the estuary where the stream meets the sea. These zoomed-in views were included to facilitate a clearer visualization of the spatial distribution and local transitions of clusters within the stream.

Table 7. Mean values and standard deviations of each spectral index for the four clusters.

Cluster	MCI		NDAI		NDCI		CI	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Cluster 1	1.0345	0.0222	1.2297	0.0780	1.1704	0.0599	1.9973	0.0141
Cluster 2	0.0000	0.0012	0.0059	0.0646	0.0000	0.0006	0.0000	0.0002
Cluster 3	1.1700	0.0620	0.6034	0.1203	0.8949	0.0967	1.7277	0.1166
Cluster 4	1.0427	0.0263	1.3348	0.0592	1.2968	0.0709	2.0205	0.0261

4. Discussion

Compared to traditional point-based measurements or single-index approaches, the multidimensional dataset employed here captures eutrophication as a spatially continuous and multivariate process, allowing subtle but environmentally meaningful patterns to emerge.

The clustering analysis for $k = 5$ revealed five distinct spectral patterns, each corresponding to different eutrophication levels and water-quality conditions in the Almyros Stream (Figure 1). The clusters displayed a coherent spatial organization, confirming the strong relationship between spectral variability and eutrophication. Cluster 2 was located primarily in the estuarine zone, where spectral indicator values were consistently low (Figure 6). This reflects areas with clearer water, reduced phytoplankton presence, and limited nutrient concentrations (Figures 12 and 13). Cluster 3 exhibited moderate indicator values and represents a transitional zone in which intermediate spectral changes occur, likely driven by localized biological activity or inputs from nearby agricultural land (Figure 6). These areas likely correspond to the mixing of relatively clear water with slightly enriched or mildly affected water masses (Figures 12 and 13). Clusters 1 and 4 presented elevated chlorophyll-related values, corresponding to medium and high eutrophication levels, respectively (Figure 6). Spatially, these clusters were associated with areas influenced by agricultural runoff, reinforcing the link between eutrophication intensity and anthropogenic activities (Figures 12 and 13). The spatial distribution of these clusters provides essential insights into the hydrological and ecological functioning of the stream. In the eastern sector, near the barrier where water depth is greater, Cluster 1 predominates, indicating medium eutrophication levels (Figures 12 and 13). In the central part of the stream, Cluster 4 becomes more prevalent, highlighting high trophic values. This pattern is likely related to shallower, more stagnant conditions—factors known to promote phytoplankton growth (Figures 12 and 13).

Along the stream banks, Cluster 4 appears frequently, whereas the central channel is mainly characterized by Cluster 1, suggesting the coexistence of different environmental regimes within the stream. These variations may be explained by local hydrodynamic differences and the influence of external nutrient sources. The cluster distribution confirms that areas with restricted flow are more vulnerable to eutrophication. In contrast, the estuary shows a sharp shift in pattern, dominated by Clusters 2 and 3, which correspond to low or moderate eutrophication levels. This is consistent with the mixing of freshwater with seawater and with salinization processes that typically inhibit chlorophyll accumulation. Overall, the predominance of Clusters 1 and 4 suggests that the stream exhibits medium to high levels of eutrophication. The widespread presence of medium values and the significant extent of high values indicate that, despite localized variations, the system maintains generally elevated trophic conditions. The limited spatial coverage of low-eutrophication areas (Clusters 2 and 3) is insufficient to influence the overall environmental characterization of the stream.

Beyond the qualitative interpretation of the spatial distribution of clusters, a quantitative synthesis was required to support their environmental classification. Within this framework, the TS proved to be a key interpretative tool, bridging the gap between unsupervised spectral clustering and environmental assessment. By integrating multiple chlorophyll-sensitive indices into a single composite metric, the TS reduced the influence of individual index variability and enabled a robust ranking of clusters along a relative eutrophication gradient. This facilitated objective comparison between spatially distinct water areas and strengthened the environmental meaning of the clustering results. The agreement between TS-based classification, cluster-distance relationships, and known hydrological and anthropogenic controls [15,16,18] further supports the validity of TS as an effective proxy for relative trophic intensity in UAV-based stream and river monitoring.

The results are consistent with previous studies demonstrating that methodologies using multispectral imagery from unmanned aerial vehicles (UAVs) are effective for high-resolution mapping of similar ecosystems [52], particularly for generating spatial distribution maps of chlorophyll-a (Chl-a) [53]. UAV-based approaches address the limitations of traditional in situ sampling, which often cannot capture complete spatial distributions, conduct large-area surveys within a single tidal or hydrological phase, or adequately represent long-term trends [54]. Other studies have shown that the application of the k-means clustering algorithm to Chl-a time-series data is effective for detecting potential impacts of external drivers on chlorophyll dynamics [55].

Beyond aquatic ecosystems, k-means clustering applied to vegetation indices has also been widely and successfully used in agricultural research. For example, Shi et al. (2024) [56], combined UAV-derived NDVI and other vegetation indices with k-means clustering to delineate management zones in crop fields, enabling targeted fertilization and variable management practices. This approach has been validated for corn and soybean fields, where clustering based on NDVI and terrain attributes facilitated effective zoning and management. Moreover, Ferro et al. (2023) [57] applied k-means clustering to multispectral UAV images and vegetation indices (NDVI, NDRE, GNDVI, MSAVI) to identify zones of low, medium, and high vegetative vigor, supporting agromomic decision-making and yield prediction. Finally, Krklješ et al. (2025) [58], defined compact zones in blueberry orchards by applying k-means clustering to NDVI derived from UAV multispectral orthomosaics, thereby supporting optimized soil sampling and orchard management.

Despite its effectiveness, the methodology also presents limitations. The presence of artefacts—represented by Cluster 5—may affect classification, as these areas cannot be fully characterized by spectral indicators alone. In the present study, indicators related to chlorophyll-a and eutrophication were not measured in situ but were indirectly estimated using UAV-based multispectral data and established remote sensing indices. Nevertheless, the findings are consistent with the spatial and hydrological characteristics of the study area as shown in the previously published work of Kokinou et al. (2023) [15]. Although no dedicated in-situ sampling campaign was carried out concurrently with the UAV survey, the plausibility of the k-means clustering results can be indirectly evaluated against the findings of [15], who performed a detailed spatiotemporal environmental assessment of the same Almyros karst system. In their study, monthly measurements of physicochemical parameters, nutrients and photosynthetic pigments, complemented by geophysical (spectral induced polarization) analyses and GIS mapping, revealed clear spatial gradients in water quality and identified specific pressure hotspots associated with agricultural activity, industrial infrastructure (power plant, desalination plant) and mixed land uses within the wetland and along the stream. When the spatial distribution of trophic clusters identified in the present work is compared qualitatively with the zones of degraded water quality

and elevated pigment or nutrient levels reported by [15], a strong correspondence emerges: areas classified here as “Medium–High” and “High” eutrophication tend to coincide with sectors that their work characterizes as environmentally stressed, whereas the “Low” and “Medium” clusters dominate in reaches where their measurements indicate comparatively better water status or stronger dilution effects. This agreement between independent, chemically and geophysical-based assessments and the unsupervised spectral partitioning applied here provides an important indirect validation of the clustering scheme and supports its interpretation as a meaningful representation of relative eutrophication levels along the Almyros Stream.

In conclusion, the gradation of clusters illustrates a spatial escalation of eutrophication from upstream to downstream, following the distribution of pollution sources and prevailing hydrological conditions. The internal cohesion and separation of the clusters underscore the stability of the k-means algorithm and its capacity to reveal essential environmental patterns.

Future studies could:

1. Incorporate concurrent field sampling of chlorophyll-a, nutrients, and physicochemical parameters during UAV campaigns. Such data would enable direct quantitative validation of remotely sensed eutrophication classes and support the calibration of spectral indicators under varying hydrological conditions.
2. Extend the methodology to repeated UAV surveys across different seasons and hydrological states would allow investigation of temporal dynamics in eutrophication patterns. This would enhance understanding of seasonal drivers, episodic nutrient inputs, and the persistence or variability of identified trophic hotspots.
3. Compare the performance of k-means with other unsupervised, semi-supervised, or hybrid machine-learning methods to better capture complex spectral–environmental relationships.
4. Apply the proposed framework to rivers and streams with different geomorphological, climatic, and optical properties. This would help assess its generalizability. Comparative studies across multiple sites could identify system-specific adaptations and support the development of standardized UAV-based eutrophication monitoring protocols.

5. Conclusions

This study presents a high-resolution, UAV-based, unsupervised framework for mapping eutrophication in stream systems, integrating multispectral imagery and chlorophyll-sensitive indices without reliance on extensive in situ calibration. The novelty lies in the spatially explicit identification of eutrophication gradients using k-means clustering, combined with an objective cluster-evaluation strategy that supports both environmental interpretation and optimized sampling design. Specifically:

- A multidimensional remote sensing framework based on four chlorophyll-related indices (NDCI, NDAI, CI and MCI) computed from UAV-based multispectral imagery was developed to assess stream eutrophication.
- Applying the k-means clustering algorithm to the four indices, the Almyros Stream was divided into five distinct spectral clusters that captured the system’s environmental variability. Based on the optimal number of clusters ($k = 5$), four clusters corresponded to meaningful trophic conditions ranging from low to high eutrophication, while the fifth cluster represented artefacts associated with shallow water, turbulence, riparian vegetation, and adjacency effects.
- Medium to high eutrophication levels dominate much of the stream, especially in low-flow, shallow, and agriculturally influenced areas, whereas low levels prevail near the estuary where seawater mixing occurs.

- The spatial distribution of the clusters and their consistency with known hydrological and anthropogenic drivers confirm the ability of the proposed approach to represent meaningful environmental conditions. Furthermore, the good agreement with previously published chemically and geophysically based assessments provides indirect validation of the approach.

Finally, the spatially explicit eutrophication maps enable evidence-based management and optimized sampling strategies, allowing monitoring efforts to focus on representative and high-risk areas, improving early detection of degradation, and supporting more effective protection and restoration of vulnerable aquatic ecosystems.

Author Contributions: Conceptualization, M.K. and C.V.; methodology, M.K., C.V., E.K. and N.G.; software, M.K., C.V., E.K. and N.G.; validation, M.K. and C.V.; formal analysis, M.K., C.V. and E.K.; investigation, M.K., C.V., E.K. and N.G.; resources, M.K., C.V. and E.K.; data curation, M.K. and C.V.; writing—original draft preparation, M.K., C.V., E.K. and N.G.; writing—review and editing, M.K., C.V., E.K. and N.G.; visualization, M.K. and C.V.; supervision, C.V. and E.K.; project administration, C.V. and E.K.; funding acquisition, C.V. and E.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study is available upon request from the corresponding author due to privacy reasons.

Acknowledgments: The authors are grateful to the editor, assistant editor and anonymous reviewers for their critical review and constructive comments.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

Clustering Results for $k = 2$ to 11 (excluding $k = 5$). These illustrations allow the visual assessment of the differentiation between clusters and the cohesion of the produced cluster groups, supporting the conclusion that $k = 5$ is the most appropriate option.

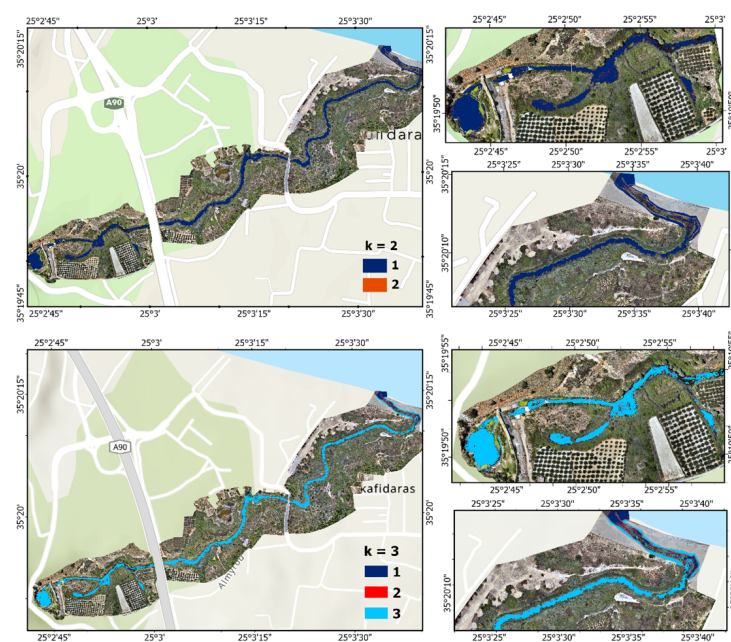


Figure A1. Cont.

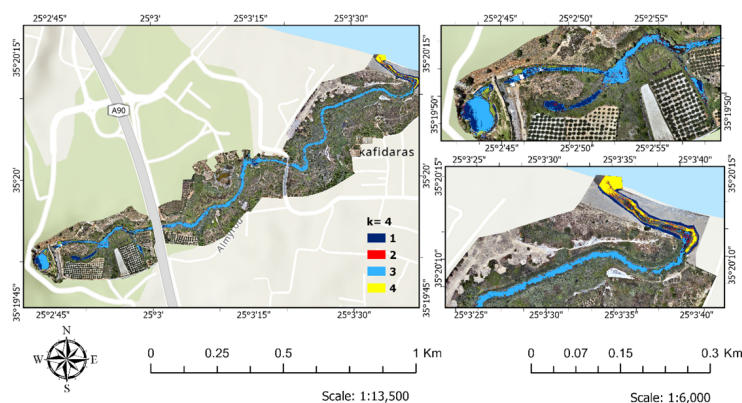


Figure A1. Cartographic representation of the k-means clustering results for $k = 2, 3$, and 4 . The maps shown in the second column present enlarged views of selected areas from the main clustering outputs for $k = 2, 3$, and 4 . These zoomed-in sections were included to provide a clearer visualization of how the clusters are distributed within the stream and to highlight small-scale spatial variations that are not easily visible in the full-resolution maps. This detailed perspective helps illustrate the degree of generalization or fragmentation produced by each k value and supports the evaluation of clustering performance.



Figure A2. Cont.

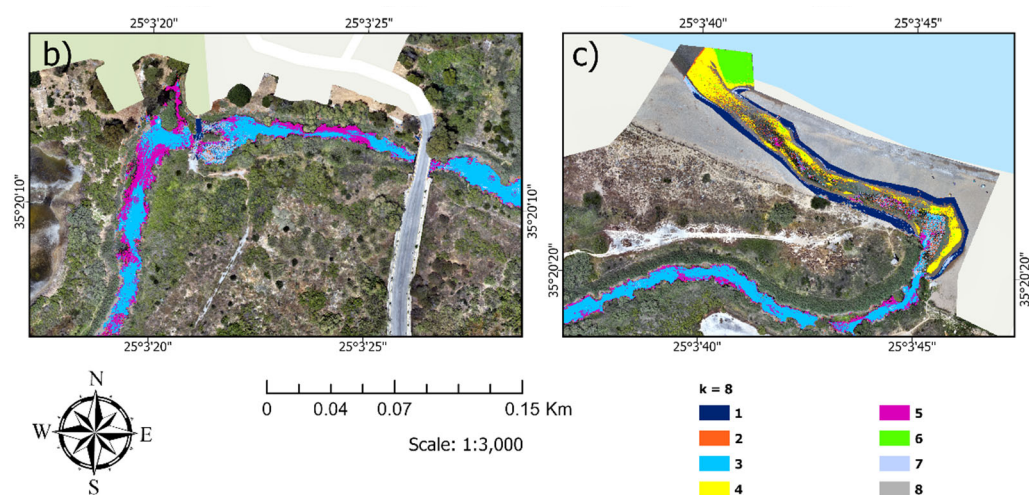


Figure A2. Cartographic representation of the k-means clustering results for $k=8$: (a) Zoomed-in view of the upper part of the stream and the dam, (b) zoomed-in view of the middle section of the stream, (c) zoomed-in view of the lower part of the stream near the estuary.



Figure A3. Cont.

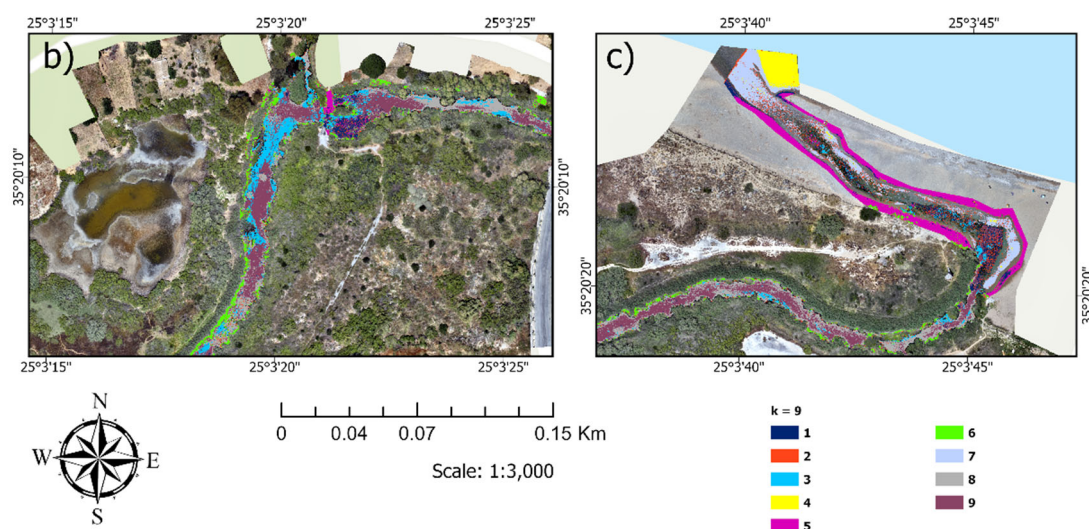


Figure A3. Cartographic representation of the k-means clustering results for $k = 9$: (a) Zoomed-in view of the upper part of the stream and the dam, (b) zoomed-in view of the middle section of the stream, (c) zoomed-in view of the lower part of the stream near the estuary.



Figure A4. Cont.

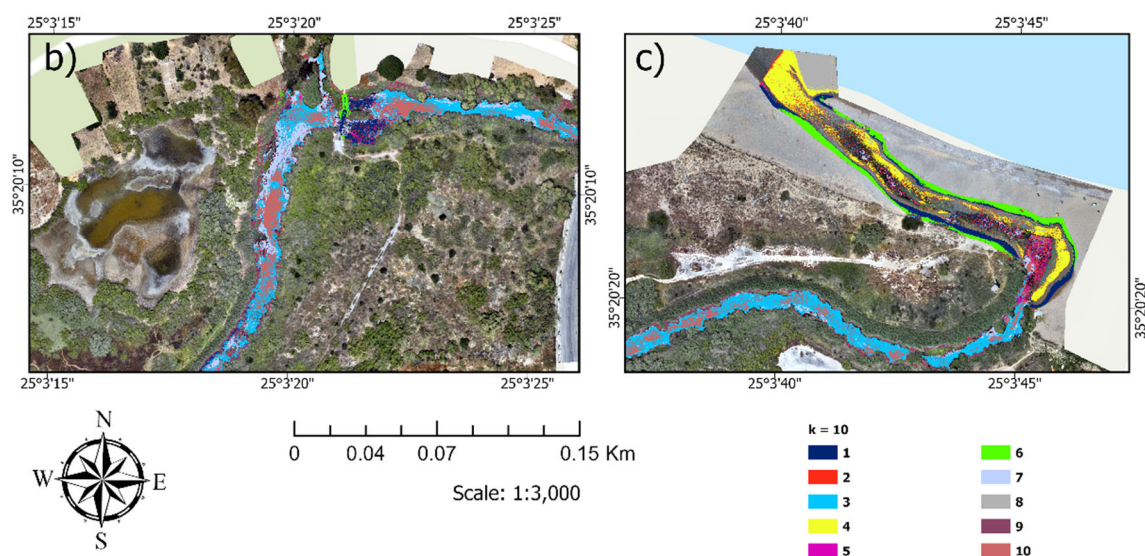


Figure A4. Cartographic representation of the k-means clustering results for $k = 10$: (a) Zoomed-in view of the upper part of the stream and the dam, (b) zoomed-in view of the middle section of the stream, (c) zoomed-in view of the lower part of the stream near the estuary.



Figure A5. Cont.

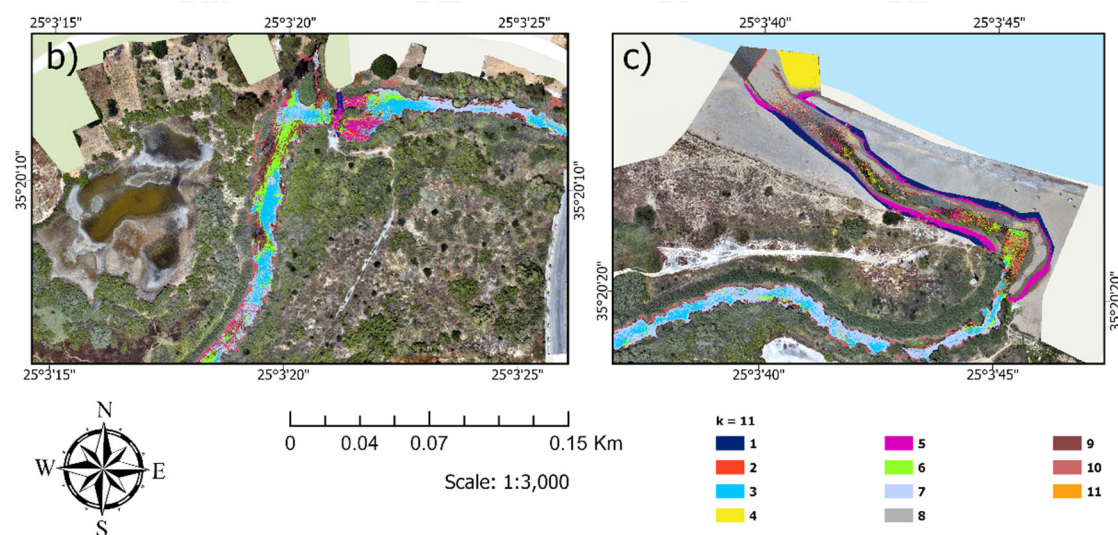


Figure A5. Cartographic representation of the k-means clustering results for $k = 11$: (a) Zoomed-in view of the upper part of the stream and the dam, (b) zoomed-in view of the middle section of the stream, (c) zoomed-in view of the lower part of the stream near the estuary.



Figure A6. Cont.

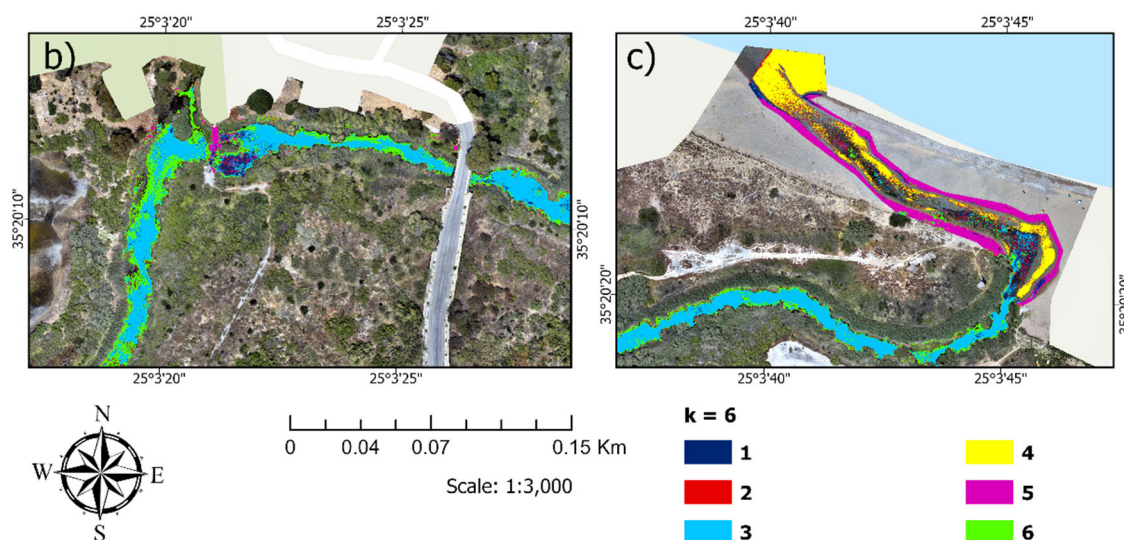


Figure A6. Cartographic representation of the k-means clustering results for $k = 6$: (a) Zoomed-in view of the upper part of the stream and the dam, (b) zoomed-in view of the middle section of the stream, (c) zoomed-in view of the lower part of the stream near the estuary.

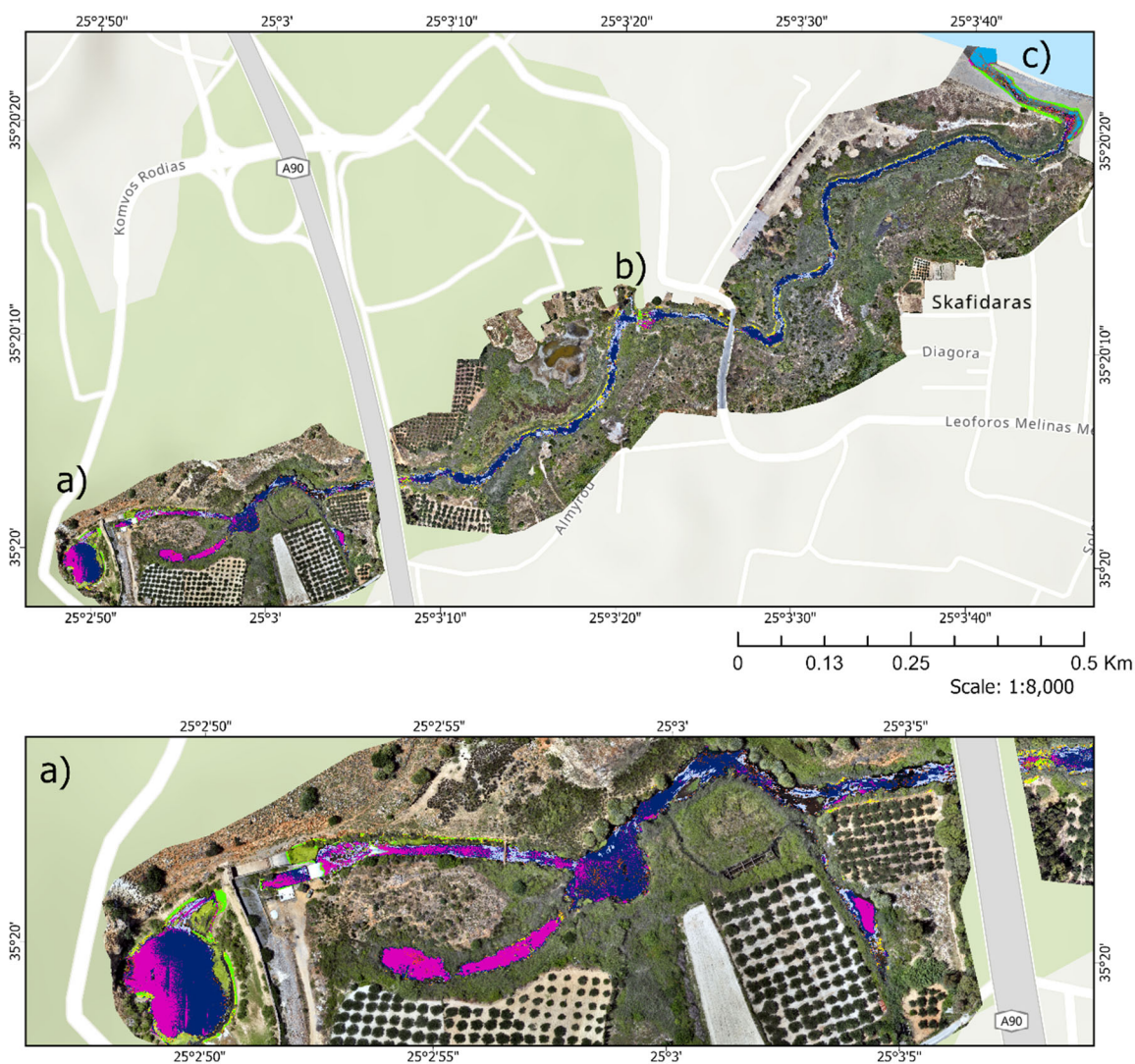


Figure A7. Cont.

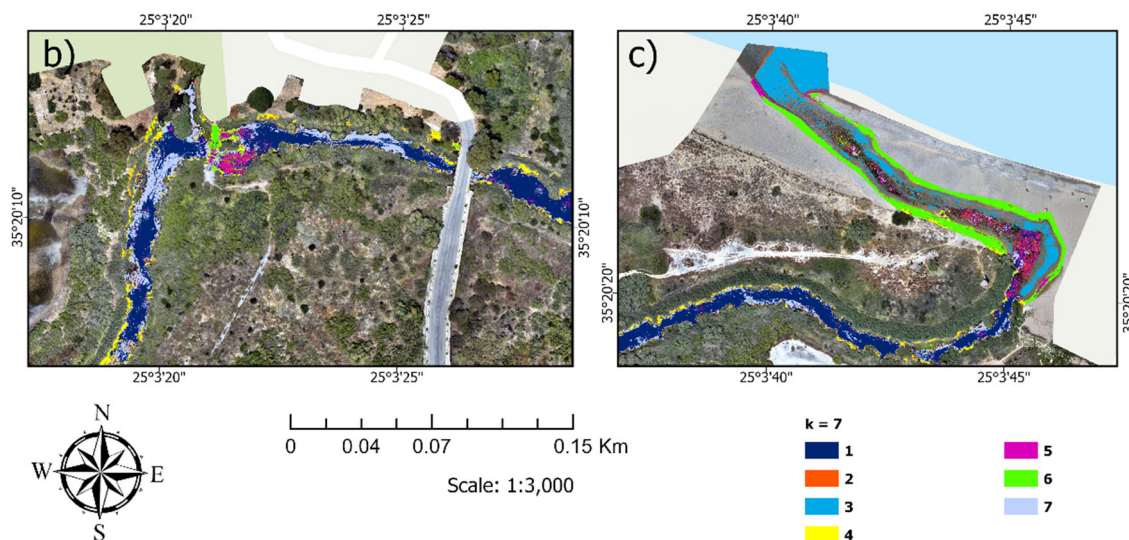


Figure A7. Cartographic representation of the k-means clustering results for $k = 7$: (a) Zoomed-in view of the upper part of the stream and the dam, (b) zoomed-in view of the middle section of the stream, (c) zoomed-in view of the lower part of the stream near the estuary.

References

1. Xiao, Q.; Xu, X.; Duan, H.; Qi, T.; Qin, B.; Lee, X.; Hu, Z.; Wang, W.; Xiao, W.; Zhang, M. Eutrophic Lake Taihu as a significant CO_2 source during 2000–2015. *Water Res.* **2020**, *170*, 115331. [CrossRef] [PubMed]
2. Smith, V.H.; Tilman, G.D.; Nekola, J.C. Eutrophication: Impacts of excess nutrient inputs on freshwater, marine, and terrestrial ecosystems. *Environ. Pollut.* **1999**, *100*, 179–196. [CrossRef] [PubMed]
3. Feng, L.; Wang, Y.; Hou, X.; Qin, B.; Kutser, T.; Qu, F.; Chen, N.; Paerl, H.W.; Zheng, C. Harmful algal blooms in inland waters. *Nat. Rev. Earth Environ.* **2024**, *5*, 631–644. [CrossRef] [PubMed]
4. Xiao, Q.; Xu, X.; Qi, T.; Luo, J.; Lee, X.; Duan, H. Lakes shifted from a carbon dioxide source to a sink over past two decades in China. *Sci. Bull.* **2024**, *69*, 1857–1861. [CrossRef]
5. Balmer, M.B.; Downing, J.A. Carbon dioxide concentrations in eutrophic lakes: Undersaturation implies atmospheric uptake. *Inland Waters* **2011**, *1*, 125–132. [CrossRef]
6. Kokolakis, S.; Kokinou, E.; Chronaki, C.; Moen, A.; Datta, G. Earth Observation and Geoinformatics to Monitoring the Environmental Status of Urban Streams Inextricably Linked to People's Mental Health. In *Digital Health and Informatics Innovations for Sustainable Health Care Systems*; IOS Press: Amsterdam, The Netherlands, 2024; pp. 1560–1564. Available online: <https://ebooks.iospress.nl/doi/10.3233/SHTI240716> (accessed on 12 March 2025).
7. Dodds, W.K.; Bouska, W.W.; Eitzmann, J.L.; Pilger, T.J.; Pitts, K.L.; Riley, A.J.; Schloesser, J.T.; Thornbrugh, D.J. Eutrophication of U.S. Freshwaters: Analysis of Potential Economic Damages. *Environ. Sci. Technol.* **2009**, *43*, 12–19. [CrossRef]
8. Paerl, H.W.; Paul, V.J. Climate change: Links to global expansion of harmful cyanobacteria. *Water Res.* **2012**, *46*, 1349–1363. [CrossRef]
9. Jain, A.K. Data clustering: 50 years beyond K-means. *Pattern Recognit. Lett.* **2010**, *31*, 651–666. [CrossRef]
10. Li, W.; Ren, M.; Zhang, H.; Duan, Y.; Chen, D.; Li, S.; Xu, M.; Wang, L.; Yang, X. Estimation of Water Quality in Coastal Aquaculture Waters Using the Combination of Machine Learning and Unmanned Aerial Vehicle Multispectral Imagery—ScienceDirect. *Aquaculture* **2026**, *611*, 743002. [CrossRef]
11. Ikotun, A.M.; Ezugwu, A.E.; Abualigah, L.; Abuhaija, B.; Heming, J. K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. *Inf. Sci.* **2023**, *622*, 178–210. [CrossRef]
12. Phiri, D.; Morgenroth, J.; Phiri, D.; Morgenroth, J. Developments in Landsat Land Cover Classification Methods: A Review. *Remote Sens.* **2017**, *9*, 967. [CrossRef]
13. Islam, M.M.; Rahman, M.S.; Kabir, M.A.; Islam, M.N.; Chowdhury, R.M. Predictive assessment on landscape and coastal erosion of Bangladesh using geospatial techniques. *Remote Sens. Appl. Soc. Environ.* **2020**, *17*, 100277. [CrossRef]
14. Gao, S.; Janowicz, K.; Couclelis, H. Extracting urban functional regions from points of interest and human activities on location-based social networks. *Trans. GIS* **2017**, *21*, 446–467. [CrossRef]
15. Kokinou, E.; Zacharioudaki, D.E.; Kokolakis, S.; Kotti, M.; Chatzidavid, D.; Karagiannidou, M.; Fanouraki, E.; Kontaxakis, E. Spatiotemporal environmental monitoring of the karst-related Almyros Wetland (Heraklion, Crete, Greece, Eastern Mediterranean). *Environ. Monit. Assess.* **2023**, *195*, 955. [CrossRef]

16. Zacharioudaki, D.-E.; Kotti, M.; Kokinou, E. Evaluation of water salinity through fluorescence: The case of Almiros River (Northeastern Crete, Greece). *Int. J. Environ. Anal. Chem.* **2021**, *101*, 2525–2538. [\[CrossRef\]](#)
17. Source Of Almyros | Unesco Sites in Crete. Available online: <https://www.unescositesincrete.gr/en/simeia-endiaferontos/source-of-almyros/> (accessed on 24 October 2025).
18. Kokolakis, S.; Kokinou, E.; Karagiannidou, M.; Gerarchakis, N.; Vasilakos, C.; Kotti, M.; Chronaki, C. From Space to Stream: Combining Remote Sensing and In Situ Techniques for Comprehensive Stream Health Assessment. *Remote Sens.* **2025**, *17*, 1532. [\[CrossRef\]](#)
19. DJI. DJI Mavic 3M—Specifications. Available online: <https://ag.dji.com/mavic-3-m?site=ag&from=nav> (accessed on 3 June 2025).
20. Konik, M.; Bradtke, K.; Stoń-Egiert, J.; Soja-Woźniak, M.; Śliwińska-Wilczewska, S.; Darecki, M. Cyanobacteria Index as a Tool for the Satellite Detection of Cyanobacteria Blooms in the Baltic Sea. *Remote Sens.* **2023**, *15*, 1601. [\[CrossRef\]](#)
21. Shi, W.; Wang, M. Green macroalgae blooms in the Yellow Sea during the spring and summer of 2008. *J. Geophys. Res. Ocean.* **2009**, *114*. [\[CrossRef\]](#)
22. Maximum Chlorophyll Index (MCI)—Euro Data Cube Public Collections. Available online: <https://collections.eurodatacube.com/xcube-gen-s2-mci/> (accessed on 5 February 2025).
23. Choudhary, A. Empirical Comparison of Performances of K-Means, K-Means++, Weighted K-Means and Hartigan and Wong K-Means Clustering Algorithms. Available online: https://www.researchgate.net/publication/283300196_Empirical_Comparison_of_Performances_of_K-Means_K-Means_Weighted_K-Means_and_Hartigan_and_Wong_K-Means_Clustering_Algorithms (accessed on 1 February 2026).
24. Normalized Difference Chlorophyll Index (NDCI)—ClimateEngine.org Support. Available online: <https://support.climateengine.org/article/127-normalized-difference-chlorophyll-index> (accessed on 4 February 2025).
25. Xu, R.; Wunsch, D. Survey of clustering algorithms. *IEEE Trans. Neural Netw.* **2005**, *16*, 645–678. [\[CrossRef\]](#)
26. Pawełek-Lubera, E.; Przyborowski, M.; Ślęzak, D.; Wasilewski, A. Multi-criteria selection of data clustering methods for e-commerce personalization. *Appl. Soft Comput.* **2025**, *182*, 113559. [\[CrossRef\]](#)
27. Tang, M.; Zhou, X.; Liao, H.; Xu, J.; Fujita, H.; Herrera, F. Ordinal consensus measure with objective threshold for heterogeneous large-scale group decision making. *Knowl.-Based Syst.* **2019**, *180*, 62–74. [\[CrossRef\]](#)
28. Dadjoo, M.; Fatemi Nasrabadi, S.B. The application of spatial domain in optimum initialization for clustering image data using particle swarm optimization. *Expert Syst. Appl.* **2021**, *168*, 114224. [\[CrossRef\]](#)
29. Ahmed, M.; Seraj, R.; Islam, S.M.S. The k-means Algorithm: A Comprehensive Survey and Performance Evaluation. *Electronics* **2020**, *9*, 1295. [\[CrossRef\]](#)
30. Wang, Y.; Li, D.; Wang, Y. Realization of remote sensing image segmentation based on K-means clustering. *IOP Conf. Ser. Mater. Sci. Eng.* **2019**, *490*, 072008. [\[CrossRef\]](#)
31. Halkidi, M.; Batistakis, Y.; Vazirgiannis, M. On Clustering Validation Techniques. *J. Intell. Inf. Syst.* **2001**, *17*, 107–145. [\[CrossRef\]](#)
32. Rousseeuw, P.J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **1987**, *20*, 53–65. [\[CrossRef\]](#)
33. Davies, D.L.; Bouldin, D.W. A Cluster Separation Measure. *IEEE Trans. Pattern Anal. Mach. Intell.* **1979**, *PAMI-1*, 224–227. [\[CrossRef\]](#)
34. Gholizadeh, M.H.; Melesse, A.M.; Reddi, L. A Comprehensive Review on Water Quality Parameters Estimation Using Remote Sensing Techniques. *Sensors* **2016**, *16*, 1298. [\[CrossRef\]](#)
35. Arandhara, B.; Shukla, J.; Dhyani, S. Damage assessment of Baghjan oil field blowout on terrestrial and aquatic ecosystems near Dibru Saikhowa Biosphere Reserve, Assam India. *Remote Sens. Appl. Soc. Environ.* **2023**, *31*, 100999. [\[CrossRef\]](#)
36. SIOS's Earth Observation (EO), Remote Sensing (RS), and Operational Activities in Response to COVID-19. *Remote Sens.* **2021**, *13*, 712. [\[CrossRef\]](#)
37. He, Y.; Leng, L.; Ji, X.; Wang, M.; Huo, Y.; Li, Z.; He, Y.; Leng, L.; Ji, X.; Wang, M.; et al. Inversion and Analysis of Global Ocean Chlorophyll-a Concentration Based on Temperature Zoning. *Remote Sens.* **2024**, *16*, 2302. [\[CrossRef\]](#)
38. Yunus, A.P.; Dou, J.; Sravanthi, N. Remote sensing of chlorophyll-a as a measure of red tide in Tokyo Bay using hotspot analysis. *Remote Sens. Appl. Soc. Environ.* **2015**, *2*, 11–25. [\[CrossRef\]](#)
39. MacQueen, J. Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics*; University of California Press: Oakland, CA, USA, 1967; pp. 281–298. Available online: <https://projecteuclid.org/ebooks/berkeley-symposium-on-mathematical-statistics-and-probability/Proceedings-of-the-Fifth-Berkeley-Symposium-on-Mathematical-Statistics-and-/chapter/Some-methods-for-classification-and-analysis-of-multivariate-observations/bsmsp/1200512992> (accessed on 18 September 2025).
40. Bogdał, A.; Wałęga, A.; Kowalik, T.; Cupak, A. Assessment of the Impact of Forestry and Settlement-Forest Use of the Catchments on the Parameters of Surface Water Quality: Case Studies for Chechło Reservoir Catchment, Southern Poland. *Water* **2019**, *11*, 964. [\[CrossRef\]](#)

41. Sasanka, A.K.R.; Kariyawasam, D.H.; Siripala, K.R.R. *Cluster Analysis of Microbial, Physico-Chemical Characteristics, Water Quality Index (WQI) and Trophic Level Index (TLI) for Water Quality Assessment of Beira Lake, Colombo, Sri Lanka*; Sri Lanka Technology Campus: Padukka, Sri Lanka, 2024.
42. Pita, A.; Rodriguez, F.J.; Navarro, J.M. Analysis and Evaluation of Clustering Techniques Applied to Wireless Acoustics Sensor Network Data. *Appl. Sci.* **2022**, *12*, 8550. [\[CrossRef\]](#)
43. Sterckx, S.; Knaeps, E.; Ruddick, K.; Sterckx, S.; Knaeps, E.; Ruddick, K. Detection and Correction of Adjacency Effects in hyperspectral airborne data of Coastal and Inland Waters: The Use of the Near Infrared Similarity Spectrum. *Int. J. Remote Sens.* **2011**, *32*, 6479–6505. [\[CrossRef\]](#)
44. Peppas, M.; Vasilakos, C.; Kavrouidakis, D.; Peppas, M.; Vasilakos, C.; Kavrouidakis, D. Eutrophication Monitoring for Lake Pamvotis, Greece, Using Sentinel-2 Data. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 143. [\[CrossRef\]](#)
45. Özdarici, O.A.; Türker, M. Field-Based Classification of Agricultural Crops Using Multi-Scale Images. 2012. Available online: https://www.isprs.org/proceedings/XXXVI/4-C42/Papers/18_Automated%20classification%20IC%20II%20-%20Agriculture/OBIA2006_Ozdarici_Turker.pdf (accessed on 1 February 2025).
46. Legleiter, C. Mapping river depth from publicly available aerial images. *River Res. Appl.* **2013**, *29*, 760–780. [\[CrossRef\]](#)
47. Jensen, J.R. *Remote Sensing of the Environment: An Earth Resource Perspective*; Prentice Hall: Hoboken, NJ, USA, 2006.
48. Chen, N.; Liu, L.; Qiao, D.; Li, Y.; Lv, Y. Seasonal succession patterns of plankton in eutrophic rivers on plains. *Ann. Limnol.-Int. J. Lim.* **2016**, *52*, 217–233. [\[CrossRef\]](#)
49. Moorhouse, H.L.; Read, D.S.; McGowan, S.; Wagner, M.; Roberts, C.; Armstrong, L.K.; Nicholls, D.J.E.; Wickham, H.D.; Hutchins, M.G.; Bowes, M.J. Characterisation of a major phytoplankton bloom in the River Thames (UK) using flow cytometry and high performance liquid chromatography. *Sci. Total Environ.* **2018**, *624*, 366–376. [\[CrossRef\]](#)
50. Akinawo, S.O. Eutrophication: Causes, consequences, physical, chemical and biological techniques for mitigation strategies. *Environ. Chall.* **2023**, *12*, 100733. [\[CrossRef\]](#)
51. Wu, C.; Liu, L.; Chen, C.; Zhang, C.; He, G.; Li, J.; Wu, C.; Liu, L.; Chen, C.; Zhang, C.; et al. Challenges of the Polarimetric Update on Operational Radars in China—Ground Clutter Contamination of Weather Radar Observations. *Remote Sens.* **2021**, *13*, 217. [\[CrossRef\]](#)
52. Román, M.; Davies, B.F.R.; Oiry, S.; Rosa, P.; Gernez, P.; Olabarria, C.; Barillé, L. Discrimination of the intertidal goose barnacle *Pollicipes pollicipes* from rocky shore invertebrates and macroalgae using in situ hyperspectral signatures. *Remote Sens. Appl. Soc. Environ.* **2025**, *39*, 101697. [\[CrossRef\]](#)
53. Chen, B.; Mu, X.; Chen, P.; Wang, B.; Choi, J.; Park, H.; Xu, S.; Wu, Y.; Yang, H. Machine learning-based inversion of water quality parameters in typical reach of the urban river by UAV multispectral data. *Ecol. Indic.* **2021**, *133*, 108434. [\[CrossRef\]](#)
54. Yuan, S.; Li, Y.; Bao, F.; Xu, H.; Yang, Y.; Yan, Q.; Zhong, S.; Yin, H.; Xu, J.; Huang, Z.; et al. Marine environmental monitoring with unmanned vehicle platforms: Present applications and future prospects. *Sci. Total Environ.* **2023**, *858*, 159741. [\[CrossRef\]](#)
55. Mercado, J.M.; Gómez-Jakobsen, F.; Cortés, D.; Yebra, L.; Salles, S.; León, P.; Putzeys, S. A method based on satellite imagery to identify spatial units for eutrophication management. *Remote Sens. Environ.* **2016**, *186*, 123–134. [\[CrossRef\]](#)
56. Shi, W.; Li, Y.; Zhang, W.; Yu, C.; Zhao, C.; Qiu, J. Monitoring and zoning soybean maturity using UAV remote sensing. *Ind. Crops Prod.* **2024**, *222*, 119470. [\[CrossRef\]](#)
57. Ferro, M.V.; Catania, P.; Micciché, D.; Pisciotto, A.; Vallone, M.; Orlando, S. Assessment of vineyard vigour and yield spatio-temporal variability based on UAV high resolution multispectral images. *Biosyst. Eng.* **2023**, *231*, 36–56. [\[CrossRef\]](#)
58. Krklješ, D.; Kitić, G.; Panić, M.; Petes, C.; Filipović, V.; Stefanović, D.; Obrenović, N.; Lalić, M.; Marko, O. Agrobot Gari, a multimodal robotic solution for blueberry production automation. *Comput. Electron. Agric.* **2025**, *237*, 110626. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.