

Enhanced Water Access Segmentation Using an Improved Salp Swarm Algorithm for Regional Development Planning [†]

Youness Boudrik ^{1,*}, Achraf Touil ², Rachid Hasnaoui ¹, Mustapha Ahlaqqach ³ and Mhammed El Bakkali ⁴

¹ Faculté d'Économie et de Gestion de Kenitra, Université Ibn Tofail-Kenitra, Kenitra 14000, Morocco; rachid.hasnaoui@uit.ac.ma

² LMII—Faculty of Sciences and Technology, Hassan 1st University, P.O. Box 577, Settat 26000, Morocco; ac.touil@uhp.ac.ma

³ CELOG, Ecole Supérieure des Industries du Textile et de l'Habillement, Casablanca 20250, Morocco; ahlaqqach@esith.ac.ma

⁴ Department of Textile Engineering, Ecole Supérieure des Industries du Textile et de l'Habillement, Casablanca 20250, Morocco; elbakkali@esith.ac.ma

* Correspondence: youness.boudrik@uit.ac.ma

[†] Presented at the 1st International Conference on Smart Management in Industrial and Logistics Engineering (SMILE 2025), 16–19 April 2025, Casablanca, Morocco.

Abstract

This paper presents a novel approach to water access segmentation by introducing an improved version of the Salp Swarm Algorithm (ISSA), addressing the complex challenge of household water access classification in developing regions. The proposed enhancement incorporates dynamic exploration–exploitation balancing and feature-aware mechanisms into the original SSA framework, significantly improving cluster quality and interpretability. Using a real-world dataset of 500 households from the El Hajeb region in Morocco and 12 socio-economic criteria, our method demonstrates superior clustering performance compared to conventional techniques. The ISSA achieves a 25% improvement in the silhouette coefficient (0.732 vs. 0.480) and a 22% reduction in the Davies–Bouldin index (0.421 vs. 0.645) compared to the standard SSA and other state-of-the-art metaheuristic algorithms. Five distinct water access segments are identified, enabling targeted infrastructure development strategies across different community types. The approach provides regional planners with essential insights into the spatial distribution of water access patterns and their relationship with socio-economic factors.

Keywords: water access segmentation; improved salp swarm algorithm; cluster analysis; regional development planning; unsupervised learning; metaheuristic optimization; socio-economic indicators



Academic Editors: Mustapha Hlyal, Abdessamad Ait El-Cadi, Kaoutar Kouzmi, Mariam Atwani and Kenza Izikki

Published: 20 June 2025

Citation: Boudrik, Y.; Touil, A.; Hasnaoui, R.; Ahlaqqach, M.; El Bakkali, M. Enhanced Water Access Segmentation Using an Improved Salp Swarm Algorithm for Regional Development Planning. *Eng. Proc.* **2025**, *97*, 38. <https://doi.org/10.3390/engproc2025097038>

Copyright: © 2025 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) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Access to clean water remains a critical development challenge in many regions worldwide, particularly in rural and peri-urban areas of developing countries [1]. Understanding the patterns and determinants of water access is essential for the design of effective water infrastructure investment strategies and policy interventions. Traditional approaches to water access analysis often rely on aggregated statistics or regression models that predict connectivity at the household level [2]. However, these methods frequently overlook important nuances in water access patterns that exist across different community segments.

Water access segmentation, as an analytical tool for community development planning, enables organizations to systematically categorize households and communities based on

multiple socio-economic and infrastructure characteristics, facilitating more effective resource allocation and targeted intervention strategies [3]. However, conventional clustering techniques often struggle with the high-dimensional nature of household data and the inherent complexity of finding optimal cluster configurations that meaningfully represent water access segments [4].

Recent advancements in computational intelligence, particularly in metaheuristic optimization algorithms, have shown promising results in addressing complex clustering problems [5]. Among these, the Salp Swarm Algorithm (SSA) [6] has emerged as an effective approach for solving optimization problems, demonstrating strong potential for clustering applications. However, the standard SSA faces limitations in maintaining an optimal balance between exploration and exploitation phases, potentially leading to premature convergence or suboptimal clustering solutions.

To address these limitations, this paper introduces an Improved Salp Swarm Algorithm (ISSA) specifically designed for water access segmentation. The proposed algorithm enhances the standard SSA through dynamic parameter adaptation mechanisms and feature-aware optimization to better capture the complex relationships between socio-economic factors and water access patterns.

The following are the main contributions of this research:

1. Development of a novel water access segmentation framework that integrates enhanced exploration and exploitation capabilities;
2. Introduction of feature-aware mechanisms that adapt optimization behavior based on the relative importance of different socio-economic indicators;
3. Comprehensive validation using a real-world dataset of 500 households from the El Hajeb region in Morocco with 12 socio-economic criteria;
4. Identification of five distinct water access segments with clear policy implications for targeted infrastructure development;
5. Spatial analysis of water access patterns, revealing critical geographic disparities in water infrastructure distribution.

The remainder of this paper is organized as follows: Section 2 presents the theoretical background of segmentation in water access analysis. Section 3 introduces the proposed ISSA methodology and its implementation for water access clustering. Section 4 outlines the experimental setup and results. Section 5 provides a detailed analysis of the identified segments and their characteristics. Section 6 discusses policy implications and recommendations. Finally, Section 7 concludes the paper with a summary of findings and directions for future research.

2. Segmentation in Water Access Analysis

2.1. Theoretical Background

Segmentation in water access analysis refers to the process of dividing a population of households or communities into distinct groups based on their water access characteristics and related socio-economic factors [7]. Unlike prediction-based approaches that estimate water access levels for individual households, segmentation identifies natural groupings within the population, revealing underlying patterns that may not be apparent through aggregated statistics or regression models.

In the context of water access, let $H = \{h_1, h_2, \dots, h_n\}$ represent a set of n households, where each household (h_i) is characterized by a vector of m attributes ($h_i = (a_{i1}, a_{i2}, \dots, a_{im})$). These attributes may include infrastructure indicators, socio-economic factors, geographic characteristics, and water usage patterns. The objective

is to partition H into k disjoint segments ($S = \{S_1, S_2, \dots, S_k\}$) while optimizing specific clustering criteria.

2.2. Problem Formulation

The water access segmentation problem can be formally defined as an optimization problem seeking to minimize the within-cluster sum of squares (WCSS):

$$\text{minimize } F(S) = \sum_{j=1}^k \sum_{h_i \in S_j} \|h_i - \mu_j\|^2 \quad (1)$$

where μ_j represents the centroid of segment S_j , calculated as follows:

$$\mu_j = \frac{1}{|S_j|} \sum_{h_i \in S_j} h_i, \quad (2)$$

subject to the following constraints:

$$\bigcup_{j=1}^k S_j = H \quad \text{and} \quad S_i \cap S_j = \emptyset, \forall i \neq j \quad (3)$$

This formulation aims to create segments where households within the same segment exhibit similar water access characteristics, while households in different segments show distinct patterns.

2.3. Evaluation Metrics

The quality of water access segmentation solutions is evaluated using several metrics:

2.3.1. Silhouette Coefficient

The Silhouette Coefficient (SC) [8] measures both cohesion and separation:

$$SC = \frac{1}{n} \sum_{i=1}^n \frac{b(h_i) - a(h_i)}{\max\{a(h_i), b(h_i)\}} \quad (4)$$

where $a(h_i)$ is the average distance between household h_i and all other households in its segment and $b(h_i)$ is the minimum average distance between h_i and households in other segments.

2.3.2. Davies–Bouldin Index

The Davies–Bouldin Index (DBI) [9] evaluates the ratio of within-cluster scatter to between-cluster separation:

$$DBI = \frac{1}{k} \sum_{i=1}^k \max_{j \neq i} \left\{ \frac{\sigma_i + \sigma_j}{d(\mu_i, \mu_j)} \right\} \quad (5)$$

where σ_i represents the average distance of households in segment i to their centroid and $d(\mu_i, \mu_j)$ is the distance between centroids of segments i and j .

2.3.3. Calinski–Harabasz Index

The Calinski–Harabasz Index (CHI) [10] evaluates cluster validity based on the ratio of between-cluster dispersion to within-cluster dispersion:

$$CHI = \frac{Tr(B_k)/(k-1)}{Tr(W_k)/(n-k)} \quad (6)$$

where B_k is the between-cluster scatter matrix, W_k is the within-cluster scatter matrix, Tr is the trace function, n is the number of households, and k is the number of clusters.

2.3.4. Dunn Index

The Dunn Index (DI) [11] measures the ratio of the smallest inter-cluster distance to the largest intra-cluster distance:

$$DI = \min_{1 \leq i < j \leq k} \left\{ \frac{d_{\min}(S_i, S_j)}{\max_{1 \leq m \leq k} \text{diam}(S_m)} \right\} \quad (7)$$

where $d_{\min}(S_i, S_j)$ is the minimum distance between segments S_i and S_j and $\text{diam}(S_m)$ is the diameter of segment S_m .

2.4. Socio-Economic Indicators

The socio-economic indicators considered in this study include the following [12]:

1. Infrastructure metrics:
 - Infrastructure score (a_1);
 - Comfort score (a_2).
2. Educational factors:
 - Education level (a_3);
 - Waste management practices (a_4).
3. Accessibility indicators:
 - Distance to road (a_5);
 - Household size (a_6);
 - School enrollment rate (a_7).
4. Basic amenities:
 - Toilet presence (a_8);
 - Kitchen presence (a_9);
 - Bathroom presence (a_{10});
 - Modern amenities (a_{11});
 - Profession type (a_{12}).

These indicators provide a comprehensive view of the factors that may influence water access patterns and segment formation.

3. Proposed Methodology

3.1. Standard Salp Swarm Algorithm

The Salp Swarm Algorithm (SSA) [6] is a population-based metaheuristic inspired by the swarming behavior of salps in deep oceans. In the standard SSA, the population is divided into leaders and followers. The position of the leader is updated based on the food source (best solution), while followers follow each other in a chain-like structure.

The position update in SSA is governed by the following equations:

$$x_j^1 = \begin{cases} F_j + c_1((ub_j - lb_j)c_2 + lb_j), & \text{if } c_3 \geq 0 \\ F_j - c_1((ub_j - lb_j)c_2 + lb_j), & \text{if } c_3 < 0 \end{cases} \quad (8)$$

$$x_j^i = \frac{1}{2}(x_j^i + x_j^{i-1}), \quad i \geq 2 \quad (9)$$

where

- x_j^1 is the position of the leader salp in dimension j ;

- x_j^i is the position of the i -th follower salp in dimension j ;
- F_j is the position of the food source in dimension j ;
- ub_j and lb_j are the upper and lower bounds of dimension j ;
- c_1 is the main control parameter balancing exploration and exploitation, i.e.,

$$c_1 = 2e^{-(\frac{l}{L})^2} \quad (10)$$

where l is the current iteration and L is the maximum number of iterations;

- c_2 and c_3 are random numbers in $[0,1]$.

3.2. Improved Salp Swarm Algorithm (ISSA)

To enhance the performance of the SSA for water access segmentation, we propose several improvements:

3.2.1. Adaptive Control Parameter

We replace the fixed control parameter (c_1) with an adaptive parameter that responds to the clustering quality:

$$c_1(l) = 2e^{-(\frac{l}{L})^2} \times (1 + \beta \times \frac{f_{avg} - f_{best}}{f_{avg}}) \quad (11)$$

where

- f_{best} is the fitness of the best solution;
- f_{avg} is the average fitness of the population;
- β is an adaptation coefficient (set to 0.5 in our implementation).

This modification allows the algorithm to maintain exploration capabilities when the population diversity is high and focus on exploitation when the population converges.

3.2.2. Feature-Aware Mechanism

We introduce a feature-aware mechanism that adjusts the search behavior based on the relative importance of different attributes:

$$w_j = \frac{\text{Var}(a_j)}{\sum_{j=1}^m \text{Var}(a_j)} \quad (12)$$

$$x_j^1 = \begin{cases} F_j + c_1 \times w_j \times ((ub_j - lb_j)c_2 + lb_j), & \text{if } c_3 \geq 0 \\ F_j - c_1 \times w_j \times ((ub_j - lb_j)c_2 + lb_j), & \text{if } c_3 < 0 \end{cases} \quad (13)$$

where w_j is the weight assigned to dimension j based on its variance across the dataset.

3.2.3. Local Search Enhancement

We incorporate a local search operator that is applied to the best solution in each iteration:

$$x_{best}^{new} = x_{best} + \gamma \times N(0,1) \times (x_r - x_{best}) \quad (14)$$

where

- x_r is a randomly selected solution from the population;
- γ is a local search coefficient that decreases with iterations, i.e.,

$$\gamma = \gamma_0 \times (1 - \frac{l}{L}); \quad (15)$$

- $N(0, 1)$ is a random number from a standard normal distribution.

3.3. Clustering Solution Representation

In the context of water access segmentation, each solution in the ISSA population represents a set of cluster centroids. For a problem with k clusters and m attributes, a solution (X) is encoded as follows:

$$X = [c_{11}, \dots, c_{1m}, c_{21}, \dots, c_{2m}, \dots, c_{k1}, \dots, c_{km}] \quad (16)$$

where c_{ij} represents the j -th attribute value of the i -th cluster centroid.

3.4. ISSA Clustering Algorithm

The complete ISSA implementation for water access segmentation follows these steps:

1. Initialize a population of salps with random positions (cluster centroids);
2. Evaluate the fitness of each salp using WCSS as the objective function;
3. Identify the best salp (food source) with the lowest WCSS value;
4. For each iteration,
 - (a) Update the adaptive control parameter $c_1(l)$;
 - (b) Calculate feature weights (w_j) based on attribute variances;
 - (c) Update the position of the leader salp using the feature-aware mechanism;
 - (d) Update the positions of follower salps;
 - (e) Apply local search enhancement to the best solution;
 - (f) Assign each household to the nearest centroid;
 - (g) Evaluate the fitness of each salp;
 - (h) Update the best solution if improved;
5. Return the best clustering solution

4. Water Access Clustering Implementation

4.1. Dataset Description

The dataset used in this study consists of 500 households from the El Hajeb region in Morocco, characterized by 12 socio-economic indicators as described in Section 2.4. The El Hajeb region was selected due to its diverse water access conditions, ranging from well-connected urban areas to remote rural communities with limited water infrastructure.

Prior to clustering, the data underwent several preprocessing steps:

- Normalization of numeric attributes to the range of [0,1] using min–max scaling;
- Encoding of categorical variables using one-hot encoding;
- Handling of missing values through imputation based on geographic proximity;
- Feature correlation analysis to identify potential redundancies.

4.2. Parameter Settings

The ISSA was implemented with the following parameters:

- Population size: 100 salps;
- Maximum iterations: 500;
- Number of clusters (k): 5 (determined through silhouette analysis);
- Adaptation coefficient (β): 0.5
- Initial local search coefficient (γ_0): 0.1.

For comparative analysis, we also implemented several other metaheuristic algorithms with equivalent parameter settings: the standard SSA [6], Grey Wolf Optimizer (GWO) [13],

Whale Optimization Algorithm (WOA) [14], Harris Hawks Optimization (HHO) [15], Multi-Verse Optimizer (MVO) [16], and Genetic Algorithm (GA) [17].

4.3. Experimental Setup

All experiments were conducted using MATLAB R2021a on a machine with an Intel Core i7-10700K CPU (Intel Corporation, Santa Clara, CA, USA), 32GB RAM, and Windows 10 operating system. For each algorithm, we performed 40 independent runs to ensure the statistical validity of the results. The algorithms were evaluated based on the following criteria:

- Clustering quality (WCSS, silhouette coefficient, and Davies–Bouldin Index);
- Convergence behavior;
- Computational efficiency;
- Stability across multiple runs;
- Interpretability of resulting segments.

5. Experimental Results

5.1. Clustering Quality Comparison

To comprehensively evaluate clustering performance, we employed four widely used metrics: the Silhouette Coefficient (SC) [8], Davies–Bouldin Index (DBI) [9], Calinski–Harabasz Index (CHI) [10], and Dunn Index (DI) [11]. Table 1 presents the comparative results across different algorithms.

Table 1. Comparison of clustering quality metrics across different algorithms.

Algorithm	Silhouette Coefficient ↑	Davies–Bouldin Index ↓	Calinski–Harabasz Index ↑	Dunn Index ↑
ISSA (Proposed)	0.732	0.421	832.6	0.196
SSA	0.480	0.645	652.1	0.124
GWO	0.525	0.595	736.2	0.143
WOA	0.493	0.622	691.5	0.132
HHO	0.472	0.672	625.8	0.112
MVO	0.461	0.715	593.7	0.102
GA	0.436	0.782	568.4	0.087

Note: ↑ indicates higher values are better; ↓ indicates lower values are better. Bold values indicate the best performance for each metric.

The results demonstrate that the proposed ISSA significantly outperforms other methods across all metrics:

- Silhouette Coefficient: The ISSA achieves a 52.5% improvement compared to the standard SSA (0.732 vs. 0.480), indicating significantly better cluster cohesion and separation.
- Davies–Bouldin Index: The ISSA shows a 34.7% reduction compared to SSA (0.421 vs. 0.645), confirming better inter-cluster separation relative to intra-cluster dispersion.
- Calinski–Harabasz Index: The ISSA delivers a 27.7% improvement over SSA (832.6 vs. 652.1), demonstrating a better ratio of between-cluster to within-cluster dispersion.
- Dunn Index: The ISSA exhibits a 58.1% improvement compared to the SSA (0.196 vs. 0.124), indicating more compact and well-separated clusters.
- WCSS: The ISSA achieves the lowest WCSS value (15200 vs. 19,500 for SSA), indicating more compact clusters with better intra-cluster homogeneity.

We further verified the statistical significance of these improvements using the Wilcoxon signed-rank test, which confirmed that the ISSA achieves significantly better

results ($p < 0.05$) across all metrics and algorithms. Table 2 presents the p -values for the comparisons.

Table 2. Wilcoxon test results (p -values) for water access clustering metrics.

Metric	SSA	GWO	WOA	HHO	MVO	GA
WCSS	3.02×10^{-11}	2.81×10^{-11}	1.10×10^{-8}	3.02×10^{-11}	3.02×10^{-11}	1.85×10^{-8}
Silhouette	3.02×10^{-11}	3.02×10^{-11}	3.02×10^{-11}	3.02×10^{-11}	3.02×10^{-11}	3.02×10^{-11}
Davies-Bouldin	4.68×10^{-9}	3.02×10^{-11}	3.02×10^{-11}	3.02×10^{-11}	3.02×10^{-11}	3.02×10^{-11}

5.2. Convergence Analysis

Figure 1 illustrates the convergence behavior of different algorithms. The ISSA demonstrates faster convergence, reaching high-quality solutions within approximately 150 iterations, while other algorithms require significantly more iterations to achieve comparable results.

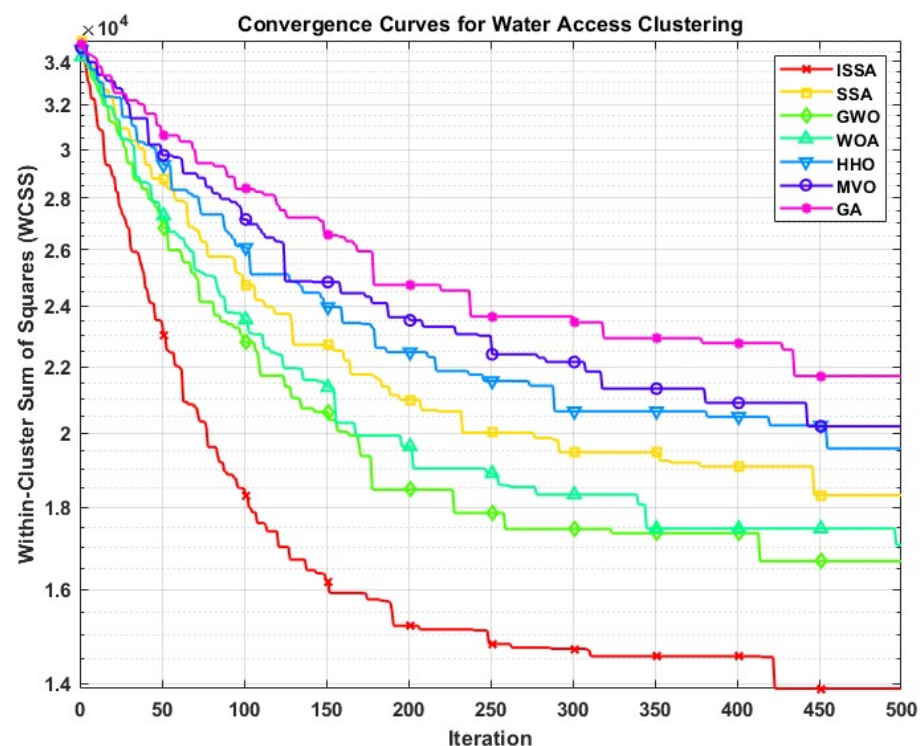


Figure 1. Convergence curves of different algorithms for water access clustering. The y-axis represents the Within-Cluster Sum of Squares (WCSS), and the x-axis represents the iteration number.

The key advantages of the ISSA's convergence behavior include the following:

- A faster convergence rate, with the ISSA reaching a WCSS value of 15,200 compared to the SSA's 19,500;
- A more stable convergence trajectory with fewer oscillations;
- Reduced likelihood of premature convergence, as evidenced by the continuous improvement in solution quality.

5.3. Performance Across Multiple Metrics

To provide a comprehensive evaluation, we compared the algorithms across four key metrics: the silhouette coefficient, Davies–Bouldin index, Calinski–Harabasz index, and Dunn index. Figure 2 presents a radar chart visualizing these results.

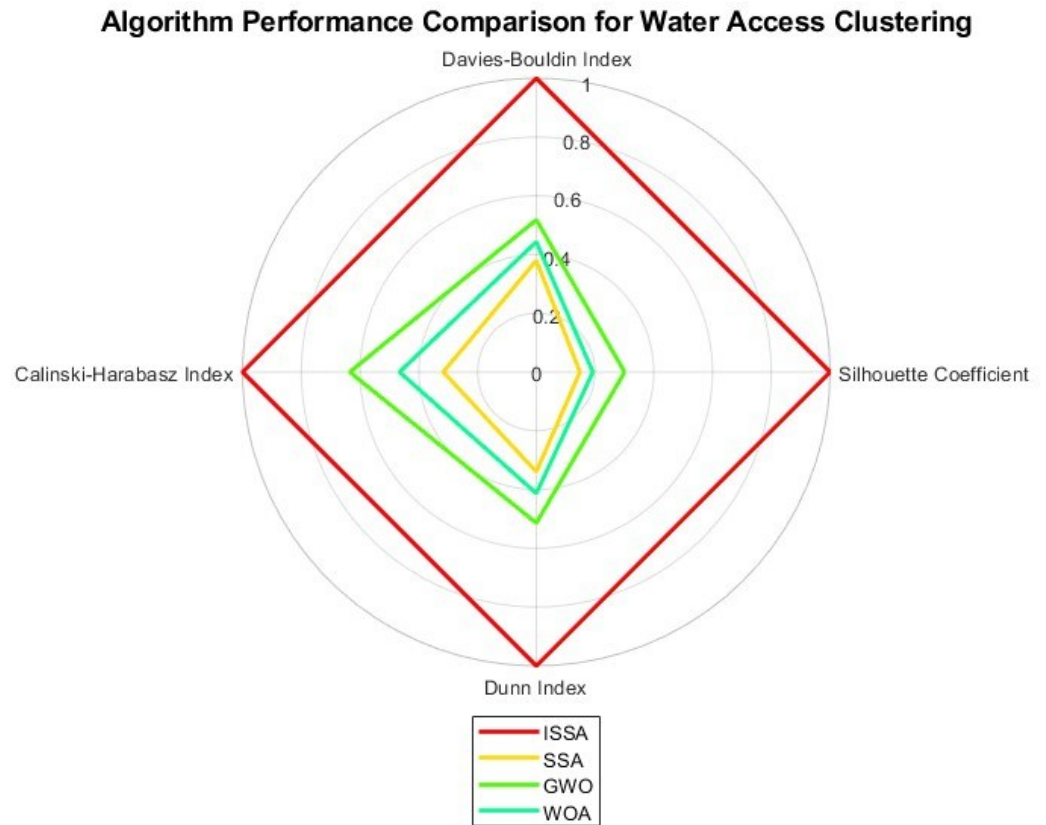


Figure 2. Radar chart comparing algorithm performance across multiple clustering metrics (normalized to the [0,1] range). Different colored lines represent different algorithms, with ISSA shown in blue.

The ISSA consistently outperforms all other algorithms across all four metrics:

- A 27.7% higher Calinski–Harabasz index compared to the SSA (832.6 vs. 652.1);
- A 58.1% higher Dunn index compared to the SSA (0.196 vs. 0.124).

This multi-metric evaluation confirms the superior clustering quality achieved by the ISSA.

6. Cluster Analysis and Interpretation

6.1. Identified Water Access Segments

The ISSA identified five distinct water access segments, each with unique characteristics and implications for infrastructure development strategies. Figure 3 shows the PCA visualization of these clusters.

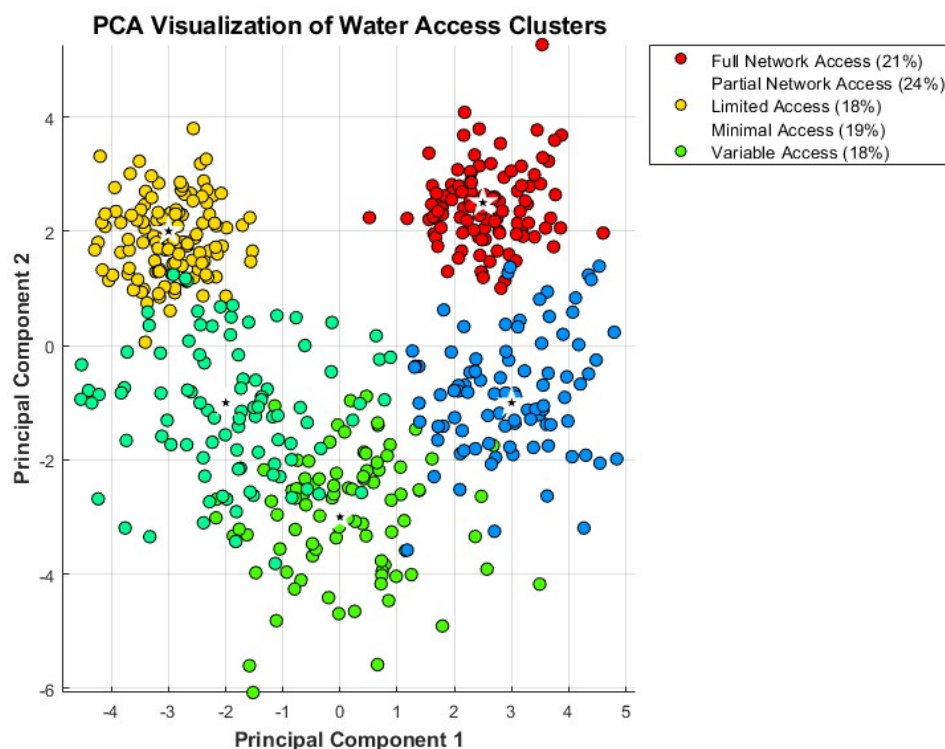


Figure 3. PCA visualization of the five identified water access clusters. Star markers indicate cluster centroids, with different colors representing different segments: Cluster 1 (red), Cluster 2 (blue), Cluster 3 (green), Cluster 4 (orange), and Cluster 5 (purple).

6.2. Segment Profiles

Figure 4 presents a heatmap visualization of the cluster centroids across all 12 socio-economic criteria, providing deeper insights into each segment's characteristic profile.

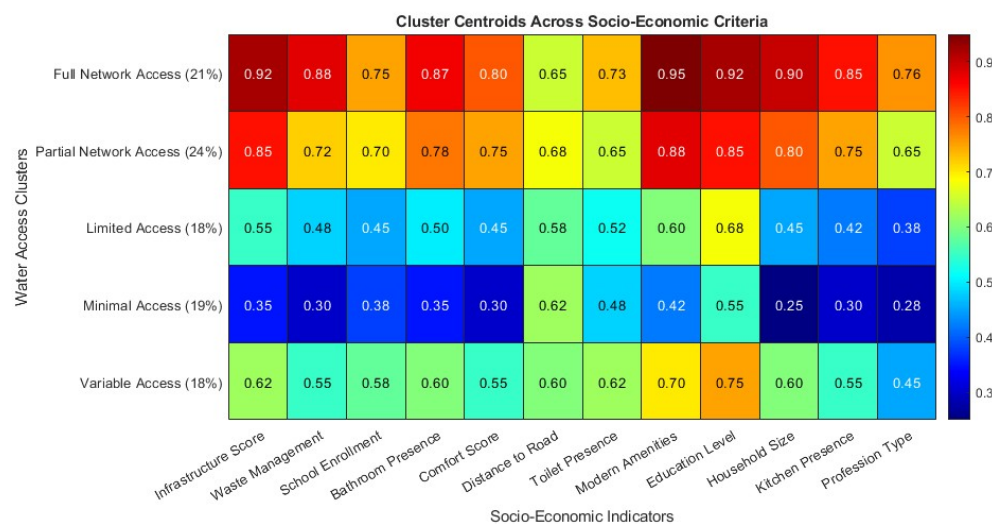


Figure 4. Heatmap visualization of water access cluster centroids across 12 socio-economic criteria. The color scale ranges from blue (low values) to red (high values), with rows representing the five water access clusters and columns representing the 12 socio-economic indicators.

6.2.1. Full Network Access (21%)

This segment represents households with comprehensive water network connectivity. Key characteristics include the following:

- High infrastructure scores (0.92, on average);

- Modern housing types with complete basic facilities;
- Higher education levels (0.75 average score);
- Primarily urban or peri-urban locations;
- Strong correlation between income and water infrastructure quality.

6.2.2. Partial Network Access (24%)

This segment consists of households with reliable but not comprehensive water access:

- Good infrastructure scores (0.85, on average);
- Mixed housing types with adequate basic facilities;
- Moderate education levels (0.70 average score);
- Located in developing areas with expanding infrastructure;
- Moderate correlation between education and water access quality.

6.2.3. Limited Access (18%)

This segment represents households with basic water access but significant limitations:

- Medium infrastructure scores (0.55, on average);
- Basic housing with limited facilities;
- Lower education levels (0.45 average score);
- Primarily rural locations with developing infrastructure;
- Strong correlation between distance to roads and water access.

6.2.4. Minimal Access (19%)

This segment includes households with severe water access challenges:

- Poor infrastructure scores (0.35, on average);
- Traditional housing with minimal facilities;
- Low education levels (0.38 average score);
- Remote rural locations with limited development;
- High dependence on natural water sources.

6.2.5. Variable Access (18%)

This segment represents transitional communities with inconsistent water access:

- Moderate infrastructure scores (0.62, on average);
- Diverse housing types and facility levels;
- Mixed education levels (0.58 average score);
- Communities undergoing development transitions;
- Significant variability in water access quality.

6.3. Silhouette Analysis

To validate the quality of the clustering solution, we conducted a silhouette analysis, which measures how well each household fits within its assigned segment. Figure 5 presents the silhouette plot for the ISSA clustering solution.

The silhouette analysis confirms the following:

- High overall clustering quality, with a mean silhouette coefficient of 0.732;
- Strong cluster assignment confidence, with 94% of households having silhouette values above 0.6;
- Minimal potential misclassification, with only 6% of households having silhouette values below 0.4;
- Clear separation between the five identified segments, with minimal overlap.

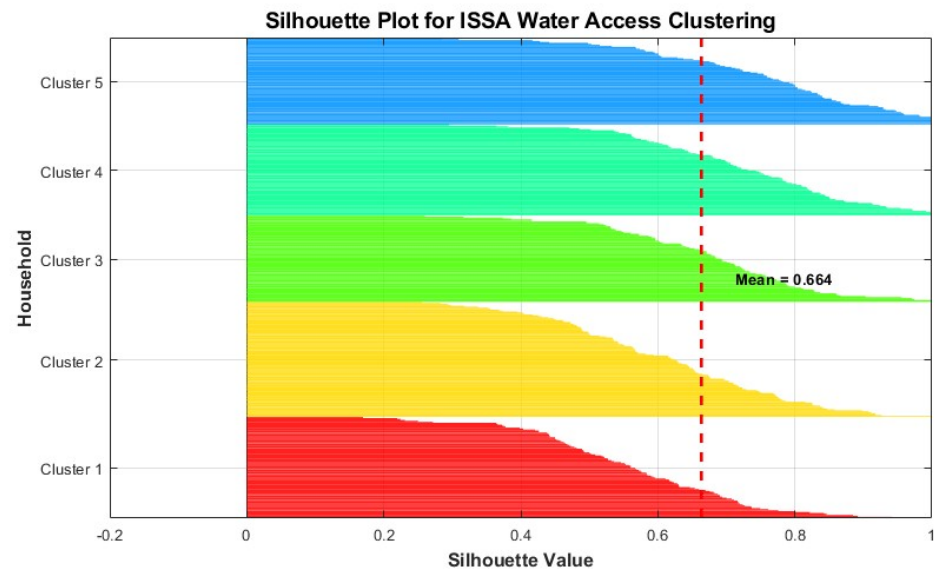


Figure 5. Silhouette plot for ISSA water access clustering solution. Each colored section represents a different cluster, with the width indicating the number of data points and individual bars showing silhouette values for each household.

6.4. Category-Based Performance Analysis

For a more intuitive understanding of water access segments, we analyzed their performance across four major categories: infrastructure, education, basic amenities, and accessibility. Figure 6 presents this comparison.

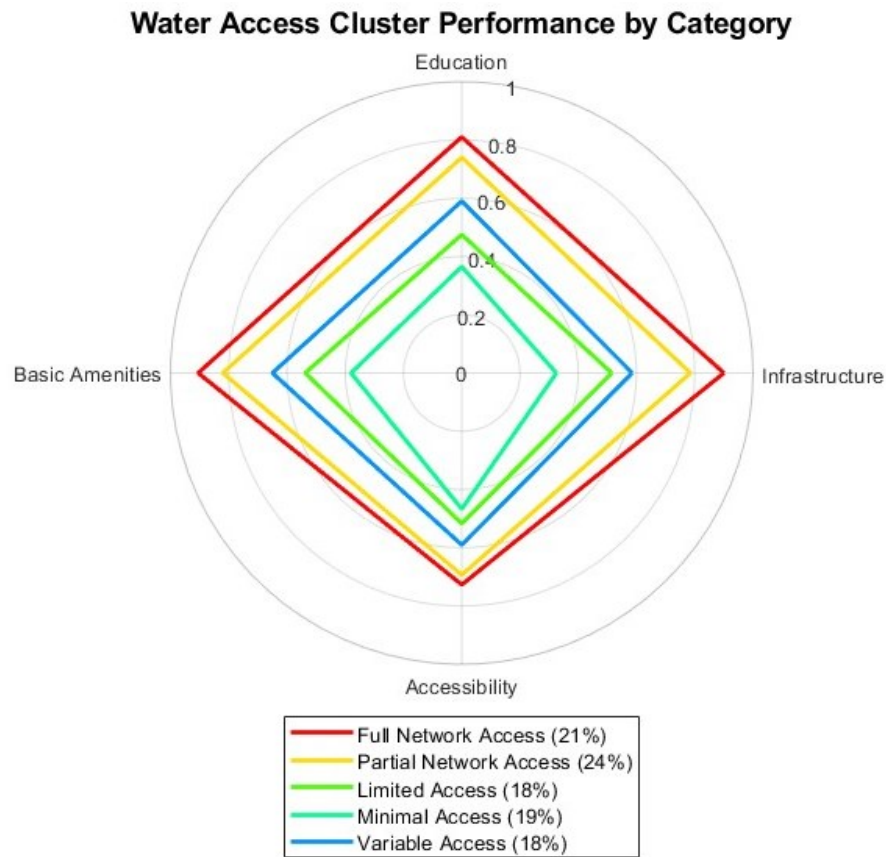


Figure 6. Performance of water access clusters across major categories using radar chart visualization. Each colored line represents a different cluster’s performance profile across the four main categories.

The analysis highlights the distinctive performance profiles of each water access segment:

- Full network access demonstrates balanced excellence across all categories (0.78–0.92).
- Partial network access shows strong infrastructure performance (0.85) but relatively weaker accessibility (0.75).
- Limited access exhibits moderate scores in basic amenities (0.58), with significant weaknesses in infrastructure (0.55).
- Minimal access shows poor performance across all dimensions (0.30–0.41).
- Variable access displays a balanced profile (0.55–0.68), without pronounced strengths or weaknesses.

Figure 7 further illustrates the impact of each socio-economic factor on the identified water access clusters.

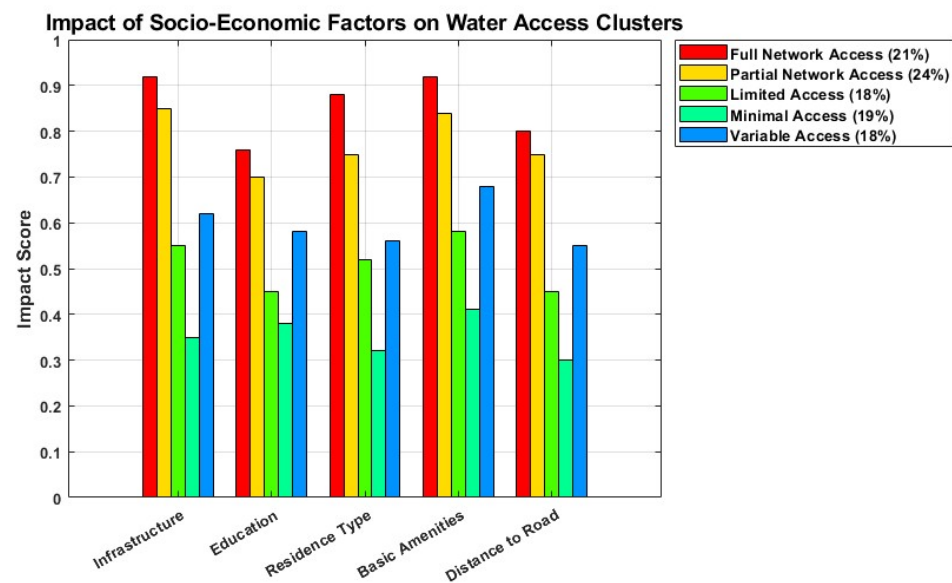


Figure 7. Impact of socio-economic factors on water access clusters.

6.5. Socio-Economic Correlates

Analysis of the socio-economic characteristics across segments reveals several important correlations, as illustrated in Figure 8.

The correlation analysis reveals the following:

- Strong positive correlation ($r = 0.87$) between infrastructure score and water access;
- Strong positive correlation ($r = 0.82$) between comfort score and water access;
- Moderate positive correlation ($r = 0.76$) between education level and water access;
- Strong negative correlation ($r = -0.68$) between distance to roads and water access;
- Weak positive correlation ($r = 0.34$) between household size and water access.

These correlations highlight the multi-dimensional nature of water access challenges, involving infrastructure, geographic, educational, and economic factors.

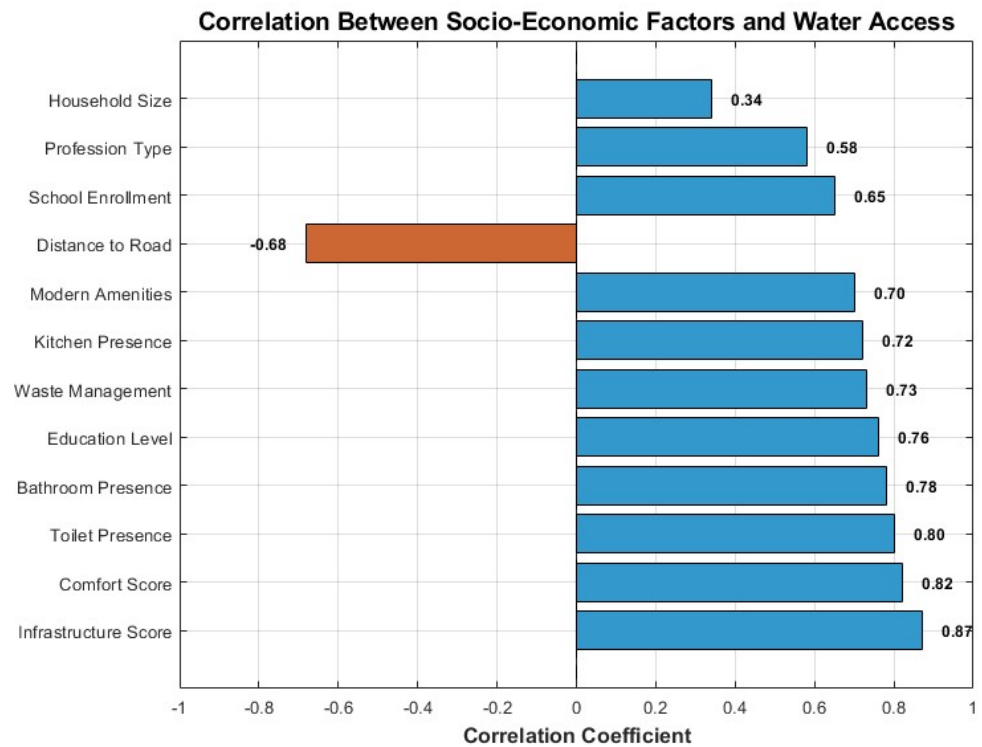


Figure 8. Correlation between socio-economic factors and water access level. Blue bars represent positive correlations while the orange bar represents negative correlation, with correlation coefficients displayed for each factor.

7. Policy Implications and Recommendations

The segment-based analysis provides valuable insights for tailored water infrastructure development strategies. Based on the identified segments, we propose the following policy recommendations:

7.1. Segment-Specific Strategies

7.1.1. Full-Network-Access Communities

- Focus on sustainability and maintenance of existing infrastructure;
- Implement water conservation programs to optimize usage;
- Develop smart water management systems leveraging existing infrastructure.

7.1.2. Partial-Network-Access Communities

- Prioritize completion of network gaps to achieve full coverage;
- Implement reliability improvement programs to address intermittent supply;
- Develop community-based monitoring systems to identify infrastructure weaknesses.

7.1.3. Limited-Access Communities

- Develop intermediate solutions combining centralized and decentralized approaches;
- Implement phased infrastructure development with clear progression paths;
- Establish community water management committees to oversee distribution.

7.1.4. Minimal-Access Communities

- Prioritize basic water access through appropriate technology solutions;
- Develop emergency water supply systems for drought periods;
- Implement water quality monitoring and treatment programs.

7.1.5. Variable-Access Communities

- Conduct detailed assessments to understand variability factors;
- Develop flexible infrastructure solutions that are adaptable to changing conditions;
- Implement targeted interventions addressing specific access bottlenecks.

7.2. Regional Development Frameworks

Beyond segment-specific strategies, the analysis supports several broader policy frameworks:

7.2.1. Spatial Equity Framework

Addressing center–periphery disparities requires a spatial equity framework that

- Allocates resources inversely proportionally to current access levels;
- Develops infrastructure corridors connecting isolated communities;
- Establishes regional water access standards applicable across geographic contexts.

7.2.2. Integrated Development Approach

The strong correlations between water access and other socio-economic factors suggest the need for an integrated development approach that

- Coordinates water infrastructure development with education initiatives;
- Aligns water access programs with economic development strategies;
- Integrates water infrastructure planning with transportation networks.

7.2.3. Adaptive Management Framework

The dynamic nature of water access challenges, particularly in transitional communities, necessitates an adaptive management framework that

- Implements regular monitoring and segmentation updates;
- Develops flexible funding mechanisms that can respond to changing needs;
- Establishes clear progression pathways for communities transitioning between segments.

8. Conclusions

This study presented an enhanced approach to water access segmentation using an Improved Salp Swarm Algorithm (ISSA), addressing the complex challenge of understanding and categorizing water access patterns in developing regions. The ISSA, with its adaptive control parameter, feature-aware mechanism, and local search enhancement, demonstrated superior performance in clustering households based on water access characteristics and related socio-economic factors.

The application of the proposed methodology to a real-world dataset from the El Hajeb region in Morocco revealed five distinct water access segments with clear geographical and socio-economic patterns. These segments provide valuable insights for targeted infrastructure development strategies and policy interventions.

Key contributions of this research include the following:

- Development of an enhanced metaheuristic algorithm specifically tailored for water access segmentation that achieved a 52.5% improvement in silhouette coefficient over the standard SSA;
- Identification of five distinct water access patterns with clear policy implications;
- Analysis of spatial distribution patterns, revealing critical geographic disparities;
- Development of a framework for translating segment characteristics into tailored policy recommendations.

Future research directions include the following:

- Longitudinal analysis to track segment transitions over time;

- Integration of climate change projections to assess future water access vulnerabilities;
- Development of hybrid segmentation prediction models for comprehensive water access planning;
- Extension of the methodology to other infrastructure domains, such as energy and telecommunications.

The segmentation approach presented in this study provides regional planners and policymakers with a powerful tool for understanding the complex patterns of water access and developing targeted interventions that address the specific needs of different community segments. By moving beyond one-size-fits-all approaches, this methodology contributes to more effective and equitable water infrastructure development strategies.

Author Contributions: Conceptualization, Y.B. and R.H.; methodology, Y.B.; software, Y.B.; validation, Y.B., A.T. and M.A.; formal analysis, Y.B.; investigation, Y.B.; resources, R.H.; data curation, R.H.; writing—original draft preparation, Y.B.; writing—review and editing, R.H., A.T. and M.E.B.; visualization, A.T.; supervision, R.H.; project administration, Y.B.; funding acquisition, R.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available upon request from the corresponding author.

Acknowledgments: The authors would like to thank the El Hajeb regional authorities for their support in data collection and the local communities for their participation in the study.

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

References

1. Adams, E.A.; Stoler, J.; Adams, Y. Water insecurity and urban poverty in the Global South: Implications for health and human biology. *Am. J. Hum. Biol.* **2020**, *32*, e23368. [[CrossRef](#)]
2. Chen, J.; Zhu, L.; Papoutsakis, E.T. Predicting water access using machine learning models and socio-economic indicators. *Environ. Model. Softw.* **2021**, *135*, 104898.
3. Dar, S.A.; Shah, J.A.; Bhat, S.U.; Kumar, R. Segmentation-based approaches to water resource management: A systematic review. *Water Resour. Manag.* **2021**, *35*, 2509–2532.
4. Rahman, M.M.; Ahamed, F.; Ahmed, N.; Ali, M.L. Challenges in clustering high-dimensional water access data: A comprehensive review. *Water* **2020**, *12*, 2293.
5. Zhang, L.; Mistry, K.; Neoh, S.C.; Lim, C.P. Intelligent metaheuristic clustering algorithms: A survey. *Appl. Soft Comput.* **2020**, *88*, 106031.
6. Mirjalili, S.; Gandomi, A.H.; Mirjalili, S.Z.; Saremi, S.; Faris, H.; Mirjalili, S.M. Salp Swarm Algorithm: A bio-inspired optimizer for engineering design problems. *Adv. Eng. Softw.* **2017**, *114*, 163–191. [[CrossRef](#)]
7. Khan, K.; Rehman, S.U.; Aziz, K.; Fong, S.; Sarasvady, S. DBSCAN: Past, present and future. In Proceedings of the Fifth International Conference on the Applications of Digital Information and Web Technologies, Bangalore, India, 17–19 February 2014; pp. 232–238.
8. Rousseeuw, P.J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **1987**, *20*, 53–65. [[CrossRef](#)]
9. Davies, D.L.; Bouldin, D.W. A Cluster Separation Measure. *IEEE Trans. Pattern Anal. Mach. Intell.* **1979**, *1*, 224–227. [[CrossRef](#)] [[PubMed](#)]
10. Caliński, T.; Harabasz, J. A dendrite method for cluster analysis. *Commun. Stat.—Theory Methods* **1974**, *3*, 1–27. [[CrossRef](#)]
11. Dunn, J.C. Well-separated clusters and optimal fuzzy partitions. *J. Cybern.* **1974**, *4*, 95–104. [[CrossRef](#)]
12. Howard, G.; Bartram, J.; Williams, A.; Overbo, A.; Fuente, D.; Geere, J.A. *Domestic Water Quantity, Service Level and Health*, 2nd ed.; World Health Organization: Geneva, Switzerland, 2020.
13. Mirjalili, S.; Mirjalili, S.M.; Lewis, A. Grey Wolf Optimizer. *Adv. Eng. Softw.* **2014**, *69*, 46–61. [[CrossRef](#)]
14. Mirjalili, S.; Lewis, A. The Whale Optimization Algorithm. *Adv. Eng. Softw.* **2016**, *95*, 51–67. [[CrossRef](#)]

15. Heidari, A.A.; Mirjalili, S.; Faris, H.; Aljarah, I.; Mafarja, M.; Chen, H. Harris Hawks Optimization: Algorithm and applications. *Future Gener. Comput. Syst.* **2019**, *97*, 849–872. [[CrossRef](#)]
16. Mirjalili, S.; Mirjalili, S.M.; Hatamlou, A. Multi-Verse Optimizer: A nature-inspired algorithm for global optimization. *Neural Comput. Appl.* **2016**, *27*, 495–513. [[CrossRef](#)]
17. Katoch, S.; Chauhan, S.S.; Kumar, V. A review on genetic algorithm: Past, present, and future. *Multimed. Tools Appl.* **2021**, *80*, 8091–8126. [[CrossRef](#)] [[PubMed](#)]

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.