ARTICLE

Simulation of multi-criteria mean Clustering method for satellite imagery of Bogd Khan mountain of Mongolia

Bayanjargal Darkhijav[1], Davaajargal Jargalsaikhan[2], Munkh-Erdene Altangerel[2] and Enkhbat Rentsen[2]*

[1]Department of Statistics and Applied Mathematics, School of Information Technology and Electronics, National University of Mongolia, Ulaanbaatar, Mongolia

[2]Mathematics division, Institute of Mathematics and Digital Technology, Mongolian Academy of Sciences, Ulaanbaatar, Mongolia

Abstract: Traditionally, a variety of unsupervised methods have been used for analyzing remotely-sensed imagery. Although popular clustering techniques such as K-means, Mini-Batch K-means, and Fuzzy C-means are widely applied, their classification accuracy is often limited. This study proposes a new clustering method, called Multi-Criteria Mean Clustering (MCMC), to classify and extract forest areas from remote sensing images. The proposed method is based on a multi-criteria optimization framework, and leverages Pareto-optimal solutions arising from multiple clustering objectives. We show that K-means and Mini-Batch K-means can be viewed as specific instances of the proposed MCMC approach. For experimental evaluation, Root Mean Square Error (RMSE), Structural Similarity Index (SSIM), Overall Accuracy (OA), and Intersection over Union (IU) were used to assess performance. The results demonstrate that the proposed MCMC method provides more accurate and reliable classification, achieving lower RMSE, higher SSIM, higher OA, and higher IU compared to conventional clustering techniques. The study area is Bogd Khan Mountain, located in Central Mongolia, on the southern fringes of Ulaanbaatar. Sentinel-2B remote sensing data were employed for this research, and all computational experiments were conducted in Python Jupyter Notebook.

Keyword: *cluster, clustering methods, remote sensing data;*

INTRODUCTION

Currently, researchers around the world are actively applying the latest advancements in remote sensing technology in various fields of study. Through automatic interpretation of remotely sensed images and advanced knowledge-based techniques, thematic results are generated and used for a variety of scientific purposes. A vast amount of remote sensing data, including optical, hyperspectral, and synthetic aperture radar (SAR) data, has been produced and is increasingly applied across different scientific domains. These data are particularly valuable for environmental and natural resource management. The diversity of research objectives and the unique characteristics of remote sensing data have led to the application of a wide range of machine-learning and signal-processing algorithms [1].

Corresponding author, email: *davaajargal_j@mas.ac.mn* *https://orcid.org/0000-0001-7585-2943*

The statistical characterization of remote sensing imagery is challenging due to the high dimensionality of pixel data, the presence of various noise sources and uncertainties, the inherent nonlinear nature of the data, and high spatial and spectral redundancies [1]. Machine-learning has been successfully applied in remote sensing for classification, regression, clustering, coding, and source separation tasks [1]. Despite the availability of automated techniques for satellite data processing, many applications still require manual intervention and human expertise for decision-making. Artificial intelligence and machine-learning techniques aim to address these challenges by enabling computational processes to function in a manner similar to human reasoning [2].

Clustering, also known as cluster analysis, has become an important machine-learning technique for discovering natural groupings within observed data [3], [4]. It serves as a powerful tool for pattern discovery, data organization, and gaining insight into the inherent structure of datasets. In remote sensing, clustering is commonly used for grouping similar pixels or objects in satellite or aerial imagery, helping to identify patterns, regions of interest, and distinguish different land-cover or land-use types. Data clustering involves identifying natural groupings or clusters within multidimensional data based on similarity measures, such as Euclidean, Manhattan, or Minkowski distance [5], [6], [7], [8]. Clustering methods are widely applied across many fields, including pattern recognition [9], [10], image segmentation [11], [12], [13], vector and color image quantization [14], [15], [16], and data mining [17], [18]. The choice of a clustering method depends on the characteristics of the remote sensing data, such as the number and distribution of classes, as well as the spatial and spectral properties of the imagery. Among these methods, K-means is one of the most widely used clustering algorithms. It is extensively applied in data mining, image segmentation, and pattern recognition, where the primary objective is to group similar data points and uncover underlying patterns or structures within the data.

The K-means clustering algorithm optimizes the following objective function:

$$J(C) = \sum_{i=1}^{n} \sum_{j=1}^{k} \left\| x_i - c_j \right\|^2 \rightarrow min \qquad (1)$$

where x_i, $i = 1, \ldots, n$, are the data points, c_j, $j = 1, \ldots, k$, are the cluster centroids, and $\|\cdot\|$ denotes the Euclidean distance. Despite its limitations, K-means is a computationally efficient and widely used clustering algorithm. It is particularly suitable for relatively well-structured datasets and is often employed as a baseline or starting point for more advanced clustering methods.

Mini-Batch K-means is a variant of the traditional K-means algorithm that employs a stochastic optimization strategy by using randomly selected subsets (mini-batches) of data to iteratively update the cluster centroids. The objective function of Mini-Batch K-means is a stochastic approximation of the original K-means objective and is defined as follows:

$$J(C) = \frac{1}{|M|} \sum_{i \in M} \sum_{j=1}^{k} r_{ij} \left\| x_i - c_j \right\|^2 \rightarrow min \qquad (2)$$

Here, M denotes a randomly selected mini-batch of data points, and $|M|$ is the size of the mini-batch. The variable r_{ij} is an indicator defined as:

$$r_{ij} = \begin{cases} 1 \ x_i \ is \ assigned \ to \ cluster \ j \\ 0 \ other \ wise \end{cases}$$

Fuzzy C-Means (FCM) is a clustering method that extends the classical K-means algorithm. It was introduced by James C. Bezdek [19], [20], following his significant contributions to pattern recognition and fuzzy systems. The objective function of FCM, which is minimized during the iterative optimization process, is expressed in terms of fuzzy membership values and cluster centroids as follows:

$$J_m(C) = \sum_{i=1}^{n} \sum_{j=1}^{k} u_{ij}^m \|x_i - c_j\|^2 \to min \qquad (3)$$

subject to the constraint:

$$J(C) = \sum_{i=1}^{k} u_{ij}^m = 1, \quad j = \overline{1,n}$$

Here, m is a weighting exponent that controls the degree of fuzziness in the partition, and u_{ij}^m represents the membership degree of data point x_i in cluster c_j. Typically, **m=2** is used in most applications. FCM has been widely applied in image processing, pattern recognition, and data mining, where soft clustering is advantageous for modeling uncertainty and overlapping data structures. In this paper, we propose a multi-criteria optimization-based clustering approach applied to Sentinel-2B remote sensing imagery over selected forested areas of Bogd Khan Mountain, Central Mongolia. The results demonstrate that the proposed method outperforms traditional clustering approaches, including K-means**,** Mini-Batch K-means, and Fuzzy C-Means.

Multi-Criteria mean clustering (MCMC)

The general formulation of clustering involves grouping a set of data points into subsets, or clusters, based on some similarity or proximity measure. Let **X={x₁,x₂,…,xₙ}** be a set of n data points in a feature space, and let **C={c₁,c₂,…,cₖ}** be a set of **k** clusters. The term **d(xᵢ,cⱼ),** for **i=1,…,n** and **j=1,…,k**, represents the distance (e.g., Manhattan, Minkowski, Euclidean) or dissimilarity between a data point **xᵢ** and a cluster centroid **cⱼ**.

The objective function J quantifies the quality of clustering and is optimized during the clustering process. It is commonly formulated as:

$$J_m(C) = \sum_{i=1}^{n} \sum_{j=1}^{k} w_{ij} d(x_i, c_j) \to min \qquad (4)$$

where c_j is the centroid of the **j-th** cluster, w_{ij} is a weight indicating the association of **xᵢ** with cluster **cⱼ,** and **d(xᵢ,cⱼ)** is the distance between **xᵢ** and **cⱼ**. The clustering process involves finding the optimal assignment of data points to clusters in a way that minimizes the objective function. This is typically achieved through an iterative optimization procedure. We propose a clustering method based on a multi-criteria optimization approach. Denote by C the set of points in a cluster. We introduce the following multi-criteria formulation in K-means clustering:

$$\varphi_j(c) = \sum_{i=1}^{n} \|x_i - c_j\|^2, \; j = 1, \dots, k, \; c \in C \tag{5}$$

This formulation allows the clustering process to simultaneously consider multiple criteria, improving the robustness and quality of the resulting clusters.

Definition: A point $c^* \in C$ is called a Pareto optimal point of problem (5), if there is no other $c \in C$ such that

$$\varphi_j(c^*) \leq \varphi_j(c), \quad j = 1, \dots, k,$$
$$\varphi_j(c^*) \neq \varphi_j(c), \quad \varphi = (\varphi_1, \varphi_2, \dots, \varphi_k)$$

The concept of Pareto optimality is the primary notion of optimality in multi-objective optimization. A common approach for finding Pareto optimal solutions is the weighted sum method.

Let $\alpha_j > 0$ for $j=1,\dots,k$ denote positive weights. Using these weights, we can formulate the following scalarized optimization problem corresponding to (5):

$$\min_{c \in C} F(c) = \sum_{j=1}^{k} \alpha_j \, \varphi_j(c) \tag{6}$$

The relationship between Pareto optimal solutions of the multi-objective problem and the solutions of this scalarized problem is given by the following assertion.

Proposition: A solution c^* to problem (6) is a Pareto optimal solution to problem (5).

Proof: Let us assume $c*$ is a solution to problem (6), but $c*$ is not a Pareto optimal solution of problem (5). Then there exists a point $\bar{c} \in C$ and an index $\bar{j} \in \{1,2,\dots,k\}$, such that

$$\varphi_j(\bar{c}) \geq \varphi_j(c^*), \quad j = 1, \dots, k,$$
$$\varphi_j(\bar{c}) > \varphi_j(c^*)$$

By the definition of the scalarized objective function F(c), we have

$$F(c^*) = \sum_{j=1}^{k} \sum_{i=1}^{n} \alpha_j \|x_i - c_j^*\|^2 = \sum_{j=1}^{k} \alpha_j \varphi_j(c^*) \leq \sum_{j=1}^{k} \alpha_j \varphi_j(\bar{c}) < \sum_{j=1}^{k} \left(\alpha_j \varphi_j(\bar{c}) + \alpha_j \varphi_{\bar{j}}(\bar{c}) \right) = F(\bar{c})$$

which contradicts the assumption that $c*$ is a solution to problem (6). Hence, $c*$ must be a Pareto optimal solution of problem (5). Proof is completed.

In the special case where $\alpha_j=1$ for all j, problem (6) reduces to the standard K-means clustering problem. The advantage of the multi-criteria optimization approach is that by selecting appropriate weights α_j, the quality of clustering can be improved.

For example, if in (6) we set $\alpha_j = \frac{1}{|M|} \sum_{i=1}^{n} r_{ij}$, the problem becomes equivalent to Mini-Batch K-means clustering. Similarly, Fuzzy C-means clustering can be regarded as a particular case of problem (6). In our numerical experiments, we considered **k=2** clusters and set the weights as $\alpha_1=0.7$ and $\alpha_2=0.3$ reflecting the relative importance of the two clustering criteria.

Test area and data

The ground truth information was obtained through a forest inventory, which, by law, must be conducted for designated forest areas in Mongolia. In the study area, the inventory was conducted from 2013 to 2014, and the data analysis and processing were carried out by the Forest Division of the Ministry of Nature and Environment (MNE). Bogd Khan Mountain, one of Mongolia's oldest protected areas, is located in Central Mongolia, south of Ulaanbaatar, the capital city (Figure 1). Historically, the mountain has been considered a sacred site by the Mongolian people. The area covers 416.51 km², of which approximately 55% consists of highly mixed forest types [21]. The mountain hosts 588 species of higher plants, belonging to 256 genera and 70 families. Among these, 135 species, such as *Carex*, *Artemisia*, and *Oxytropis*, from 11 major genera, account for 22.9% of all plant species distributed across the mountain. The elevation of the forest spans from 1,400 m to 2,150 m above sea level and is classified into three sub-zones: mountain plateau, taiga, and taiga-type forests. Cedar and larch dominate the forest cover, with pine, birch, spruce, and poplar also present [22]. For this study, Sentinel-2B multispectral satellite imagery, acquired on 1 October 2023, was used. Figure 1 shows the Sentinel-2B image of the test area. Sentinel-2B, part of the European Space Agency's (ESA) Copernicus program, carries a Multi-Spectral Instrument (MSI) capturing imagery in 13 spectral bands.

This satellite provides high-quality, high-resolution Earth observation data essential for environmental monitoring, resource management, and scientific research. The data are widely used by researchers, scientists, and the public for applications that enhance understanding of Earth's dynamics and environmental changes over time.
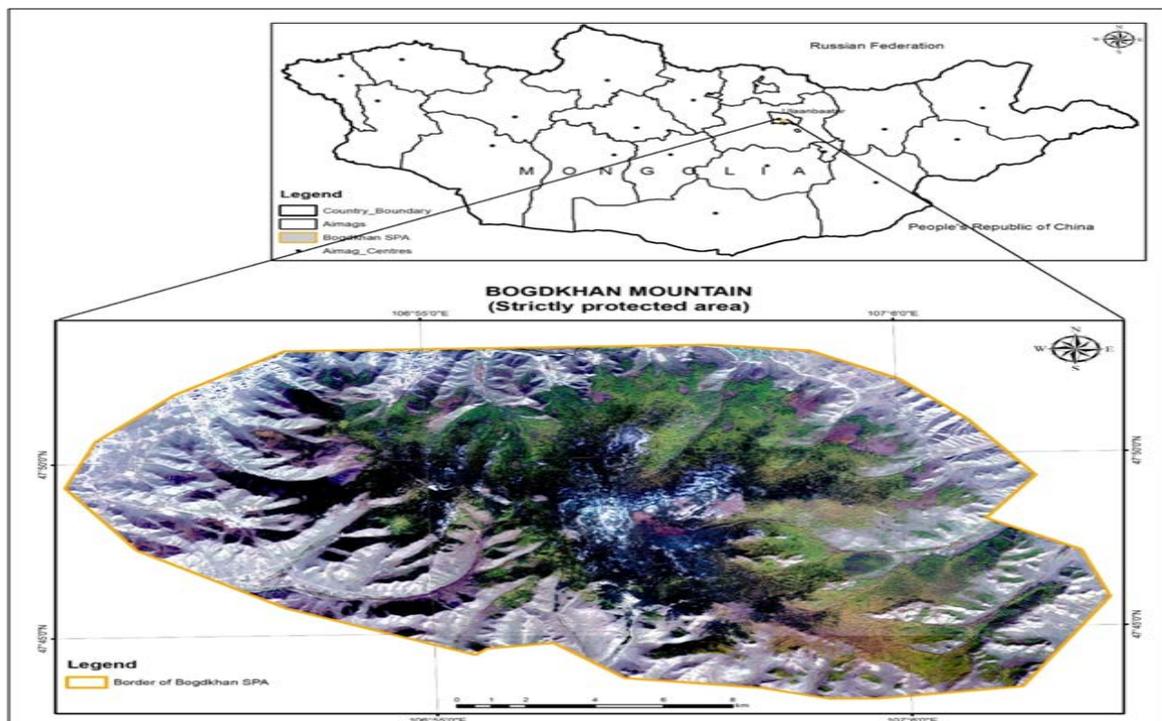

Figure 1. Test Area.

RESULTS

In the numerical experiments, the performance of the clustering methods was evaluated using Root Mean Square Error (RMSE), Structural Similarity

Index (SSIM), Overall Accuracy (OA), and Intersection over Union (IU), treating the remote sensing image as a pixel-wise classification problem. RMSE is a widely used metric to measure the differences between predicted values and observed values. It

is particularly useful in regression analysis and image processing to evaluate the accuracy of models or comparisons. In our case, RMSE will compare the corresponding pixels of two images and compute the errors. If RMSE = 0, two images are identical (perfect match), if RMSE = 1, two images are completely different. Each pixel has either 0 (white, non-forest area) or 1 (green, forest area). RMSE measures the pixel-wise differences between predicted and observed values:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x}_i)^2} \quad (7)$$

where **n** is the number of pixels, x_i is the predicted pixel value, and $\overline{x}_i$ is the corresponding ground truth value. SSIM evaluates structural similarity, capturing perceptual and spatial information. SSIM is a method for measuring the similarity between two images. SSIM

values range from -1 to 1, where 1 indicates perfect structural similarity, 0 indicates no structural similarity, negative values can occur but are less common in typical comparisons. The SSIM is defined mathematically as follows:

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_x^2 + C_1)(\sigma_x^2 + \sigma_x^2 + C_2)} \quad (8)$$

where **x** and **y** are the two images being compared, μ_x, μ_y are mean intensities, σ_x, σ_y are variances, σ_{xy} is covariance, and $C_1 \approx 0.0001, C_2 \approx 0.0009$ are small constants.

In Addition, the inclusion of OA and IU strengthens the evaluation, providing a comprehensive assessment

of both pixel-wise accuracy and spatial consistency, confirming that the proposed multi-criteria clustering method outperforms conventional methods. Overall Accuracy (OA) measures the proportion of correctly classified pixels:

$$OA = \frac{\text{Number of correctly classified pixels}}{\text{Total number of pixels}} \quad (9)$$

and IU quantifies the spatial overlap between predicted clusters and ground truth:

$$IU = \frac{TP}{TP + FP + FN} \quad (10)$$

where TP (true positives), FP (false positives), and FN (false negatives) are the numbers of pixels corresponding to the target class.

The experiments were conducted in Python Jupyter Notebook on an image

containing 506,834 pixels (899 × 566). Figure 2 shows the results of applying the proposed multi-criteria clustering method in comparison with K-means, Mini-Batch K-means, and Fuzzy C-means on the ground truth image. Table

1 summarizes the quantitative evaluation of all clustering methods. From the results, it is evident that the proposed multi-criteria method outperforms the other three approaches in terms of both accuracy and structural preservation.
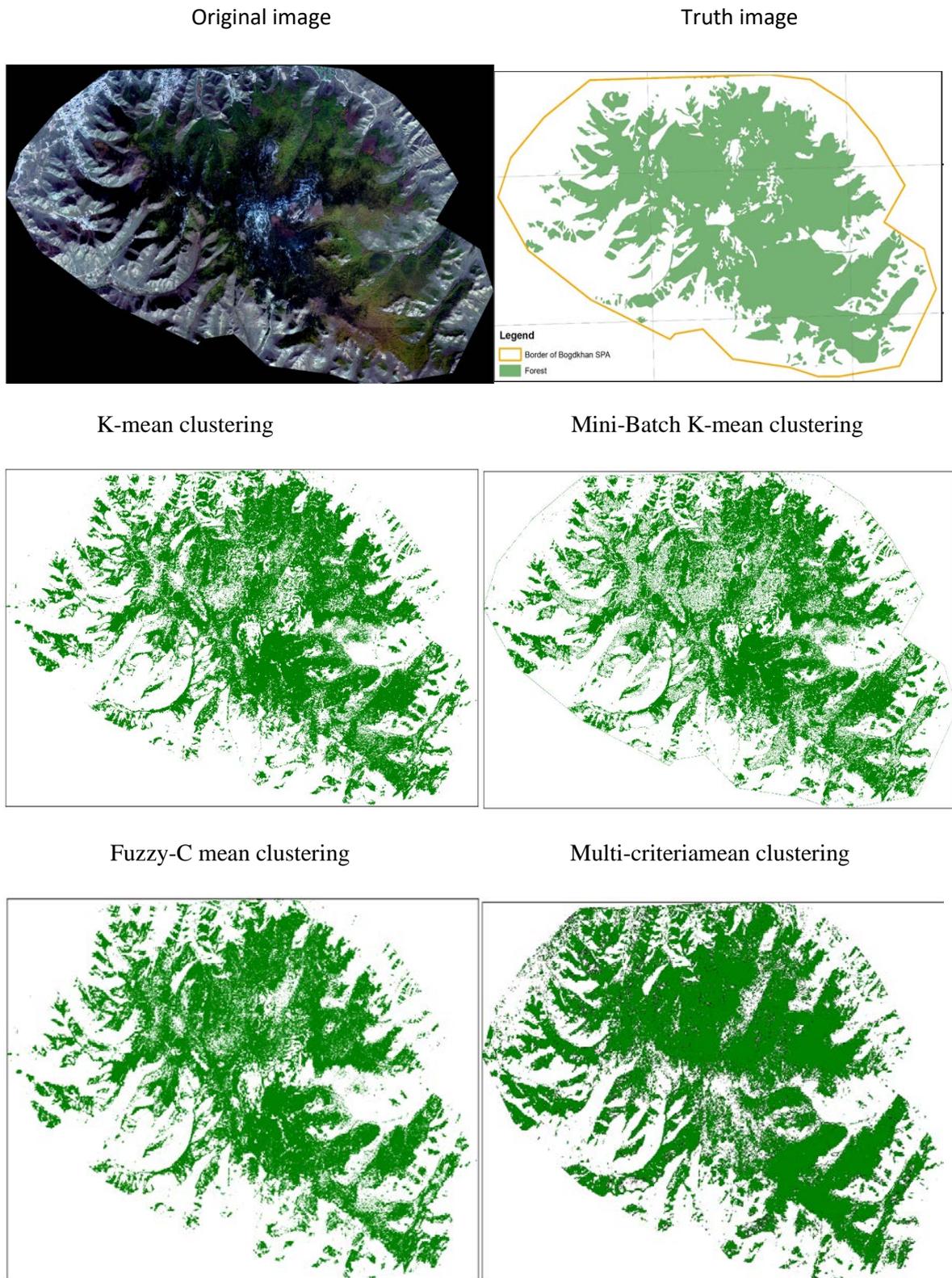
Original image

Truth image

K-mean clustering

Mini-Batch K-mean clustering

Fuzzy-C mean clustering

Multi-criteriamean clustering

Figure 2. Compared results.

Table 1.Comparative results.

Clustering methods	RMSE with truth image	SSIM with truth image	OA with truth image	IU with truth image
K-mean	0.413	0.606	0.71	0.67
Mini-Batch	0.4199	0.627	0.71	0.68
Fuzzy-C	0.4114	0.678	0.73	0.69
MCMC	0.4093	0.683	0.74	0.69

CONCLUSIONS

In this paper, we proposed a new clustering method called Multi-Criteria Mean Clustering (MCMC)**.** The method was applied to Sentinel-2B remote sensing images and compared with K-means**,** Mini-Batch K-means, and Fuzzy C-means using RMSE**,** SSIM**,** Overall Accuracy (OA)**,** and Intersection over Union (IU) as evaluation metrics.

The results in Table 1 demonstrate that MCMC effectively enhances clustering performance by balancing multiple objectives. By incorporating weighted criteria, the proposed method improves pixel-wise accuracy, structural consistency, and spatial overlap, making it a robust alternative to traditional unsupervised clustering approaches for forest mapping from Sentinel-2 imagery. The experiments were conducted using Python Jupyter Notebook, confirming that the multi-criteria optimization framework effectively enhances clustering performance for remote sensing image analysis.

Acknowledgements

Ethical approval

Ethical approval was not required for this study because it exclusively used publicly available data and did not involve human participants or identifiable personal information.

Source of funding

Conflict of interest

We have no conflict of interest to disclose. The authors declare that they have no conflicts of interest.

Authors contributions

The authors confirm contribution to the paper as follows: BD- supervised and reviews. DJ- writing original draft and visualization. MA- data curation. ER- methodology and editing. All authors reviewed the results and approved the final version of the article.

REFERENCES

1. Camps-Valls, G., "Machine learning in remote sensing data processing," IEEE International Workshop on Machine Learning for Signal Processing, pp. 1-6, 2009. https://doi.org/10.1109/MLSP.2009.5306233.

2. Sisodiya, N., Dube, N., Thakkar, P., "Next-generation artificial intelligence techniques for satellite data processing," In Artificial Intelligence Techniques for Satellite Image Analysis, Springer, Vol. 24, pp. 235-254, 2020. https://doi.org/10.1007/978-3-030-24178-0_11.

3. Han, J., Pei, M., Kamber, M., "Data Mining: Concepts and Techniques," Elsevier, New York, 2011.

4. X. Wu, X. Zhu, G. Q. Wu, and W. Ding, "Data mining with big data," IEEE Transaction on Knowledge and Data Engineering, vol. 26, no. 1, pp. 97-107, 2014. https://doi.org/10.1109/tkde.2013.109.

5. Jain, A., Murty, M., Flynn, P., "Data Clustering: A Review," CM Computing Surveys, Vol. 31, no. 3, pp. 264-323, 1999. https://doi.org/10.1145/331499.331504.

6. A. S. Shirkhorshidi, S. Aghabozorgi, T. Y. Wah, and T. Herawan, "Big Data Clustering: A Review," Lecture Notes in Computer Science, vol. 8583, pp. 707-720, 2014. https://doi.org/10.1007/978-3-319-09156-3_49.

7. Jain, A., Duin, R., Mao, J., "Statistical Pattern Recognition: A Review," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no.1, pp. 4-37, 2000. https://doi.org/10.1109/34.824819.

8. D. T. Nguyen, L. Chen, and C. K. Chan, "Clustering with Multi-viewpoint-Based Similarity Measure," IEEE Transactions on Knowledge and Data Engineering, vol. 24, no. 6, pp. 988-1001, 2012. https://doi.org/10.1109/tkde.2011.86.

9. Hamerly, G., Elkan, C., "Learning the K in K-means," In The Seventh Annual Conference on Neural Information Processing Systems, 2003.

10. C. Chen, L. Pau, and P. Wang, "Hand book of Pattern Recognition and Computer Vision , Eds., World Scientific, Singapore, pp. 3-32. R.Dubes, "Cluster analysis and related issue".

11. Coleman, G., Andrews, H., "Image Segmentation by Clustering," In Proceedings of IEEE, vol. 67, pp. 773-785, 1979. https://doi.org/10.1109/proc.1979.11327.

12. Jain, A., Dubes, R., "Algorithms for Clustering Data," Prentice Hall, New Jersey, USA, 1988.

13. Turi, R.H., "Clustering-Based Color Image Segmentation," PhD Thesis, Monash University, Australia, 2001.

14. Kaukoranta, T., Franti, P., Nevalainen. O., "A New Iterative Algorithm for VQ Codebook Generation," International Conference on Image Processing, pp. 589-593, 1998. https://doi.org/10.1109/icip.1998.723533.

15. Baek, S., Jeon, B., Lee. D., Sung. K., "Fast Clustering Algorithm for Vector Quantization," Electronics Letters, vol. 34, no. 2, pp. 151-152, 1998. https://doi.org/10.1049/el:19980217.

16. Xiang, Z., "Color Image Quantization by Minimizing the Maximum Inter-cluster Distance," ACM Transactions on Graphics, vol. 16, no. 3, pp. 260-276, 1997. https://doi.org/10.1145/256157.256159.

17. D. L. Olson and D. Delen, "Advanced Data Mining Techniques," Springer, 1st edition, 2008.

18. Judd, D., Mckinley, P., Jain, A., "Large-scale Parallel Data Clustering," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 20, no. 8, pp. 871-876, 1998. https://doi.org/10.1109/34.709614.

19. Bezdek, J., "A Convergence Theorem for the Fuzzy ISODATA Clustering Algorithms," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 2, pp. 1-8, 1980. https://doi.org/10.1109/tpami.1980.4766964.

20. Bezdek, J., "Pattern Recognition with Fuzzy Objective Function Algorithms," Plenum Press, 1981. https://doi.org/10.1007/978-1-4757-0450-1.

21. Amarsaikhan, D., Munkh-Erdene, A., Enkhjargal, D., Nyamjargal, E., Tsogzol, G., "Forest change study in Mongolia using optical and microwavers," First International Conference on Climate Change and Environment in Central and North-East Asia, October, 2022.

22. Nyamjargal, E., Amarsaikhan, D., Munkh-Erdene, A., Enkhjargal, D., Battsengel, V., Bolorchuluun, Ch., "Object-based classification of mixed forest types in Mongolia," Geocarto International, 35:14, 1615-1626, 2020. https://doi.org/10.1080/10106049.2019.1583775.