DEVELOPMENT OF A NOVEL MULTI-CRITERIA METHOD USING DEEP LEARNING AND OPTIMIZATION FOR IMAGE CLASSIFICATION

Tsolmonbayar.Sh[1], Bayanjargal.D[2*], Batchuluun.Ts[1], Tsolmon.R[3], Davaajargal.J[2], Selenge.M[1] and Bayanmunkh.N[4,5]

[1]School of Engineering and Applied Science, National University of Mongolia, Ulaanbaatar, Mongolia
[2]School of Information and Electronic Engineering Applied Matematics National University of Mongolia, Ulaanbaatar, Mongolia
[3]School of Art and Sciences Physics department, National University of Mongolia, Ulaanbaatar, Mongolia
[4]Centre for Policy Research and Analysis of Ulaanbaatar Municipality, Ulaanbaatar, Mongolia
[5]Laboratory of Geo-Virtual, Mongolian National University, Ulaanbaatar, Mongolia
*bayanjargal@num.edu.mn

Abstract: *Recent advancements in remote sensing image analysis have increasingly utilized deep learning models, resulting in notable improvements in classification accuracy and computational efficiency. Among these approaches, hybrid methods that combine deep learning with optimization techniques have shown superior performance over conventional single-model algorithms. In this study, we propose a novel classification algorithm called Multi-Criteria Mean Clustering (MCMC). This method integrates deep learning-based feature extraction with a multi-objective optimization framework, enabling it to better capture the diverse characteristics of high-dimensional and heterogeneous remote sensing data. By considering multiple criteria—such as spectral separability, spatial coherence, and class distribution—MCMC enhances clustering robustness and interpretability. The proposed method was applied to a case study in Dornod Province, Mongolia, a region along the Siberian forest boundary known for its complex land cover structure and ecological significance. We used Sentinel-2B multispectral imagery to perform land cover classification. To validate the classification performance, results from MCMC were compared against NDVI-based ground truth data. Correlation analysis revealed a 98% agreement between the MCMC outputs and the NDVI-derived reference map. Additionally, MCMC was benchmarked against two commonly used techniques: Mini-Batch K-Means, known for its scalability, and Random Forest, a widely adopted supervised classification method. Comparative results showed that MCMC either matched or exceeded the performance of these methods, particularly in terms of class boundary delineation and intra-class homogeneity. These findings demonstrate the potential of the MCMC approach in addressing the limitations of existing clustering and classification techniques, especially for complex and heterogeneous remote sensing datasets. The integration of deep learning with multi-criteria optimization provides a flexible and scalable framework for accurate land cover mapping. The method holds promise for future applications in environmental monitoring, land use analysis, and geospatial intelligence.*

Keywords: Multi-Criteria, deep learning, optimization, classification

Introduction

The development of classification methods in remote sensing dates back to the early days of aerial photography in the mid-20th century, when visual interpretation was the primary approach. Analysts manually identified land cover types based on color, tone, and texture. With the advent of satellite imagery in the 1970s, such as the Landsat program, digital classification techniques emerged, enabling automated analysis of multispectral data. Initially, unsupervised classification methods were widely used, grouping pixels into clusters based on statistical similarities without prior knowledge. Later, supervised classification techniques, such as maximum likelihood and minimum distance classifiers, became standard, relying on training data for known land cover types. From the 1990s onward, the integration of machine learning algorithms revolutionized classification. Techniques like decision trees, support vector machines (SVM), and random forests provided higher accuracy and flexibility in handling complex, high-dimensional datasets. More recently, deep learning approaches, including convolutional neural networks (CNNs), have shown promising results, particularly for high-resolution imagery and large-scale land cover mapping. During the 1990s, statistical and machine learning methods, such as decision trees and support vector machines, were introduced to improve classification accuracy in heterogeneous forest landscapes. The 2000s onward saw the integration of spectral indices like NDVI and enhanced multispectral data, allowing more precise discrimination of forest types and monitoring of forest health. Recently, optimization approaches combined with machine learning and deep learning techniques have emerged, providing more robust and scalable solutions for complex forestry applications.

Overall, the evolution of classification methods reflects a progression from manual interpretation to sophisticated automated approaches leveraging advanced computational techniques, enabling more precise and scalable analysis of Earth’s surface. Currently, researchers around the world are actively applying the latest advancements in remote sensing technology to various fields of research. Through the automatic interpretation of various remotely sensed images and advanced techniques based on knowledge, thematic results are obtained and utilized for various scientific purposes. A vast amount of remote sensing data, including optical, hyperspectral, and synthetic aperture radar (SAR) data, has been produced and is increasingly being utilized in various scientific domains. These data are particularly useful in environmental and natural resource management.

The diversity of objectives and the special characteristics of the data give rise to the use of a wide range of machine learning and signal processing algorithms (Camps-Valls, 2009). The statistical characterization of remote sensing images turns out to be difficult because of the pixel's high dimensionality, the presence of different kinds of noise sources and uncertainty, their inherent non-linear nature, and the high spatial and spectral redundancy (Camps-Valls, 2009). Machine learning has been successfully applied in remote sensing for classification, regression, clustering, coding, or source separation (Camps-Valls, 2009). Apart from the current automated techniques for satellite data processing, still, a large number of applications require manual intervention and the use of human intelligence for decision-making. Artificial intelligence and machine learning techniques are trying to address these issues with machine (computer) processes like humans (Sisodiya, Dube, & Thakkar, 2020).

Clustering, also known as cluster analysis, has become an important technique in machine learning used to discover the natural grouping of the observed data (Han, Pei, & Kamber, 2011). It is also a powerful tool for discovering patterns, organizing data, and gaining insights into the inherent structures within datasets. Clustering in remote sensing data is a common technique used to group similar pixels or objects in satellite or aerial imagery. This helps in identifying patterns, regions of interest, and distinguishing different land cover or land use types.

Data clustering is the process of identifying natural groupings or clusters within multidimensional data based on some similarity measure (e.g. Euclidean, Manhattan, and Minkowski distance) (Jain, Murty, & Flynn, 1999; Jain, Duin, & Mao, 2000). Clustering method is used in many applications, such as in pattern recognition (Hamerly & Elkan, 2003), image segmentation (Coleman & Andrews, 1979; Jain & Dubes, 1988; Turi, 2001), vector and color image quantization (Kaukoranta, Fränti, & Nevalainen, 1998; Baek, Jeon, Lee, & Sung, 1998; Xiang, 1997), and data mining (Judd, McKinley, & Jain, 1998), etc. The choice of clustering method depends on the characteristics of the remote sensing data, such as the number of classes, distribution of classes, and the spatial and spectral properties of the data

Objective

In this paper, we propose a Multi-criteria Mean Clustering (MCMC) method that integrates multiple optimization criteria into the clustering process, allowing for improved accuracy and robustness compared to tradtional approaches. The method was applied to Sentinel-2B multispectral imagery of a forested area in Dornod province, Mongolia, a region with complex land cover patterns. Experimental results show that MCMC outperforms widely used clustering algorithms such as K-means, Mini-Batch K-means, providing higher clustering quality and more reliable classification of remote sensing data.

Methodology

This section introduces the methodological framework used in this study. We first present the proposed Multi-Criteria Mean Clustering (MCMC) method, which generalizes traditional clustering by incorporating multiple optimization objectives. To evaluate its effectiveness, we compare its performance with two baseline methods: Mini-Batch K-Means and Random Forest**.** An optimization approach was applied in the development of the MCMC model. In mathematics, an optimization problem involves finding the best solution according to a specific criterion, typically by maximizing or minimizing an objective function. Despite its potential, there have been relatively few studies that combine optimization techniques with machine learning for land cover classification using remote sensing data.

Multi-Criteria Mean Clustering (MCMC): Traditional clustering algorithms such as K-Means aim to minimize a single objective function, usually the intra-cluster variance. However, remote sensing data often requires the simultaneous consideration of multiple criteria, such as spatial smoothness, spectral separability, and inter-class dissimilarity. To address this, we propose a multi-objective formulation of the clustering problem. (Deb, 2014).

Let X= $\{x_1, x_2,\ldots, x_n\}$ be a dataset in R^n, and let C=$\{c_1,c_2,\ldots,c_k\}$ denote the centroids of k clusters. Each data point x_i is assigned a membership weight w_{ij}, and the distance function d (x_i, c_j) may represent various similarity measures (Euclidean, Manhattan, etc).

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

$$J(C) = \sum_{i=1}^{n}\sum_{j=1}^{k} w_{ij}\ \mathrm{d}(\mathrm{x}_i, \mathrm{c}_j) \longrightarrow \min \tag{1}$$

where c_j is the centroid of the j-th cluster, w_{ij} is a weight indicating the association of x_i with cluster c_j, and $d(x_i,c_j)$ is the distance between x_i and the centroid c_j.

We introduce the following multi-criteria approach in K-mean clustering, that is:

$$\varphi_j(c) = \sum_{i=1}^{n} \left\| \mathrm{x}_i - \mathrm{c}_j \right\|^2, j = 1,2,\ldots,k, \qquad c \in \mathrm{C} \tag{2}$$

and

$$\min_{c \in C} F(c) = \sum_{j=1}^{k} \alpha_j \, \varphi_j(c) \tag{3}$$

where $\alpha_j > 0$, j=1,…, k is weight of the multi-criteria. In particular case, if we set (3) in α_j = 1, then the problem become K-mean clustering.

Advantage of multi-criteria optimization approach to clustering is that by choosing appreciate weights α_j we can improve quality of clustering.

This allows dynamic control over the importance of each clustering criterion. Notably:

- If α_j=1 for all j, this reduces to the classical K-Means objective.
- If $\alpha_j=\frac{1}{|M|}\sum_{i=1}^{n} r_{ij}$, it resembles Mini-Batch K-Means.
- Fuzzy C-Means and other soft clustering methods also become special cases under appropriate weights.

This flexible formulation allows MCMC to adapt to specific application requirements and data characteristics, making it highly suitable for remote sensing analysis.

Mini-Batch K-Means: Mini-Batch K-Means is a variant of the K-Means algorithm designed for scalability and computational efficiency. Instead of using the entire dataset for updating centroids, it selects random mini-batches of data points at each iteration. This reduces memory usage and improves processing speed, particularly for large remote sensing datasets. (Sculley, 2010). Mini-Batch K-means is a variant of the traditional K-means clustering algorithm that employs a stochastic optimization approach by using randomly selected subsets (mini-batches) of the data to update the cluster centroids iteratively. The objective function for Mini-Batch K Means is a stochastic approximation of the original K-Means objective. The objective function for Mini-Batch K-Means is defined as follows:

$$J(C) = \frac{1}{|M|} \sum_{i \in M} \sum_{j=1}^{k} r_{ij} \|x_i - c_j\|^2 \longrightarrow \min \tag{4}$$

Here M is the randomly selected mini-batch of data points, M is the size of the mini-batch. r_{ij} is indicator variable defined as following equation

$$r_{ij} = \begin{cases} 1 & xi \quad is\ assigned\ to\ cluster \\ 0 & otherwise \end{cases}$$

While clustering methods are effective for exploratory analysis and data grouping, supervised classification methods are often preferred when labelled data are available. One of the most powerful and widely used supervised learning algorithms in remote sensing is the Random Forest (RF) classifier. Random Forest is an ensemble learning method that builds multiple decision trees using randomly selected subsets of the training data and features. The final prediction is made by aggregating the outputs of all individual trees, typically via majority voting in classification tasks. This ensemble strategy improves generalization and reduces over fitting, making Random Forest particularly suitable for high-dimensional, noisy, and complex remote sensing datasets.

The advantages of Random Forest include:

- High classification accuracy even with limited training samples
- Ability to handle nonlinear relationships between features
- Tolerance to noisy and missing data
- Capability to estimate variable importance for feature selection

Random Forest has been successfully applied in various remote sensing tasks such as land use/land cover classification, vegetation type mapping, soil classification, and change detection. Its robustness and ease of use make it a preferred choice among researchers and practitioners working with satellite and aerial imagery. In summary, both unsupervised clustering techniques like Mini-Batch K-means and supervised classifiers such as Random Forest offer complementary strengths in the analysis of remote sensing data. The appropriate method should be selected based on the specific application goals, data characteristics, and availability of labelled training samples. Despite being computationally efficient, this method may still suffer from sensitivity to initialization and inability to incorporate multiple objectives.

Random Forest Classification: Random Forest is a robust ensemble learning method widely used for supervised classification tasks. It constructs multiple decision trees during training and aggregates their outputs to make final predictions. In the context of remote sensing, Random Forest is effective in handling high-dimensional data and heterogeneous feature sets (spectral bands, vegetation indices, texture measures, etc.) (Breiman, 2001).

Key characteristics of Random Forest:

- Handles non-linear class boundaries well.
- Provides internal feature importance ranking.
- Resilient to over fitting, especially with high-dimensional input.

In this study, Random Forest is used as a supervised benchmark to evaluate the performance of the unsupervised MCMC and Mini-Batch K-Means clustering results. Label data obtained from reference land cover maps are used to train and validate the classifier. The Multi-Criteria Mean Clustering (MCMC) method is introduced as a generalization of conventional clustering techniques through a multi-objective optimization framework. It enables the integration of multiple data characteristics, improving clustering accuracy and interpretability. The method is compared against Mini-Batch K-Means, which is efficient but limited in scope, and Random Forest, which provides a supervised benchmark. The comparative analysis highlights the strengths and applicability of MCMC for complex remote sensing tasks. (Hastie, Tibshirani, & Friedman, 2009).

Study area

The research area is located in Dornod Province, which shares its northern border with the Russian Federation. Geographically, Dornod lies at the intersection of Mongolia's taiga, forest-steppe, and steppe zones. The northern part of the province is predominantly characterized by forested mountain ranges, consisting mainly of coniferous and mixed forests that represent the southern fringes of the Siberian taiga. These forested highlands gradually transition southward into the arid steppe plains of the Central Mongolian Plateau. The forest ecosystems in the northern region are of particular ecological significance due to their biodiversity, water resources, and soil stability. They also play a crucial role in maintaining regional ecological balance and climatic regulation. (Figure 1).

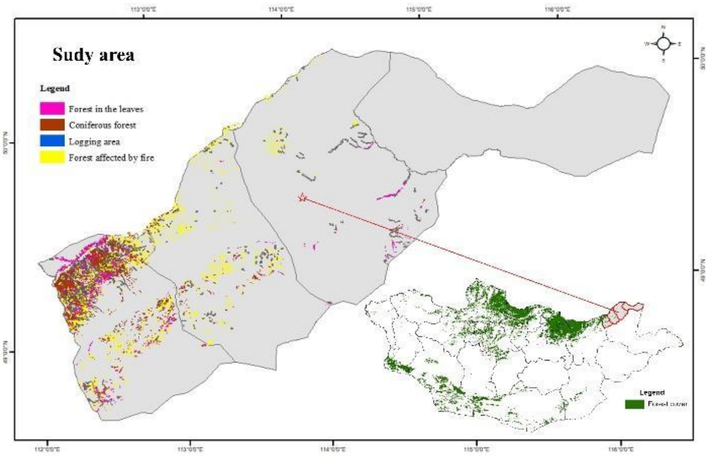


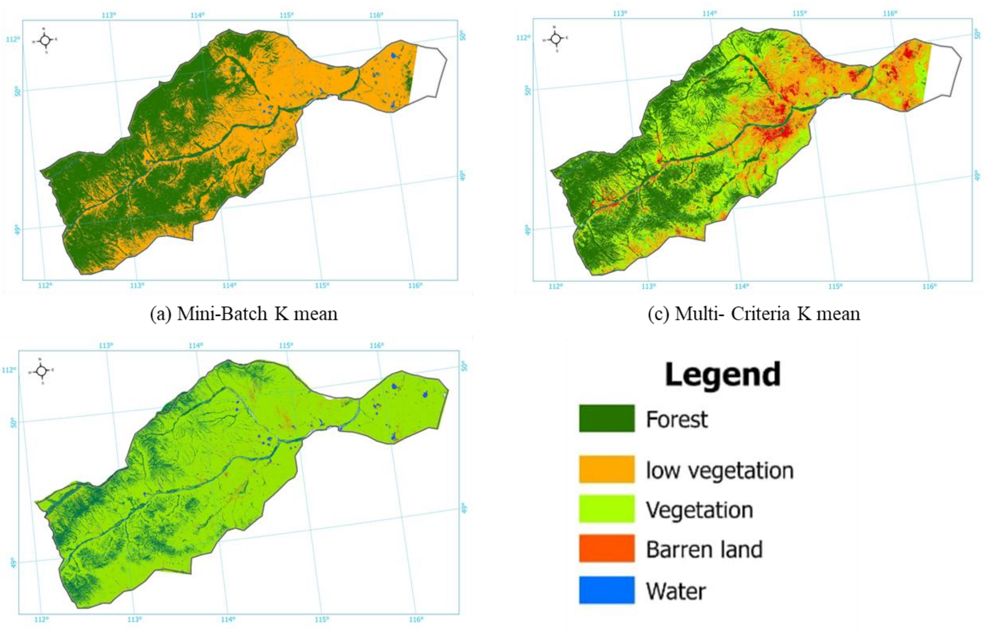
Figure 1: Location map: (a) Distribution of natural zones of Mongolia; (b) Dornod Province showing Dashbalbar, Chuluunkhoroot, Bayandun, and Bayan-Uul soums.

Satellite data

In this study, satellite imagery from the Sentinel-2 mission of the European Space Agency (ESA) Copernicus program was employed. Sentinel-2 consists of two satellites (Sentinel-2A and Sentinel-2B) carrying high-resolution multispectral sensors designed for land monitoring and environmental studies. The Multispectral Instrument (MSI) onboard Sentinel-2 provides data in 13 spectral bands with spatial resolutions of 10 m, 20 m, and 60 m. This capability makes Sentinel-2 suitable for identifying different land cover types, including forests, grasslands, croplands, and water bodies. The revisit frequency is approximately five days, enabling the monitoring of seasonal and climatic dynamics.

Accuracy Assessment and NDVI-Based Validation: To assess classification accuracy and reliability, 250 randomly distributed sample points were selected across the study area. The classification outputs from Multi-Criteria K-Means, Mini-Batch K-Means, and Random Forest methods were extracted at each sample point and compared against ground truth values derived from NDVI-based classification. The Normalized Difference Vegetation Index (NDVI) is one of the most widely used remote sensing indices for monitoring vegetation and classifying land cover. NDVI measures the difference between near-infrared (NIR) and red-light reflectance from vegetation, providing a quantitative estimate of vegetation health, density, and cover.
NDVI values range from -1 to 1, where higher values indicate denser and healthier vegetation, while lower or negative values correspond to barren land, water, or urban areas. High-resolution sensors and global datasets (MODIS, Sentinel-2) have allowed large-scale NDVI-based mapping with improved temporal frequency. NDVI has been combined with machine learning and optimization techniques to enhance land cover classification accuracy, particularly in heterogeneous landscapes. Overall, NDVI remains a cornerstone in remote sensing-based land cover classification due to its simplicity, robustness, and strong correlation with vegetation properties.

The comparison results demonstrate that the Multi-Criteria K-Means method more accurately delineates the boundary between forest and dense vegetation classes, exhibiting higher class separability and stronger spatial coherence. This highlights the advantage of the multi-objective optimization approach, which effectively captures transitional and mixed-class zones.
Furthermore, the strong correspondence between the Multi-Criteria K-Means classification and NDVI-based ground truth indicates the method’s robustness and reliability in reflecting real-world land cover characteristics.



(a) Mini-Batch K mean (c) Multi- Criteria K mean

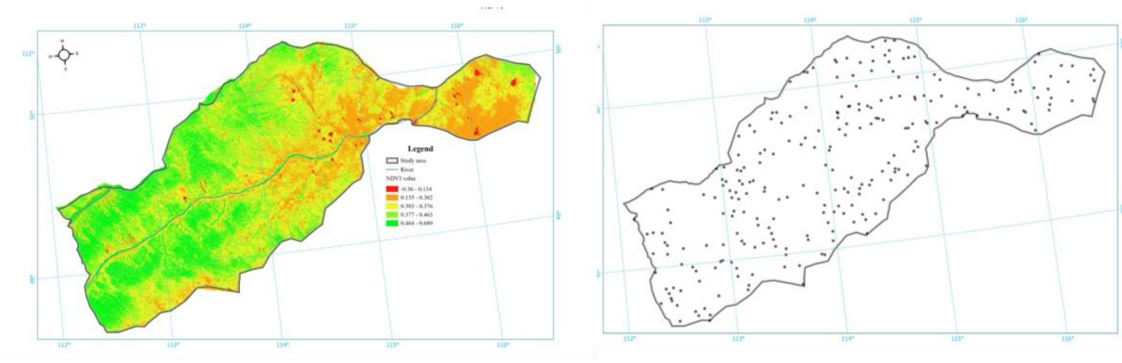
(b) Random forest

Figure 2: Compared result

(a)Mini-Batch K mean, (b) Random forest, (c) Multi- Criteria K mean

The land cover classification of the study area was processed using three different methods, and the results were analyzed for each of the six distinct land cover classes such as forest, low vegetation, Baren Land and water.

Figure 2 (a) shows the results obtained from the Multi-Criteria K-Means method. The classification map reveals a low distinction between forest and dense vegetation classes, with some vegetation classes being incorporated into the forest category. In Figure 2 (b), the opposite trend is observed, where the forest area decreases and the vegetation class expands, indicating a reversed pattern in the differentiation between forest and vegetation. Figure 2 (c) illustrates the land cover classification using the Multi-Criteria K-Means algorithm, where clear boundaries and high separability between each class are evident, demonstrating improved class discrimination.



(a) NDVI classification map (b) Randomly distributed sampling points

Figure 3: NDVI classification map

To assess the accuracy of the classification results produced by each method, a total of 250 sample points were randomly distributed across the study area. At each point, the class values obtained from the different classification methods were extracted and compared against the NDVI-based classification values.

Table 1. Matrix results

	NDVI	**Multi-Criteria K mean**	**MiniBatch clustered**	**Random porest**
NDVI	1			
Multi- Criteria K mean	0.9871615	1		
Mini-Batch K-mean	0.9667798	0.915098	1	
Random forest	0.8440049	0.747847	0.947952572	1

To validate the classification accuracy, the land cover map generated in Figure 3 was compared with the NDVI-based classification map shown in Figure 4. Using correlation matrix analysis, the highest agreement was observed between the NDVI classification and the Multi-Criteria K-Means (MCMC) method, with a correlation of 98%. Mini-Batch K-Means followed closely with 96% agreement, while the Random Forest algorithm showed the lowest correlation at 84%. These results indicate that the MCMC method provides more accurate classification performance, particularly in datasets characterized by high spectral and spatial variability. Furthermore, a strong agreement of 96% was also observed between the Multi-Criteria K-Means and Mini-Batch K-Means methods, suggesting that both approaches share a similar clustering logic and are capable of capturing structural patterns within the data (Table 1).

Results and Discussion

The classification of land cover using remote sensing plays a critical role in environmental monitoring and resource management. The results of this study demonstrate a strong agreement between the proposed methodology, other classification methods, and NDVI-based analysis. Integrating optimization techniques with machine learning provides a powerful and effective tool for improving the accuracy and efficiency of land cover classification. The land cover classification of the study area was conducted using three different methods, with the results analyzed across six distinct land cover classes. This section presents a detailed comparative analysis of the classification outcomes obtained from the proposed Multi-Criteria K-Means (MCMC) algorithm alongside baseline approaches, including Mini-Batch K-Means and Random Forest. The evaluation focuses on differences in classification structure, accuracy, and each method's capability to effectively capture the complex spectral and spatial variability inherent in remote sensing datasets. Multi-Criteria K-Means Clustering: The classification results of the proposed Multi-Criteria K-Means algorithm for land cover are illustrated in the figures below. In Figure (a), the boundary between forest and dense vegetation classes appears indistinct, with some vegetated areas being assigned to the forest class. Conversely, in Figure (b), the opposite trend is observed where forested areas are included in the vegetation class, indicating reduced separability between classes. This suggests some overlap and mixing in the classification results. Figure (c) presents the final land cover classification obtained using the specially developed Multi-Criteria K-Means approach, showing clearer and better-defined class boundaries. The separability between forest, vegetation, and other land cover classes is significantly improved, facilitating easier interpretation and more precise mapping. In this study, we proposed a novel clustering approach called Multi-Criteria Mean Clustering (MCMC) to address the limitations of conventional single-objective clustering algorithms when applied to high-dimensional and complex remote sensing data. Unlike traditional techniques such as K-Means or Mini-Batch K-Means, the MCMC algorithm incorporates multiple optimization objectives into the clustering process, allowing for a more nuanced analysis of spatial heterogeneity, spectral variability, and inter-class separability. To evaluate the performance and practical applicability of the proposed method, we conducted experiments using Sentinel-2B multispectral imagery of a forested area in Dornod province, Mongolia. This study area was selected due to its diverse land cover characteristics, which pose significant challenges for standard clustering and classification techniques. The results of MCMC were compared against two baseline methods—Mini-Batch K-Means and Random Forest—using both visual interpretation and a correlation matrix derived from ground-truth NDVI classification data. Numerous remote sensing-based classification methods have been developed for forestry applications, ranging from traditional pixel-based approaches to advanced machine learning techniques. These methods are used to map forest types, monitor vegetation health, estimate biomass, and detect changes in forest cover over time. The integration of spectral indices, such as NDVI, with classification algorithms has further improved the accuracy of forest mapping. In recent years, combining optimization techniques with machine learning has emerged as a promising approach to enhance classification performance, particularly in complex and heterogeneous forest landscapes. In remote sensing and geographic studies, land cover classes for standardization refer to a set of predefined categories used to consistently classify the Earth's surface. Standardization ensures that land cover mapping is comparable across regions, studies, and time periods. The comparison revealed that the MCMC method achieved higher classification accuracy, improved class boundary separation, and better agreement with NDVI-derived reference data, with a correlation coefficient of 98%, outperforming both Mini-Batch K-Means (96%) and Random Forest (84%). These findings demonstrate the potential of MCMC as a robust, unsupervised technique for land cover classification in complex remote sensing tasks.

All model simulations, clustering operations, and result analyses were conducted in the Python Jupyter Notebook environment, ensuring a reproducible and scalable workflow. Overall, the proposed MCMC algorithm offers a flexible and effective alternative for clustering heterogeneous remote sensing datasets and can be extended to other domains requiring multi-objective optimization. The integration of deep learning with multi-criteria optimization offers a flexible and scalable framework for accurate land cover mapping. A limitation of this study is the need for additional ground truth data collection and validation to further improve accuracy. Moreover, applying this methodology to other regions with similar climates worldwide is necessary to assess its generalizability. It is also important to consider that land cover classes may differ in other regions, which could affect classification performance.

Reduces ambiguity in classification. The method can be used to support automated classification algorithms for the Common Standard Land Cover Classes Standardized systems, such as the FAO Land Cover Classification System (LCCS) or CORINE Land Cover and Enables global land cover monitoring and modeling of environmental processes.

Acknowledgements

This work was conducted under project P2019-3753, supported by the National University

of Mongolia.

References

Camps-Valls, G. (2009). Machine learning in remote sensing data processing. In *IEEE International Workshop on Machine Learning for Signal Processing* (pp. 1–6). IEEE.

Sisodiya, N., Dube, N., & Thakkar, P. (2020). Next-generation artificial intelligence techniques for satellite data processing. In *Artificial Intelligence Techniques for Satellite Image Analysis* (Vol. 24, pp. 235–254). Springer. https://doi.org/10.1007/978-3-662-49096-9

Han, J., Pei, M., & Kamber, M. (2011). *Data mining: Concepts and techniques*. Elsevier.

Jain, A. K., Murty, M. N., & Flynn, P. J. (1999). Data clustering: A review. *ACM Computing Surveys, 31*(3), 264–323.

Jain, A. K., Duin, R. P. W., & Mao, J. (2000). Statistical pattern recognition: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 22*(1), 4–37.

Hamerly, G., & Elkan, C. (2003). Learning the K in K-means. In *Proceedings of the 17th International Conference on Neural Information Processing Systems*.

Coleman, G. B., & Andrews, H. C. (1979). Image segmentation by clustering. *Proceedings of the IEEE, 67*(5), 773–785.

Jain, A. K., Dubes, R. C. (1988). *Algorithms for clustering data*. Prentice Hall.

Turi, R. H. (2001). *Clustering-based colour image segmentation* (Doctoral dissertation, Monash University, Australia).

Kaukoranta, T., Fränti, P., & Nevalainen, O. (1998). A new iterative algorithm for VQ codebook generation. In *Proceedings of the IEEE International Conference on Image Processing* (pp. 589–593).

Baek, S., Jeon, B., Lee, D., & Sung, K. (1998). Fast clustering algorithm for vector quantization. *Electronics Letters, 34*(2), 151–152. https://doi.org/[DOI]

Xiang, Z. (1997). Color image quantization by minimizing the maximum inter-cluster distance. *ACM Transactions on Graphics, 16*(3), 260–276.

Judd, D., Mckinley, P., & Jain, A. (1998). Large-scale parallel data clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 20*(8), 871–876.

Bezdek, J. C. (1980). A convergence theorem for the fuzzy ISODATA clustering algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 2*, 1–8.

Deb, K. (2014). *Multi-objective optimization*. Springer. https://doi.org/10.1007/978-3-642-40860-2

Sculley, D. (2010). Web-scale k-means clustering. In *Proceedings of the 19th International Conference on World Wide Web* (pp. 1177–1178). ACM. https://doi.org/10.1145/1772690.1772862

Breiman, L. (2001). *Random forests*. Machine Learning, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer. https://doi.org/10.1007/978-0-387-84858-7

European Space Agency. (2015). *Sentinel-2 user handbook* (ESA Standard Document ESA/PHC/1410). https://sentinels.copernicus.eu/documents/247904/685211/Sentinel-2_User_Handbook