Detection of Changes in Land Cover in a Mining Area of Mexico Using Remote Sensing and Machine Learning

Saúl Dávila-Cisneros *, Ana G. Castañeda-Miranda *, Carlos Francisco Bautista-Capetillo, Erick Dante Mattos-Villarroel, Víktor Iván Rodríguez-Abdalá, Cruz Octavio Robles Rovelo

Posted Date: 26 March 2026

doi: 10.20944/preprints202603.2049.v1

Keywords: mining; land cover; remote sensing; machine learning; environmental monitoring

 doi:10.20944/preprints202603.2049.v1

Article

Detection of Changes in Land Cover in a Mining Area of Mexico Using Remote Sensing and Machine Learning

Saúl Dávila-Cisneros [1,*], Ana G. Castañeda-Miranda [2,*], Carlos Francisco Bautista-Capetillo [3], Erick Dante Mattos-Villarroel [3], Víktor Iván Rodríguez-Abdalá [2] and Cruz Octavio Robles Rovelo [3]

[1] Unidad Académica de Ingeniería I, Universidad Autónoma de Zacatecas, Av. Ramón López Velarde No. 801, Col. Centro, C.P. 98060, Zacatecas, Zacatecas, México

[2] Posgrado en Ingeniería para la Innovación Tecnológica, Unidad Académica de Ingeniería Eléctrica, Universidad Autónoma de Zacatecas, C.P. 98160, Zacatecas, Mexico

[3] Doctorado en Ciencias de la Ingeniería, Universidad Autónoma de Zacatecas, Campus UAZ Siglo XXI, Carretera Zacatecas-Guadalajara Km. 6, Ejido La Escondida C.P. 98160, Mexico

* Correspondence: author: davila84@uaz.edu.mx (S.D.-C.); agmiranda@uaz.edu.mx (A.G.C.-M.)

Highlights

What are the main findings?

- Tests performed on eleven classification algorithms showed that the Spectral Angle Mapping (SAM) algorithm achieved the best results with an accuracy of 85.16% and a Kappa coefficient of 0.79.
- Measurements of changes in land cover revealed increases of +556.83 Ha in water bodies, +1729.35 Ha in mining cover, +2.61 Ha in infrastructure, and +1488.15 Ha in bare soil. Losses of -2372.49 Ha of soil, -1444.59 Ha of shrubland, and -9.45 Ha of vegetation were also detected.

What are the implications of the main findings?

- It is of utmost importance to carefully select the input data, the classification algorithm, the training data, and the validation data.
- It is possible to detect changes in land cover in bodies of water, mining cover, infrastructure, bare soil, soil, scrubland and vegetation; using satellite images and classification algorithms in mining areas of Mexico.

Abstract

Mining generates various alterations to the environment, affecting flora, fauna, morphology, and soil. To contribute to solving this problem, this study measured land cover (LC) changes induced by open-pit mining in Zacatecas, Mexico. Remote sensing techniques were applied using multitemporal Landsat 5 and 8 satellite imagery, along with supervised classification, to detect land cover variations. Tests performed on eleven classification algorithms showed that the Spectral Angle Mapper (SAM) obtained the best results, with an accuracy of 85.16% and a Kappa coefficient of 0.79. Measurements of changes in land cover revealed an increase in the surface area of water bodies of +556.83 Ha, in mining cover of +1729.35 Ha, in infrastructure of +2.61 Ha, and of bare soil of +1488.15 Ha, and a loss of soil of -2372.49 Ha, of scrubland of -1444.59 Ha, and of vegetation of -9.45 Ha. The use of supervised classification for multi-temporal satellite imagery allowed for the measurement of alterations to land cover. These alterations highlight the need for sustainable management strategies, environmental restoration, and the importance of continued monitoring for informed decision-making. It is recommended to explore new categories and techniques, such as deep learning, to improve the accuracy of land cover classification.

 doi:10.20944/preprints202603.2049.v1

Keywords: mining; land cover; remote sensing; machine learning; environmental monitoring

1. Introduction

Mining operations generate alterations to the flora, fauna, morphology, and hydrology, and cause soil erosion [1,2]. Mining areas undergo continuous transformation due to excavation, waste disposal, and land reclamation [3,4]. Therefore, monitoring Land Cover (LC) in mining regions is of utmost importance. An accurate land cover classification facilitates Environmental Impact Assessment (EIA), monitoring of Land Use and Land Cover (LULC) changes, planning, rehabilitation assessment, and regulatory compliance [5,6].

Currently, land cover monitoring is conducted using remote sensing, which is the most efficient tool for detecting and quantifying land cover changes over time [7]. Some studies highlight that LC change in mining areas reflects long-term human-environment interactions and is a key indicator of ecological disturbance and recovery [8,9].

Due to the usefulness of LC classification maps in mining areas, numerous studies have been conducted worldwide using different methodologies, techniques, and algorithms. For example, some researchers have proposed a new time-series LC classification method that uses spectral, temporal, and spatial information in three regions of the USA, demonstrating greater classification accuracy, both overall and by class, and higher kappa index values [10]. Landsat Image Classification Algorithms (LICA) have also been proposed to automatically detect LC information, perform multi-temporal and multi-sensor analysis using Google Earth Engine (GEE) and cloud processing; in the Apulia region (southern Italy) based on the sequential application of two new indices: SwirTirRed and the SwiRed index [7]. Other authors have proposed a new carbonate index based on WorldView-3, using three optimization methods: a genetic algorithm (GA), k-fold cross-validation (CV), and particle swarm optimization (PSO), to obtain the optimization parameters of the Support Vector Machine (SVM); achieving a significant increase in the overall classification accuracy of fine LC by 8.40% [11]. In another study, changes in LC related to a coal mine in Bangladesh were assessed using remote sensing techniques and geographic information systems (GIS). These studies employed temporal Landsat satellite imagery and a supervised classification maximum likelihood matrix with an overall accuracy of 70% and a kappa coefficient of 0.60 [12]. Historical impacts of mining activities on biophysical surface characteristics have also been evaluated, and predictions of future changes in vegetation cover pattern and Land Surface Temperature (LST) have been made using Landsat satellite imagery and meteorological data from Iran, Canada, India, and Germany. The homogeneity distance classification algorithm (HDCA) and a CA-Markov model were used to predict future LC changes [13]. Another study performed a LC classification, estimating surface temperature, soil moisture, atmospheric water vapor density, eco-environmental evolution, and cumulative ecological effects (CEEs) of mining operations, using remote sensing and mining data in Inner Mongolia, China [14]. Fine land cover classification (FLCC) has also been performed in open-pit mining areas using convolutional neural networks (CNN) with a multimodal, multi-stream, multi-scale kernel-based model and remote sensing data in Wuhan, China [15]. A new CNN model called ResCapsNet has been created for LC classification in the city of Wuhan and Xinjiang province, China. obtaining average overall accuracies of 98.45% and 82.80% for the Wuhan study area, and 92.82% and 70.88% for the Xinjiang study area [16]. The principal component analysis method and the Remote Sensing Ecological Index (RSEI), integrated with four indices of greenness, humidity, dryness, and heat, have been used to assess and study the ecological disturbance of the mining area [17]. New models for the classification of the LC have been proposed using three stages: (1) execution of a multimodal feature learning method of three flows and post-melt; (2) integration of deep separable asymmetric convolution blocks and (3) use of a bidirectional long-term memory (BiLTM) network to learn contextual features across channels more deeply with an overall accuracy of 98.65% ± 0.05% in Wuhan, China [18]. Other authors are using the random forest algorithm and cloud processing with GEE to achieve time-series LC classification at different scales based on classification sampling points

 doi:10.20944/preprints202603.2049.v1

of optical, radar, and digital elevation models (DEMs); with limited overall improvement [19]. Landsat satellite image time series have been extracted from GEE to monitor the effects of mining on the local forest (LC) around the city of Ouagadougou, Burkina Faso, in West Africa. The LandTrendr algorithm and the Normalized Difference Vegetation Index (NDVI) were applied to map annual vegetation disturbance and recovery, achieving over 90% accuracy across 257 quarries [20]. Changes in the local LULC have been mapped using Landsat imagery, the random forest model, and GEE, yielding a classification with an overall accuracy and a kappa coefficient of 82.47% and 0.78, respectively [21]. Other authors have obtained accuracies of 92–97% using the same methodology [22]. A new Transformer-based semantic segmentation model, called Segmentation for Mine (SegMine), has also been proposed, improving the network's ability to obtain local spatial detail information [23]. A methodological framework for mapping open-pit mines in complex areas in Shanxi, China, using Landsat and GEE imagery, has been presented. The workflow includes preprocessing, classification, and postprocessing, achieving average accuracies of 79.4% for the producer and 83.2% for the user [24]. Multitemporal machine learning using Smile Random Forest (SRF) and Smile Gradient Tree Boost (SGTB) models has been employed to analyze and predict mining-induced LULC changes in Itagunmodi, Osun State, Nigeria, using multitemporal Landsat satellite imagery and ancillary data such as DEMs, road networks, and precipitation data, with overall accuracies of 82%, 81%, and 85% for 2002, 2016, and 2024, respectively [25]. The growth of mining activities in Benguerir, Morocco, has been studied using remote sensing and ancillary data by measuring spatiotemporal LULC changes using the maximum likelihood method [26]. Others have used small unmanned aerial vehicles (UAVs) equipped with multispectral sensors for LC classification and vegetation monitoring [27]. Additionally, the impact of semi-mechanized and artisanal mining activities on LC has been assessed using remote sensing data and photogrammetric analysis in the locality of Mbale, northern Cameroon. This assessment employed a maximum likelihood classification algorithm combined with field studies, based on Sentinel-2 imagery, NDVI, Normalized Difference Water Index (NDWI), Brightness Index (BI), and Soil Crust Index (SCI) to evaluate changes in vegetation, bare soil, water bodies, and exploited areas [28].

As we can see from the background review, many studies have been conducted on land cover and land use classification using remote sensing, GIS, and artificial intelligence methodologies. A wide range of algorithms and their variations have been used, including Genetic Algorithms, LandTrendr, Maximum Likelihood (MLi), Random Forest (RF), Smile Gradient Tree Boost (SGTB), Support Vector Machine (SVM), Homogeneity Distance Classification Algorithm (HDCA), OB-SVM Classifier, and Graph Segmentation. Various spectral indices have been employed, such as NDVI, NDWI, BI, SCI, Carbonate Index, SwirTirRed Index, SwiRed Index, and indices of greenness, humidity, dryness, heat, and the RSEI. Satellite image classification has been performed primarily by pixels and by objects. Various models and variants of artificial and convolutional neural networks have been used. Remote sensing has also been combined with photogrammetry using UAVs and GIS.

However, monitoring changes in land cover due to mining activities still faces certain limitations. One of the most significant limitations is the spatial resolution of some satellite images and the lack of detailed information on surface morphology, hindering the detection and mapping of artisanal and small-scale mining [27,29]. Other limitations include the fact that time-series classification methods for satellite images do not fully utilize the available data to extract the effective features that allow for the identification of different land cover types [10]. Tracking the dynamics of LULC is difficult in large-scale underground mining areas with wide LULC heterogeneity and a history of multiple disturbances [8]; the spectral indices developed are suitable for extracting a specific class, but not for fully classifying the study area [7]. One area of opportunity among all the research is the scarcity of comparative studies between the different algorithms and techniques for LC classification. Therefore, it would be beneficial to develop comparative studies of the different methodologies, techniques, and algorithms used. On the other hand, these studies have only been carried out in a few countries worldwide, such as West Africa, Austria, Bangladesh, Brazil, Cameroon, the Philippines, Italy, Morocco, Nigeria, the USA, and Zambia, with each country

 doi:10.20944/preprints202603.2049.v1

accounting for 5% of the studies, while China accounts for 50% of the total research. Therefore, it would be beneficial to implement studies of this type in Mexico.

The main challenges in LC classification in mining areas are mining landscapes present classification challenges due to their unique physical characteristics. Some of these characteristics include complex three-dimensional terrain that affects lighting and shadows; similarity between barren soil, waste piles, and exposed rock; and rapid temporal changes that require frequent monitoring.

Therefore, we have set the general objective of detecting and measuring land cover changes in a mining area of Mexico using supervised classification algorithms such as: Binary Encoding (BE), Parallelepiped (P), Minimum Distance (MD), MLi, Spectral Angle Mapper (SAM), Decision Tree (DT), K-Nearest Neighbours (K-NN), Maximum Entropy (ME), Normal Bayes (NB), RF and SVM, and comparing the different algorithms.

The main contribution of this article is the application of these technologies in Mexico, as similar studies have not yet been conducted, and Mexico has significant mining activity. Therefore, it is essential to monitor changes in land cover in mining areas. Another contribution of this research is the comparison between the different algorithms used, both traditional and machine learning (ML).

This article is divided into several sections: the introduction describes the problem studied, its importance, defines the current state of the research field, states the main objective, and highlights the principal contributions of this work. The materials and methods section describe the data and instruments used, as well as the methodology applied. The results section presents the experimental results and their interpretation. The discussion section discusses the results and how they can be interpreted considering previous studies. Finally, the conclusions section makes recommendations for future research, analyzes the implications of the research, and explains how the objectives were met.

2. Methodology

2.1. Study Area

This research was conducted in a mining region of Mexico, located in the state of Zacatecas, municipality of Mazapil. The study area is located between latitudes 24°33’57.57” and 24°42’14.56” North, and between longitudes 101°50’45.10” and 101°27’10.32” West (Figure 1) [4,30].

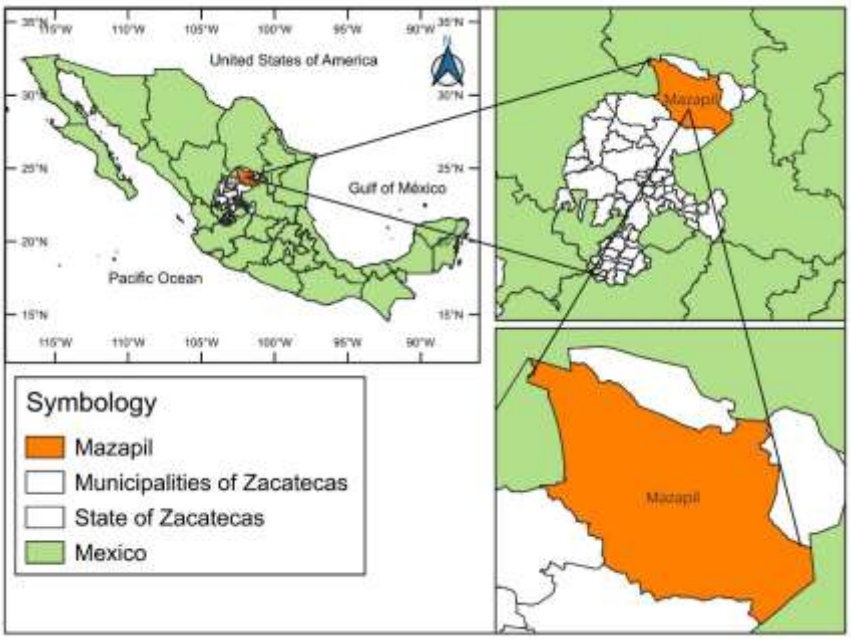

Figure 1. Location map of the study area (Source: Own creation with maps extracted from [31]).

 doi:10.20944/preprints202603.2049.v1

2.2. Methodology

The present research employed a quantitative approach for land cover variables. The proposed methodology is a combination and adaptation of methodologies proposed by [32,33]. The procedure was as follows:

1. Acquisition of satellite images before and after mining activity.
2. Radiometric and geometric corrections to the images.
3. Land cover classification using Landsat images with different methods and algorithms.
4. Evaluation and comparison of various methodologies and algorithms for land cover classification in mining areas.
5. Measurement and quantification of land cover changes in the mining area.

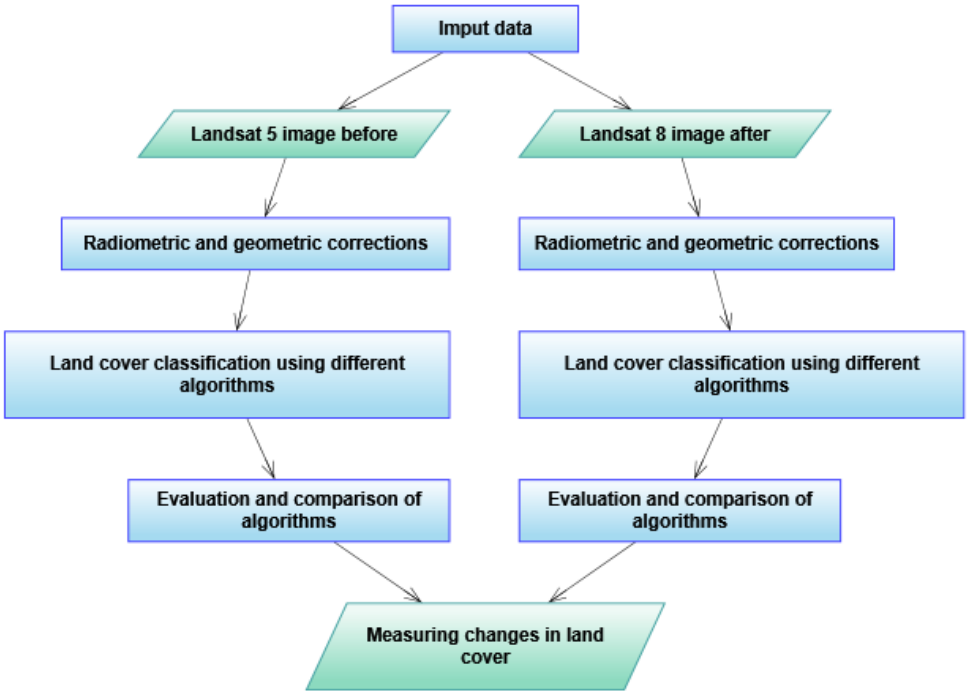

Figure 2. Methodology for the detection and measurement of changes in land cover in a mining area.

2.2.1. Acquisition of Satellite Images Before and After Mining Activity

Images from the Landsat 5 satellite (2006) and the Landsat 8 satellite (2025) were obtained to document the before and after of the start of large-scale mining activity. These images were downloaded from the United States Geological Survey (USGS) website [34].

2.2.2. Radiometric and Geometric Corrections to the Images

The Landsat 5 and 8 satellite images were converted to WGS 84 coordinates and projected onto UTM Zone 14 North. Digital Numbers (DN) were converted to Top of Atmospheric reflectance (TOA) using the equations recommended by the source [34,35]. The images were then cropped to the extent of a window covering the study area.

2.2.3. Land Cover Classification Using Landsat Images with Different Methods and Algorithms

To highlight the different land cover types, combinations of Shortwave Infrared 1 (SWIR 1), Near-Infrared (NIR), and Red bands were used. This false-color band combination is used for vegetation analysis and shows greater differentiation between soil, vegetation, and water bodies [36]. Seven categories were chosen to classify LC over the mining area of interest: water, mining cover,

 doi:10.20944/preprints202603.2049.v1

infrastructure, soil, shrubland, bare soil, and vegetation. These categories were obtained and adapted from [28], and their descriptions are shown in Table 1.

Table 1. Land Cover Classes in Mining Areas.

Land Cover Class	Description	Digital number
Water	Mining stockpiles and bodies of water in general	0
Mining Cover	Exposed mining excavations and waste rock, and tailings dams.	1
Infrastructure	Urban, road, and industrial buildings	2
Soil	Natural soil	3
Shrubland	Low-density, low-lying vegetation	4
Bare Soil	Cleared and deforested land due to human activities	5
Vegetation	Areas with dense vegetation	6

The following supervised image classification methods and algorithms available in SAGA 9.11 software were also used: BE, P, MD, ML, SAM, DT, K-NN, ME, NB, RF, and SVM.

The BE algorithm transforms image data and reference spectra into binary codes to perform efficient pixel classification. The process includes selecting regions of interest (ROIs), generating binary codes for each class, comparing using functions, and assigning based on band matching [37]. The P algorithm is a simple and computationally efficient supervised image classification method, primarily used in remote sensing. It classifies pixels by defining an n-dimensional box (parallelepiped) in the spectral feature space for each class, based on the statistical range of the training data [38]. The MD algorithm estimates the Euclidean distance between the pixel being studied and the classified pixels, assigning it to the class with the shortest distance between the compared pixels [39]. The ML classifier uses statistics for each class in each band, with a normal distribution, and calculates the probability that a given pixel belongs to a specific class, assigning it to the class with the highest probability [39]. The SAM classifier is a spectral classification based on the angles generated by the reflectance differences between two compared pixels, assigning the pixel to the classification with the smallest spectral angle [39]. DTs can be used for image classification as part of a traditional machine learning process involving the explicit extraction of features or attributes [40]. The KNN algorithm is a widely used nonparametric method for classification in pattern recognition. The fundamental principle of KNN is that the category of a data point is determined by the classification of its K nearest neighbors [41]. The ME algorithm is a probabilistic approach that selects from all models consistent with the known data, the one with the highest entropy (i.e., the least biased or informative). In remote sensing, it is used to derive class probabilities for each pixel (or object) under constraints derived from training data or features [42]. The NB Classifier is a supervised learning technique that uses Bayes' Theorem to assign labels to images based on statistical probabilities and can manage correlations between them if the data in each class follows a Gaussian distribution [43]. The RF method, proposed by Breiman, is one of the most important machine learning methods. RF has the advantages of being suitable for high-dimensional data, which can alleviate the influence of the curse of dimensionality; it is very fast because it is implemented in parallel; it is effective at handling data with unbalanced class labels or missing values; it is not prone to overfitting; it automatically measures the importance of variables; and it can achieve higher classification accuracy compared to other classifiers [44]. SVM is a theoretically superior machine learning methodology with excellent results in pattern recognition. It is especially useful for the supervised classification of high-dimensional datasets and has been shown to compete with the best machine learning algorithms [45].

 doi:10.20944/preprints202603.2049.v1

Results

3.1. Acquisition of Satellite Images Before and After Mining Activity

Satellite images were obtained from bands one to five and seven of Landsat 5, with a geometric resolution of 30m and from band one to seven for the Landsat 8 satellite, with a geometric resolution of 30m (Table 2).

Table 2. Information on the input data used in this research.

Input data	Year	Number of Bands	Pixel Resolution	Sources of Information
Landsat 5 images	2006	1-5 y 7	30m	[35]
Landsat 8 images	2025	1-7	30m	[34]

3.2. Radiometric and Geometric Corrections to the Images

The satellite images were corrected radiometrically and geometrically. Combinations of bands of Shortwave Infrared 1 (SWIR 1), Near-Infrared (NIR), and Red bands were also performed (Figures 3 and 4).

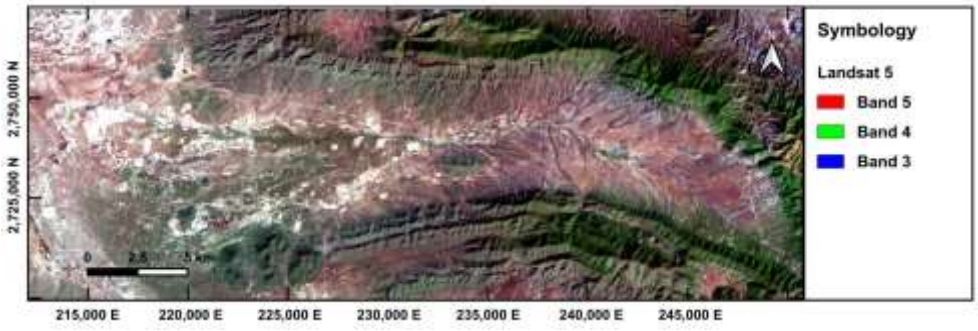

Figure 3. Combination of bands B5 Shortwave Infrared 1 (SWIR 1), B4 Near Infrared (NIR) and B3 Red, from 2006 and Landsat 5.

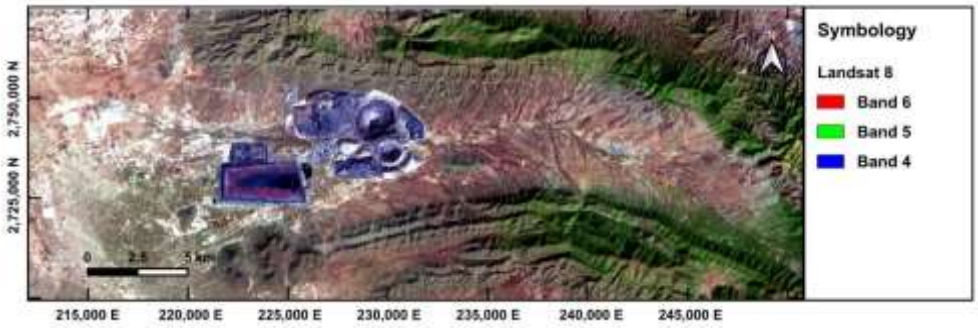

Figure 4. Combination of bands B6 Shortwave Infrared 1 (SWIR 1), B5 Near Infrared (NIR) and B4 Red, from the year 2025 and from Landsat 8.

3.3. Land Cover Classification Using Landsat Images with Different Methods and Algorithms

To classify the different land cover types, the first step was to perform a supervised pixel-based classification of the images. Pixel-based classification involves statistically analyzing the spectral values of each pixel, regardless of its position, and assigning them to a category based on their similarity to other pixels already classified in each category [46,47]. Regions of interest (ROIs), consisting of a set of pixels with similar reflectance, were defined as sampling data to define each classification category (Figures 5 and 6).

 doi:10.20944/preprints202603.2049.v1

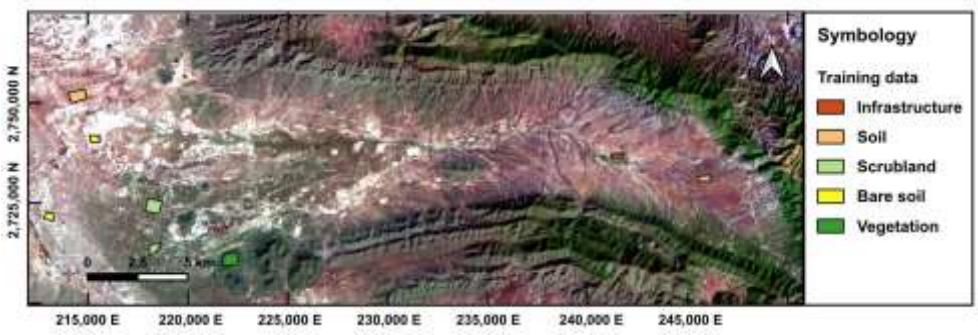

Figure 5. Classified training data from the 2006 Landsat 5 image.

Figure 6. Classified training data of the Landsat 8 image from the year 2025.

A supervised classification of the 2025 Landsat 8 images was performed using the BE, P, MD, ML, SAM, DT, K-NN, ME, NB, RF, and SVM algorithms (Figure 7). Different combinations of spectral bands were also used, such as Band 6 (B6), Band 5 (B5), and Band 4 (B4); Bands 1 through 7 (B1-B7); and bands one through seven with spectral indices (B1-B7, NDVI, SAVI, and BSI) (Table 3).

3.4. Evaluation and Comparison of Various Methodologies and Algorithms for Land Cover Classification in Mining Areas

To evaluate the reliability of the land cover classification, the image classified with the SAM algorithm was transformed into a vector layer and a visual and manual supervision of 6911 polygons was conducted, verifying if it was classified correctly, otherwise the classification was corrected by assigning the correct number from zero to six (Figure 8).

The classic confusion matrix technique was used to evaluate the results, which is the most widespread and recognized method for evaluating classification and change results by the scientific community [48–50]. The confusion matrix yielded the overall accuracy and Kappa coefficient for each classification method used (Table 3).

Table 3. Overall accuracy and Kappa coefficient for each classification method used.

	B1-B7, NDVI, SAVI and BSI		B1-B7		B6, B5 y B4	
Method or algorithm	**Overall Accuracy**	**Kappa**	**Overall Accuracy**	**Kappa**	**Overall Accuracy**	**Kappa**
Decision Tree	0.5685	0.4288	0.5287	0.3583	0.5503	0.4089
K-Nearest Neighbours	0.5847	0.4486	0.6158	0.4710	0.6222	0.4981
Maximum Entropy	0.4938	0.2993	0.6150	0.4701	0.6213	0.4971
Normal Bayes	0.4603	0.3274	0.3814	0.2691	0.5477	0.4314
Random Forest	0.5940	0.4472	0.5835	0.4240	0.5874	0.4478
Support Vector Machine [Polynomial]	0.6883	0.5360	0.6724	0.5125	0.6620	0.4933

 doi:10.20944/preprints202603.2049.v1

Binary Encoding	0.6749	0.5684	0.2769	0.1698	0.1917	-0.0181
Parallelepiped	0.4421	0.3092	0.4028	0.2365	0.2681	0.0996
Minimum Distance	0.6214	0.4749	0.4906	0.3259	0.4504	0.2695
Maximum Likelihood	0.4431	0.3160	0.3814	0.2691	0.5477	0.4314
Spectral Angle Mapping	0.6880	0.5815	0.6822	0.5673	0.8516	0.7928

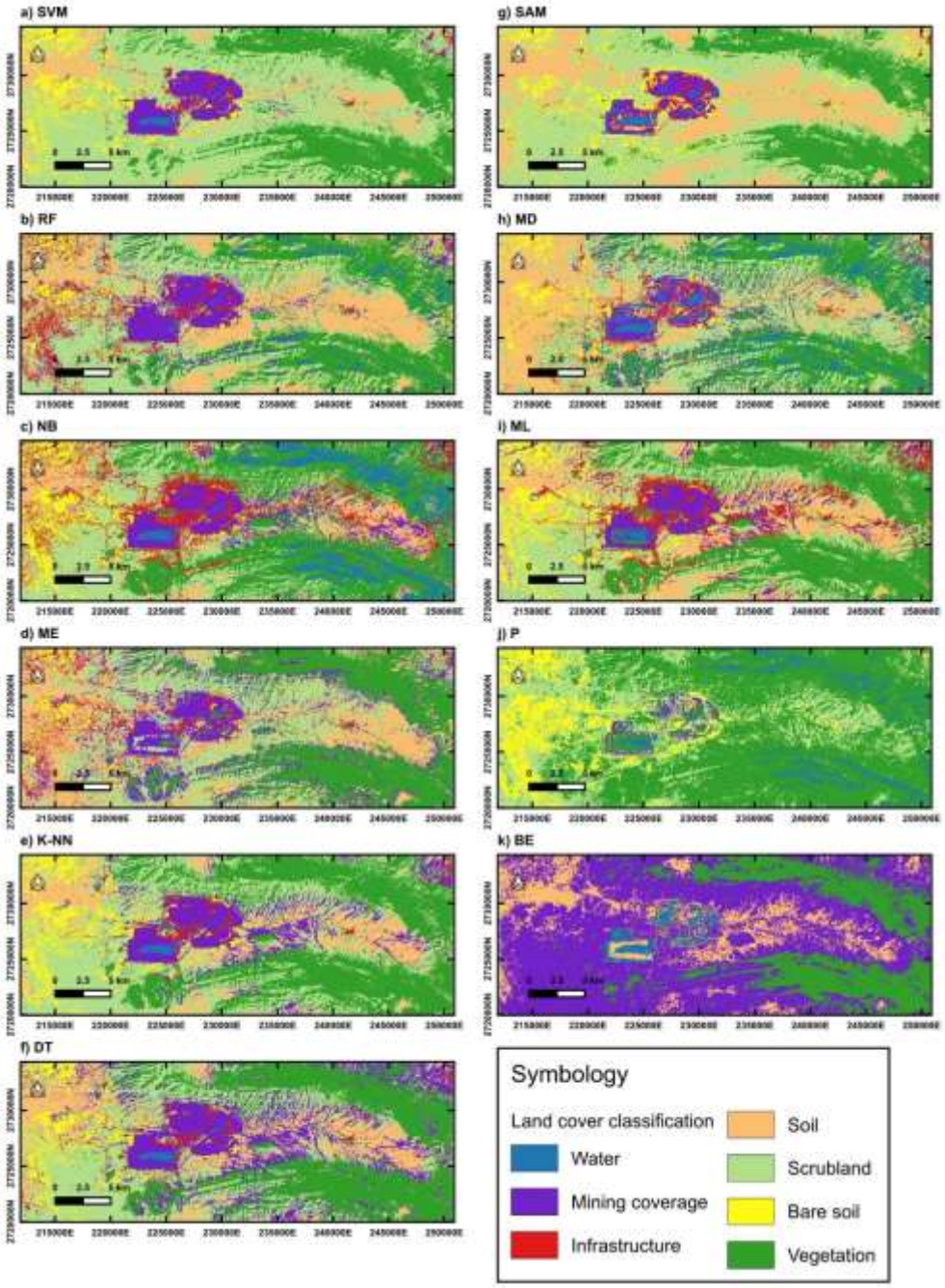

Figure 7. a) Supervised classification using the SVM method. b) Supervised classification using the RF method. c) Supervised classification using the NB method. d) Supervised classification using the ME method. e) Supervised classification using the K-NN method. f) Supervised classification using the DT method. g) Supervised classification using the SAM algorithm. h) Supervised classification using the MD algorithm. i) Supervised classification using the ML algorithm. j) Supervised classification using the P algorithm. k) Supervised classification using the BE algorithm.

 doi:10.20944/preprints202603.2049.v1

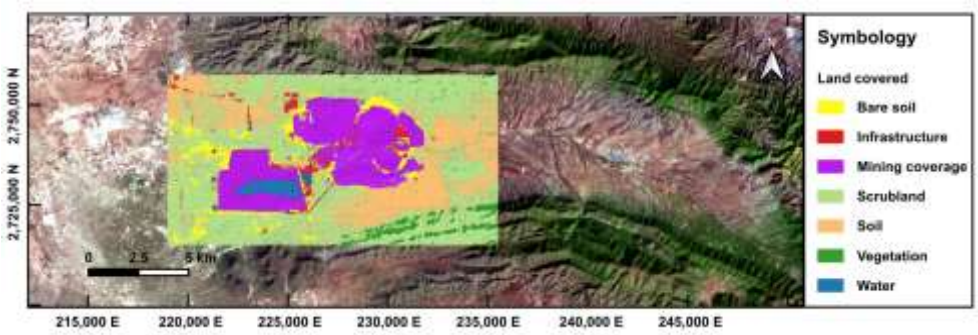

Figure 8. Classified and reviewed polygons for the validation of the generated classification maps.

After evaluating eleven supervised classification methods or algorithms using the confusion matrix and their overall accuracy, we observed that the most accurate algorithm was SAM, with an overall accuracy of 85.16% and a "Kappa" coefficient of 0.79 (Table 3). However, this overall accuracy varies for each classification category, as can be seen in the confusion matrix of supervised classification with the SAM algorithm for the year 2025 (Table 4).

Table 4. Confusion matrix of land cover classification for the year 2025.

Category	0	1	2	3	4	5	6	User's Total	User accuracy
0	2668	3447	11	3	2	56	0	6187	43.12%
1	123	18410	647	9	11	14	1	19215	95.81%
2	9	4763	2665	30	154	630	48	8299	32.11%
3	3	1121	4	29204	2107	34	2	32475	89.93%
4	0	535	1	19	68196	62	0	68813	99.10%
5	25	2084	803	1348	3745	8834	952	17791	49.65%
6	0	0	2	30	377	2	3293	3704	88.90%
Producer's Sum	2828	30360	4133	30643	74592	9632	4296	156484	
Producer Accuracy	94.34	60.64	64.48	95.30	91.43	91.72	76.65		85.16%

3.5. Measurement and Quantification of Changes in Land Cover over the Mining Area

The measurement and quantification of changes in LC was conducted by subtracting the area difference between the 2006 map and the 2025 map. The maps used were generated with the SAM algorithm, as they demonstrated the best results (Figures 9 and 10).

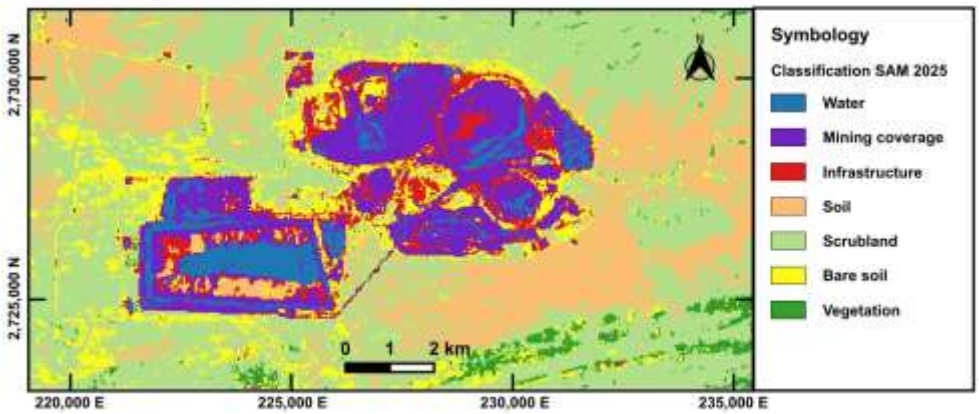

Figure 9. Supervised classification of LC with the SAM classifier in 2025.

 doi:10.20944/preprints202603.2049.v1

After evaluating which algorithm or methodology yielded the best results, the land cover classification map for the year 2006 was obtained (Figure 10).

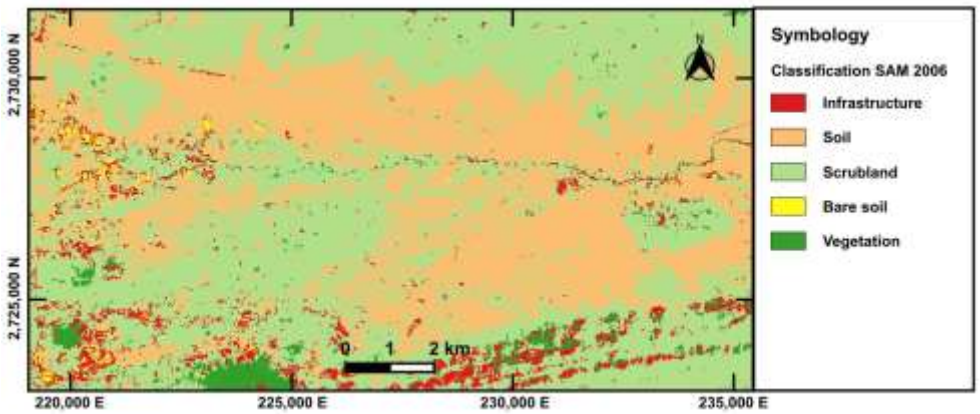

Figure 10. Supervised classification of LC with the SAM classifier in 2006.

After classifying the images and validating the data obtained, the number of pixels for each LC category was quantified (Table 5).

Table 5. Number of pixels classified for each LC category.

Type of Land Cover	2006 (Number of Pixels)	2025 (Number of Pixels)
Water		6187
Mining Cover		19215
Infrastructure	7719	8299
Soil	58836	32475
Shrubland	84864	68813
Bare Soil	1256	17791
Vegetation	3809	3704
Total	156484	156484

Next, the coverage areas for each category were calculated by multiplying the area of each pixel by the total number of pixels in each category. The area difference between 2006 and 2025 was also calculated. This allows us to estimate the increases or decreases in area for each category (Table 6).

Table 6. Land cover areas for the years 2006, 2025 and their differences.

Type of Land Cover	2006 Area (Ha)	2025 Area (Ha)	2006-2025 Area (Ha)
Water		556.83	556.83
Mining Cover		1729.35	1729.35
Infrastructure	694.71	746.91	52.2
Soil	5295.24	2922.75	-2372.49
Shrubland	7637.76	6193.17	-1444.59
Bare Soil	113.04	1601.19	1488.15
Vegetation	342.81	333.36	-9.45
Total	14083.56	14083.56	

From Table 6 we can observe that the total changes in LC from 2006 to 2025 had an increase in water bodies of 556.83 Ha, mining cover increased to 1729.35 Ha, infrastructure increased by 52.20 Ha, soil had a loss of -2372.49 Ha, scrubland had a loss of -1444.59 Ha, bare soil had an increase of 1488.15 Ha and vegetation had a loss of -9.45 Ha.

 doi:10.20944/preprints202603.2049.v1

4. Discussion

Detecting changes in LULC using satellite imagery is a technique that has been widely used in various research studies [51–54]. Landsat satellite imagery offers a temporal monitoring tool for over 50 years, with data available every 14 days, to track changes in LC [36]. Furthermore, different indices (NDVI, SAVI, BSI, among others [55]) and various band combinations in satellite imagery allow for highlighting specific land features. These indices and band combinations are used for monitoring vegetation health, deforestation, changes in soil and water bodies [36]. For this reason, this research used Landsat multi-temporal images due to their extensive temporal coverage for detecting changes in LC.

After conducting numerous tests with different combinations of information and algorithms, we observed that, in image classification, the proposed classification categories and training data are crucial for LC classification. If appropriate training data and classification categories are not selected, confusion can arise with the classification algorithms, increasing the classification error. We also observed different LC types with similar reflectance, leading to classification errors. For example, water categories were confused with the shadows of excavations, bare soil was confused with certain types of infrastructure with similar reflectance, and rock outcrops of the Earth's crust were confused with industrial and road infrastructure (Figure 7). This behavior has been reported in complex mining landscapes, where spectral heterogeneity and terrain geometry increase ambiguity in classification [5].

The algorithm that achieved the best results with bands B6, B5, and B4 was SAM, with an overall accuracy of 85.16%. With bands B1 through B7, the algorithm with the best results was SVM (Polynomial) with 67.24% accuracy, and for the combined bands and indices B1-B7, NDVI, SAVI, and BSI, the algorithm with the best results was SVM (Polynomial) with 68.83% accuracy (Table 3). Other authors have found similar accuracies, but with different algorithms and methodologies. For example, Isidro et al. (2017) obtained between 87% and 89% overall accuracy with an OB-SVM classifier in a large-scale mine in Didipio, Philippines, using Pleiades and SPOT images [29]. Similarly, Mi et al. (2019) obtained an accuracy between 81.92% and 86.6% for LULC changes in the Nanjiao mining area, Shanxi Province, China, between 1987 and 2017 using RF classifiers and continuous Landsat imagery, applying a Savitzky-Golay filter and an NDVI-based approach [8]. Therefore, it is confirmed that the selection of the optimal algorithm depends strongly on the type of landscape, the data resolution, and the training strategy. In particular, the good performance of the SAM algorithm in this study can be attributed to its ability to discriminate spectral differences in heterogeneous materials, such as those found in mining areas, where mixtures of minerals, exposed soils, and disturbed surfaces predominate [11].

As we can see from Table 3, including more input data as spectral indices does not always improve the results. This may be due to the introduction of redundancy or noise in the data, which negatively affects the performance of some classification algorithms. This behavior has been reported in studies where increasing the number of variables without proper selection can reduce the predictive capacity of the models [56–59].

From an environmental perspective, the changes detected in LC between 2006 and 2025 reflect a significant landscape transformation associated with mining activity. The increase in mining cover (+1729.35 Ha) and bare soil (+1488.15 Ha) suggests an active expansion of mining areas, as well as the removal of vegetation cover and direct exposure of the soil. This type of change increases susceptibility to erosion processes and promotes the dispersion of fine particles into the atmosphere and water bodies, which has been linked to the transport of heavy metals in mining areas [1]. Likewise, the increase in water bodies (+556.83 Ha) could be associated with the formation of tailings dams or artificial water accumulations, which implies modifications to the local hydrology and potential contamination risks.

On the other hand, the loss of soil (-2372.49 Ha), scrubland (-1444.59 Ha), and vegetation (-9.45 Ha) demonstrates a reduction in natural cover, which can have significant implications for biodiversity, soil stability, and ecosystem services. The decrease in vegetation reduces carbon

 doi:10.20944/preprints202603.2049.v1

sequestration capacity, alters hydrological cycles, and diminishes the ecosystem's resilience to disturbances [2]. Furthermore, the increase in bare surfaces in arid areas, such as Zacatecas, can favor the emission of particulate matter, which represents a potential risk to human health and air quality, especially in regions close to human settlements.

From the evaluation of the reliability of the obtained data, we can observe that different user accuracy percentages were obtained for each category. For example, for Water it was 43.12%, Mining cover 95.81%, Infrastructure 32.11%, Soil 89.93%, Soil with scrubland 99.10%, Bare soil 49.65% and Vegetation 88.90% (Table 4). The low accuracy in categories such as water and infrastructure reflects the difficulty of discriminating spectrally similar classes in mining environments, suggesting the need to incorporate auxiliary data or more advanced techniques, such as object-oriented classification or deep learning.

Finally, it is important to note some limitations of the study. The spatial resolution of Landsat images (30 m) can generate spectral mixing problems in heterogeneous areas, affecting classification accuracy. Likewise, the validation process based on visual supervision can introduce a degree of subjectivity. The lack of field data limits independent validation of the results. Therefore, future research should consider using higher-resolution imagery (e.g., Sentinel-2 or UAVs), as well as integrating in-situ data to improve the accuracy and robustness of the models. Furthermore, incorporating deep learning approaches could allow for better discrimination of complex classes in dynamic mining landscapes.

5. Conclusions

As the results show, it was possible to remotely monitor LC changes in a mining area. Pixel-supervised classification using different algorithms effectively differentiated LC types in the mining zone. Changes were detected over the 19 years analyzed in the following categories: Water bodies, mining cover, infrastructure, soil, scrubland, bare soil, and vegetation. Among the classifiers evaluated, the SAM method proved to be the most accurate in identifying changes related to mining activity, demonstrating its ability to discriminate spectral signatures in highly heterogeneous environments such as mining landscapes.

The methodology used to assess data reliability proved effective, achieving an overall accuracy of 85.16% and a Kappa coefficient of 0.79 with the SAM algorithm. These results reflect that the visual supervision process and the application of the confusion matrix are robust tools for validating the generated classes and ensuring the quality of the information obtained. Thus, it is ensured that decisions based on this data have reliable and adequate scientific backing for future analyses and applications. Furthermore, the systematic comparison of eleven algorithms constitutes a significant contribution, as it allows for the identification of optimal configurations for LC classification in mining contexts in Mexico, where this type of study is still limited.

The analysis of satellite images reveals significant changes in LC between 2006 and 2025. The increase in bodies of water and the expansion of mining cover and infrastructure reflect a major landscape transformation, primarily associated with the intensification of extractive activities. In particular, the increase in bare soil and mining areas indicates processes of vegetation removal and soil exposure, which increases susceptibility to erosion and particle dispersion. On the other hand, the loss of soil, scrubland, and vegetation indicates a reduction in natural ecosystems, which can have significant environmental implications such as decreased biodiversity, disruption of hydrological cycles, and degradation of ecosystem services. Taking together, these changes reflect cumulative environmental pressure that could intensify in the absence of appropriate management strategies.

From an applied perspective, the results obtained demonstrate the potential of remote sensing and machine learning as key tools for environmental monitoring in mining areas. This information can be used to support environmental impact assessment processes, land-use planning, monitoring of mining concessions, and the design of ecological restoration strategies. In this way, the systematic monitoring of land cover is positioned as a fundamental input for informed decision-making and the development of public policies oriented toward sustainability.

 doi:10.20944/preprints202603.2049.v1

Future research could improve the definition of the categories by observing the results and conducting tests with other, more detailed classifications. To improve the accuracy of LC classification, it is recommended to include other sources of information, such as higher spatial resolution images (e.g., Sentinel-2 or UAVs), as well as auxiliary variables (topography, climate, or in-situ data). Furthermore, the incorporation of advanced techniques such as deep learning could allow for better discrimination of complex and dynamic LC classes in mining environments.

The use of supervised classification and satellite image analysis enabled the precise identification and quantification of LC changes in a mining area over nearly two decades. The reliability of the data and the methodological rigor support the validity of the results, demonstrating the transformation of the landscape and the decline of natural ecosystems. These findings underscore the urgency of adopting sustainable management measures, ecological restoration, and continuous monitoring. In this sense, the proposed approach not only contributes to methodological advancements in LC classification but also represents a strategic tool for assessing and mitigating the environmental impacts of mining activity, promoting a balance between resource use and environmental conservation.

Author Contributions: Conceptualization, S.D-C. and A-G.C-M.; methodology, S.D-C. and A-G.C-M.; validation, C-F.B-C., E-D.M-V. and V-I.R-A.; formal analysis, C-O.R-R.; investigation, S.D-C. and A.G. C-M.; data curation, S.D-C. and A.G. C-M.; writing—original draft preparation, S.D-C. and A-G. C-M.; writing—review and editing, S.D-C. and A-G. C-M.; visualization, E-D.M-V. and V-I.R-A.; supervision, C-O.R-R., C-F.B-C. and E-D.M-V. All authors have read and agreed to the published version of the manuscript".

Funding: This research received no external funding.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author(s).

Acknowledgments: The authors thank the Secretariat of Science, Humanities, Technology, and Innovation (SECIHTI) of Mexico for the scholarship awarded to Master of Science Saúl Dávila Cisneros to pursue doctoral studies and generate this research.

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

Abbreviations

The following abbreviations are used in this manuscript:

LC	Land Cover
SAM	Spectral Angle Mapper
EIA	Environmental Impact Assessment
LULC	Land Use and Land Cover
LICA	Landsat Image Classification Algorithms
GEE	Google Earth Engine
GA	Genetic Algorithm
CV	Cross-Validation
PSO	Particle Swarm Optimization
SVM	Support Vector Machine
GIS	Geographic Information Systems
LST	Land Surface Temperature
HDCA	homogeneity distance classification algorithm
CEEs	Cumulative Ecological Effects
FLCC	Fine Land Cover Classification
CNN	Convolutional Neural Network
RSEI	Remote Sensing Ecological Index
BiLTM	Bidirectional Long-Term Memory
DEMs	Digital Elevation Models
NDVI	Normalized Difference Vegetation Index

 doi:10.20944/preprints202603.2049.v1

SRF	Smile Random Forest
SGTB	Smile Gradient Tree Boost
UAVs	Unmanned Aerial Vehicles
NDWI	Normalized Difference Water Index
BI	Brightness Index
SCI	Soil Crust Index
BE	Binary Encoding
P	Parallelepiped
MD	Minimum Distance
MLi	Maximum Likely-hood
DT	Decision Tree
K-NN	K-Nearest Neighbours
ME	Maximum Entropy
NB	Normal Bayes
RF	Random Forest
ML	Machine Learning
USGS	United States Geological Survey
DN	Digital Numbers
TOA	Top of Atmospheric reflectance
SWIR 1	Shortwave Infrared 1
NIR	Near-Infrared
ROIs	Regions Of Interest

References

1. Mardonova, M.; Han, Y.-S. Environmental, Hydrological, and Social Impacts of Coal and Nonmetal Minerals Mining Operations. *J. Environ. Manage.* **2023**, *332*, 117387, doi:https://doi.org/10.1016/j.jenvman.2023.117387.
2. Prakash, V.; Budhwan, R. Impact of Stone Mining on Vegetation and Biodiversity and Its Restoration Approaches. In *Ecological Impacts of Stone Mining: Assessment and Restoration of Soil, Water, Air and Flora*; Sharma, G.K., Rashmi, I., Ali, S., Kala, S., Kumar, A., Madhu, M., Eds.; Springer Nature Singapore: Singapore, 2024; pp. 187–228 ISBN 978-981-97-4746-7.
3. Pagouni, C.; Pavloudakis, F.; Kapageridis, I.; Yiannakou, A. Transitional and Post-Mining Land Uses: A Global Review of Regulatory Frameworks, Decision-Making Criteria, and Methods. *Land* **2024**, *13*, doi:10.3390/land13071051.
4. Dávila-Cisneros, S.; Castañeda-Miranda, A.G.; Bautista-Capetillo, C.F.; Mattos-Villarroel, E.D.; Rodríguez-Abdalá, V.I.; Robles Rovelo, C.O.; Pinedo-Torres, L.A.; Rodríguez-Trejo, A.; Ibarra-Delgado, S. Monitoring Landform Changes in a Mining Area in Mexico Using Geomatic Techniques. *Geomatics* **2025**, *5*, doi:10.3390/geomatics5040063.
5. Zhao, S.; Tu, K.; Ye, S.; Tang, H.; Hu, Y.; Xie, C. Land Use and Land Cover Classification Meets Deep Learning: A Review. *Sensors* **2023**, *23*, doi:10.3390/s23218966.
6. Srinivasa Raju, G.; Vanka, T.S.T.; Ajay, Ch.; Ameer, S.; Sunil, R. Machine Learning Techniques for Land-Use Land-Cover Classification from Satellite Images Using Hybrid Models. In Proceedings of the Smart Computing Paradigms: Intelligence and Network Applications; Bhateja, V., Hoong, A.L.S., Kong, J.D., Urooj, S., Eds.; Springer Nature Switzerland: Cham, 2026; pp. 247–256.
7. Capolupo, A.; Monterisi, C.; Tarantino, E. Landsat Images Classification Algorithm (LICA) to Automatically Extract Land Cover Information in Google Earth Engine Environment. *Remote Sens.* **2020**, *12*, doi:10.3390/rs12071201.
8. Mi, J.; Yang, Y.; Zhang, S.; An, S.; Hou, H.; Hua, Y.; Chen, F. Tracking the Land Use/Land Cover Change in an Area with Underground Mining and Reforestation via Continuous Landsat Classification. *Remote Sens.* **2019**, *11*, doi:10.3390/rs11141719.
9. Wei, Z.; Liu, T.; Liang, H.; Zhang, Z.; Wang, C.; Lv, Y.; Sun, J.; Wang, Q. Uncovering the Impacts of LUCC on Ecological Connectivity in Suburban Open-Pit Mining Concentration Areas: A Pattern Collection of Changing Relationships between Ecological Resistance and Ecological Network Elements. *Ecol. Model.* **2024**, *496*, 110815, doi:https://doi.org/10.1016/j.ecolmodel.2024.110815.

 doi:10.20944/preprints202603.2049.v1

10. Zhai, Y.; Qu, Z.; Hao, L. Land Cover Classification Using Integrated Spectral, Temporal, and Spatial Features Derived from Remotely Sensed Images. *Remote Sens.* **2018**, *10*, doi:10.3390/rs10030383.
11. Chen, W.; Li, X.; Wang, L. Fine Land Cover Classification in an Open Pit Mining Area Using Optimized Support Vector Machine and WorldView-3 Imagery. *Remote Sens.* **2020**, *12*, doi:10.3390/rs12010082.
12. Arifeen, H.M.; Chowdhury, Md.S.; Zhang, H.; Suepa, T.; Amin, N.; Techato, K.; Jutidamrongphan, W. Role of a Mine in Changing Its Surroundings—Land Use and Land Cover and Impact on the Natural Environment in Barapukuria, Bangladesh. *Sustainability* **2021**, *13*, doi:10.3390/su132413602.
13. Firozjaei, M.K.; Sedighi, A.; Firozjaei, H.K.; Kiavarz, M.; Homaee, M.; Arsanjani, J.J.; Makki, M.; Naimi, B.; Alavipanah, S.K. A Historical and Future Impact Assessment of Mining Activities on Surface Biophysical Characteristics Change: A Remote Sensing-Based Approach. *Ecol. Indic.* **2021**, *122*, 107264, doi:https://doi.org/10.1016/j.ecolind.2020.107264.
14. Li, C. A Developed Method of Water Pollution Control Based on Environmental Capacity and Environmental Flow in Luanhe River Basin. *Water* **2022**, *14*, 730, doi:10.3390/w14050730.
15. Qian, M.; Sun, S.; Li, X. Multimodal Data and Multiscale Kernel-Based Multistream CNN for Fine Classification of a Complex Surface-Mined Area. *Remote Sens.* **2021**, *13*, doi:10.3390/rs13245052.
16. Guan, R.; Li, Z.; Li, T.; Li, X.; Yang, J.; Chen, W. Classification of Heterogeneous Mining Areas Based on ResCapsNet and Gaofen-5 Imagery. *Remote Sens.* **2022**, *14*, doi:10.3390/rs14133216.
17. Tang, H.; Fang, J.; Xie, R.; Ji, X.; Li, D.; Yuan, J. Impact of Land Cover Change on a Typical Mining Region and Its Ecological Environment Quality Evaluation Using Remote Sensing Based Ecological Index (RSEI). *Sustainability* **2022**, *14*, doi:10.3390/su141912694.
18. Zhang, J.; Liu, M.; Song, Y. Human-Dominated Land Use Change in a Phosphate Mining Area and Its Impact on the Water Environment. *Water* **2022**, *14*, doi:10.3390/w14071074.
19. Yan, X.; Li, J.; Smith, A.R.; Yang, D.; Ma, T.; Su, Y. Rapid Land Cover Classification Using a 36-Year Time Series of Multi-Source Remote Sensing Data. *Land* **2023**, *12*, doi:10.3390/land12122149.
20. Himmy, O.; Nguyen, T.T.; Hemmler, K.S.; Loulad, S.; Rhinane, H.; Buerkert, A. Leveraging Machine Learning and Landsat Time Series for High-Resolution Mapping of Mining-Induced Vegetation Changes in Ouagadougou, Burkina Faso. *Environ. Chall.* **2024**, *17*, 101026, doi:https://doi.org/10.1016/j.envc.2024.101026.
21. Mutimba, K.; Watanabe, T.; Chand, M.B. Land Use Land Cover (LULC) Change Dynamics Associated with Mining Activities in Kitwe District and Adequacy of the Legal Framework on Mine Closure in Zambia. *Earth* **2024**, *5*, 110–132, doi:10.3390/earth5020006.
22. Xu, H.; Cheng, W. Landscape Analysis and Ecological Risk Assessment during 1995–2020 Based on Land Utilization/Land Coverage (LULC) and Random Forest: A Case Study of the Fushun Open-Pit Coal Area in Liaoning, China. *Sustainability* **2024**, *16*, doi:10.3390/su16062442.
23. Qiao, Q.; Li, Y.; Lv, H. Open-Pit Mining Area Extraction Using Multispectral Remote Sensing Images: A Deep Learning Extraction Method Based on Transformer. *Appl. Sci.* **2024**, *14*, doi:10.3390/app14146384.
24. Wang, Y.; Qin, K.; Zhang, Z.; He, Q.; Cohen, J. Mapping Open-Pit Mining Area in Complex Mining and Mixed Land Cover Zone Using Landsat Imagery. *Int. J. Appl. Earth Obs. Geoinformation* **2024**, *129*, 103782, doi:https://doi.org/10.1016/j.jag.2024.103782.
25. Ibukun, J.A.; Olubaju, A.E.; Thomas, S.F.; Sodipo, E.O.; Akinbiola, S.A.; Oyetunji, S.O.; Shitu, K.; Kucher, D.E.; Tariq, A. Modeling Mining-Induced Land Degradation in Itagunmodi: A Multi-Temporal Machine Learning Approach with Random Forest and Gradient Boosting. *Trees For. People* **2025**, *21*, 100926, doi:https://doi.org/10.1016/j.tfp.2025.100926.
26. Loukili, I.; Laamrani, A.; Ghorfi, M.E.; Moutak, S.E.; ghafiri, A. Monitoring Land Changes at an Open Mine Site Using Remote Sensing and Multi-Spectral Indices. *Heliyon* **2025**, *11*, e41845, doi:https://doi.org/10.1016/j.heliyon.2025.e41845.
27. Reinprecht, V.; Kieffer, D.S. Application of UAV Photogrammetry and Multispectral Image Analysis for Identifying Land Use and Vegetation Cover Succession in Former Mining Areas. *Remote Sens.* **2025**, *17*, doi:10.3390/rs17030405.

 doi:10.20944/preprints202603.2049.v1

28. Blanche, M.F.; Dairou, A.A.; Juscar, N.; Romarice, O.M.F.; Arsene, M.; Bernard, T.L.; Leroy, M.N.L. Assessment of Land Cover Degradation Due to Mining Activities Using Remote Sensing and Digital Photogrammetry. *Environ. Syst. Res.* **2024**, *13*, 41, doi:10.1186/s40068-024-00372-5.
29. Isidro, C.M.; McIntyre, N.; Lechner, A.M.; Callow, I. Applicability of Earth Observation for Identifying Small-Scale Mining Footprints in a Wet Tropical Region. *Remote Sens.* **2017**, *9*, doi:10.3390/rs9090945.
30. INEGI Compendio de Información Geográfica Municipal 2010. Mazapil, Zacatecas 2010.
31. INEGI MARCO GEOESTADÍSTICO INTEGRADO, DICIEMBRE 2021 2021.
32. Ruiz Fernández, L.Á. Métodos de Detección de Cambios En Teledetección 2017.
33. Angeles, G.R.; Geraldi, A.M.; Marini, M.F. *PROCESAMIENTO DIGITAL DE IMÁGENES SATELITALES. METODOLOGÍAS Y TÉCNICAS*; BAHIA BLANCA, ARGENTINA, 2020; ISBN 978-987-86-4773-9.
34. Earth Resources Observation and Science Center. Landsat 8-9 Operational Land Imager / Thermal Infrared Sensor Level-1, Collection 2 2020.
35. Earth Resources Observation and Science (EROS) Center Landsat 4-5 Mapeador Temático Nivel-1 2020.
36. USGS What are the band designations for the Landsat satellites? 2025.
37. Xie, H.; Tong, X. A Probability-Based Improved Binary Encoding Algorithm for Classification of Hyperspectral Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 2108–2118, doi:10.1109/JSTARS.2013.2273795.
38. Xiang, M.; Hung, C.-C.; Pham, M.; Kuo, B.-C.; Coleman, T. A Parallelepiped Multispectral Image Classifier Using Genetic Algorithms. In Proceedings of the Proceedings. 2005 IEEE International Geoscience and Remote Sensing Symposium, 2005. IGARSS '05.; 2005; Vol. 1, p. 4 pp.-.
39. Sunil, J.; Singh, A. Evaluation of the Performance of Supervised Classification Alogorithums in Image Classification. *Int J Curr Microbiol Appl Sci* **2019**, *8*, 2634–2643.
40. Agarwal, C.; Sharma, A. Image Understanding Using Decision Tree Based Machine Learning. In Proceedings of the ICIMU 2011 : Proceedings of the 5th international Conference on Information Technology & Multimedia; 2011; pp. 1–8.
41. Guo, Y.; Han, S.; Li, Y.; Zhang, C.; Bai, Y. K-Nearest Neighbor Combined with Guided Filter for Hyperspectral Image Classification. *Procedia Comput. Sci.* **2018**, *129*, 159–165, doi:https://doi.org/10.1016/j.procs.2018.03.066.
42. Qi, Y. The Impact of Artificial Intelligence on Environmental Protection. *Highlights Sci. Eng. Technol.* **2024**, *96*, 152–156, doi:10.54097/nzbf8s91.
43. Ruiz, P.; Mateos, J.; Camps-Valls, G.; Molina, R.; Katsaggelos, A.K. Bayesian Active Remote Sensing Image Classification. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 2186–2196, doi:10.1109/TGRS.2013.2258468.
44. Zhang, Y.; Cao, G.; Li, X.; Wang, B.; Fu, P. Active Semi-Supervised Random Forest for Hyperspectral Image Classification. *Remote Sens.* **2019**, *11*, doi:10.3390/rs11242974.
45. Tzotsos, A.; Argialas, D. Support Vector Machine Classification for Object-Based Image Analysis. In *Object-Based Image Analysis: Spatial Concepts for Knowledge-Driven Remote Sensing Applications*; Blaschke, T., Lang, S., Hay, G.J., Eds.; Springer Berlin Heidelberg: Berlin, Heidelberg, 2008; pp. 663–677 ISBN 978-3-540-77058-9.
46. Jog, S.; Dixit, M. Supervised Classification of Satellite Images. In Proceedings of the 2016 Conference on Advances in Signal Processing (CASP); 2016; pp. 93–98.
47. Suman, S.; Srivastava, P.K.; Petropoulos, G.P.; Avtar, R.; Prasad, R.; Singh, S.K.; Mustak, S.K.; Faraslis, I.N.; Gupta, D.K. Chapter 12 - Performance Assessment of Phased Array Type L-Band Synthetic Aperture Radar and Landsat-8 Used in Image Classification. In *Radar Remote Sensing*; Srivastava, P.K., Gupta, D.K., Islam, T., Han, D., Prasad, R., Eds.; Earth Observation; Elsevier, 2022; pp. 219–244 ISBN 978-0-12-823457-0.
48. Markoulidakis, I.; Kopsiaftis, G.; Rallis, I.; Georgoulas, I. Multi-Class Confusion Matrix Reduction Method and Its Application on Net Promoter Score Classification Problem. In Proceedings of the Proceedings of the 14th PErvasive technologies related to assistive environments conference; 2021; pp. 412–419.
49. Heydarian, M.; Doyle, T.E.; Samavi, R. MLCM: Multi-Label Confusion Matrix. *Ieee Access* **2022**, *10*, 19083–19095.
50. Sathyanarayanan, S.; Tantri, B.R. Confusion Matrix-Based Performance Evaluation Metrics. *Afr. J. Biomed. Res.* **2024**, *27*, 4023–4031.

 doi:10.20944/preprints202603.2049.v1

51. Majeed, M.; Tariq, A.; Anwar, M.M.; Khan, A.M.; Arshad, F.; Mumtaz, F.; Farhan, M.; Zhang, L.; Zafar, A.; Aziz, M.; et al. Monitoring of Land Use–Land Cover Change and Potential Causal Factors of Climate Change in Jhelum District, Punjab, Pakistan, through GIS and Multi-Temporal Satellite Data. *Land* **2021**, *10*, 1026.
52. Mashala, M.J.; Dube, T.; Mudereri, B.T.; Ayisi, K.K.; Ramudzuli, M.R. A Systematic Review on Advancements in Remote Sensing for Assessing and Monitoring Land Use and Land Cover Changes Impacts on Surface Water Resources in Semi-Arid Tropical Environments. *Remote Sens.* **2023**, *15*, doi:10.3390/rs15163926.
53. Taiwo, B.E.; Kafy, A.-A.; Samuel, A.A.; Rahaman, Z.A.; Ayowole, O.E.; Shahrier, M.; Duti, B.M.; Rahman, M.T.; Peter, O.T.; Abosede, O.O. Monitoring and Predicting the Influences of Land Use/Land Cover Change on Cropland Characteristics and Drought Severity Using Remote Sensing Techniques. *Environ. Sustain. Indic.* **2023**, *18*, 100248.
54. Yuh, Y.G.; Tracz, W.; Matthews, H.D.; Turner, S.E. Application of Machine Learning Approaches for Land Cover Monitoring in Northern Cameroon. *Ecol. Inform.* **2023**, *74*, 101955.
55. Tran, T.V.; Reef, R.; Zhu, X. A Review of Spectral Indices for Mangrove Remote Sensing. *Remote Sens.* **2022**, *14*, doi:10.3390/rs14194868.
56. Belgiu, M.; Drăguţ, L. Random Forest in Remote Sensing: A Review of Applications and Future Directions. *ISPRS J. Photogramm. Remote Sens.* **2016**, *114*, 24–31, doi:https://doi.org/10.1016/j.isprsjprs.2016.01.011.
57. Ramezan, C.A.; Warner, T.A.; Maxwell, A.E.; Price, B.S. Effects of Training Set Size on Supervised Machine-Learning Land-Cover Classification of Large-Area High-Resolution Remotely Sensed Data. *Remote Sens.* **2021**, *13*, doi:10.3390/rs13030368.
58. Maharana, K.; Mondal, S.; Nemade, B. A Review: Data Pre-Processing and Data Augmentation Techniques. *Glob. Transit. Proc.* **2022**, *3*, 91–99, doi:https://doi.org/10.1016/j.gltp.2022.04.020.
59. Loog, M.; Krijthe, J.H.; Bicego, M. Also for K-Means: More Data Does Not Imply Better Performance. *Mach. Learn.* **2023**, *112*, 3033–3050, doi:10.1007/s10994-023-06361-6.