Cuadernos de Investigación Geográfica *Geographical Research Letters*	2024	Nº 50 (2)	pp. 157-177	EISSN 1697-9540

Copyright © 2024, The authors. This work is licensed under a Creative Commons Attribution 4.0 International License.

http://doi.org/10.18172/cig.6273

REMOTE SENSING OF ILLEGAL DUMPS THROUGH SUPERVISED CLASSIFICATION OF SATELLITE IMAGES: APPLICATION IN OAXACA, MEXICO

JAVIER GÓMEZ MATURANO[1]*, JOSÉ DAVID MENDOZA SANTANA[2], ANA LILIA AGUILAR-GARCÍA[2], MAYRA SERNA HERNÁNDEZ[2]

[1] *Centro de Investigación en Materiales Avanzados, Chihuahua, México.*

[2] *Centro de Investigaciones y Estudios Superiores en Antropología-Golfo, México.*

ABSTRACT. Various economic, social, and cultural factors have contributed to the proliferation of illegal dumps, causing urban image degradation, population health impacts, and soil, air, and water contamination. Scientists developed remote sensing techniques to identify these red spots and thus contribute to their mitigation and control. They recently used these techniques to detect large areas of illegal waste dumping instead of using expensive field monitoring. Artificial intelligence algorithms have been used to process satellite images due to the availability of satellite images and the increase in the processing capacity of computer systems. This work presents the results of a satellite remote-sensing procedure to detect illegal dumps in one hydrographic subbasin in Oaxaca, Mexico, through a supervised land cover classification using a Random Forest classifier. Two hundred and fifty-six control polygons were used to train the classifier. The classification criteria were the twelve bands of the Sentinel 2A satellite images with a spatial resolution of 10x10 meters, the spectral indices NDVI, MNDWI, SAVI, NDBI, BSI, and the surface slope. Google Earth Engine platform was used to process satellite images. There were 288,100 hectares classified in this way: 65.4% classified as vegetation, 31.5% like bare soil, 2.7% was urban soil and the rest was classified as water or garbage. A confusion matrix calculated the accuracy of the model in 0.9517. The model was not able to accurately distinguish between urban soil, bare soil and garbage due to the similarity of their spectral fingerprints. NDVI and SAVI were the most important spectral indices for detecting litter, and those might contribute to building a spectral fingerprint of litter in the future. Poorly classified areas were discarded through photointerpretation work and post-processing. Finally, thirty-two probable illegal dumps were identified, twelve of which were confirmed on the territory.

Teledetección de vertederos ilegales mediante clasificación supervisada de imágenes de satélite: aplicación en Oaxaca, México

RESUMEN. Diversos factores económicos, sociales y culturales han contribuido a la proliferación de vertederos ilegales, ocasionando degradación de la imagen urbana, afectaciones a la salud de la población y contaminación del suelo, aire y agua. Diversas técnicas de percepción remota se han desarrollado para identificar estos focos rojos y así contribuir a su mitigación y control. La percepción remota de satélites ha sido utilizada en los últimos años para detectar amplias zonas de vertido ilegal de residuos, en lugar de los costosos monitoreos en campo. Se han utilizado algoritmos de inteligencia artificial para procesar imágenes de satélite gracias a su disponibilidad y al aumento en la capacidad de procesamiento de los sistemas informáticos. Este trabajo presenta los resultados de un procedimiento de teledetección por satélite para detectar vertederos clandestinos en una subcuenca hidrográfica en Oaxaca, México, a través de una clasificación supervisada de cobertura terrestre utilizando el clasificador *Random Forest*. Se utilizaron doscientos cincuenta y seis polígonos de control para entrenar al clasificador. Los criterios de clasificación fueron las doce bandas de las imágenes del Sentinel 2ª, con una resolución espacial de 10x10 metros, los índices espectrales NDVI, MNDWI, SAVI, NDBI, BSI y la pendiente de la superficie. Para el procesamiento de las imágenes de satélites se utilizó la plataforma Google Earth Engine. Se obtuvieron 288.100

hectáreas clasificadas de esta manera: 65,4% clasificadas como vegetación, 31,5% como suelo desnudo, 2,7% como suelo urbano y el resto como agua o basura. Una matriz de confusión calculó la precisión del modelo en 0,9517. El modelo no fue capaz de distinguir con precisión entre suelo urbano, suelo desnudo y basura debido a la similitud de sus huellas espectrales. Los índices espectrales más importantes para detectar basura fueron el NDVI y SAVI, los cuales podrían contribuir a construir una huella espectral de basura en el futuro. Las áreas mal clasificadas se descartaron mediante trabajos de fotointerpretación y posprocesamiento. Finalmente, se identificaron treinta y dos probables vertederos clandestinos, doce de los cuales fueron confirmados en el territorio.

Keywords: remote sensing, illegal dumps, supervised classification, urban solid waste, artificial intelligence.

Palabras clave: teledetección, vertederos clandestinos, clasificación supervisada, residuos sólidos urbanos, inteligencia artificial.

Received: 30 May 2024

Accepted: 15 October 2024

***Corresponding author:** Javier Gómez Maturano, Centro de Investigación en Materiales Avanzados, Chihuahua, México. E-mail: jgomezma@uaemex.mx

1. Introduction

In developing countries like Mexico, it is difficult to find cases in which urban solid waste (MSW) is adequately confined to avoid impacts on the natural and social environment. MSW Final Disposal Sites (FDS) have become one of the main sources of air, soil, and water pollution with strong health, economic, and legal impacts (Semarnat, 2020; Sedema, 2021). Instead, open-air landfills predominate without any type of control.

The lack of space for the final deposit of waste generated in cities, the high costs of managing these sites (Kaza *et al.*, 2018), and their poor operation (Gill *et al.,* 2019) contribute to the formation of illegal dumps. An illegal dump is a place that, without environmental considerations, is chosen by a group of people to deposit their waste without obtaining permission from the competent authority (Karimi and Richter, 2022). Likewise, the deficiency of waste collection systems and the involvement of uncontrolled private agents cause the proliferation of illegal dumps on the outskirts of urban centers (Ferronato *et al.*, 2021; Vu *et al.*, 2019). Rapid industrial and urban growth during the last century has also contributed to the emergence of illegal dumps, making them a central environmental problem in all developed and developing countries (You *et al.*, 2020; Karimi and Richter, 2022; Silvestri and Omri, 2008).

Mexico reports that 16% of the MSW generated is not collected (Semarnat, 2020), so a fraction of it could end up in illegal dumps. These illegal dumps increased by 10.8% between 2019 and 2021 in Mexico City (Sedema, 2021), and they are generally found on public roads, common areas, parks, vacant lots, ravines, and streams. The State of Mexico illegally deposes of more than 220 tons/day of waste in 50 landfills and pours nearly 5 thousand tons on the metropolitan area's periphery (Rodríguez, 2023).

Illegal dumps are a source of air, soil, and water pollution (Karimi and Richter, 2022). The pollutants generated in these sites are mobilized and dispersed in broad areas due to topographic conditions, drainage systems, and soil types (Angelino *et al.*, 2018; Glanville and Chang, 2015). Illegal dumps pollute due to the emission of greenhouse gases generated by decomposing the organic fraction of the dumped waste (Cusworth *et al.*, 2020). They also pollute by the dissolution of solid polluting materials, such as heavy metals, by meteoric water that falls on them and its subsequent integration into surface or underground water currents (Padubidri *et al.,* 2022; Mahmood *et al.*, 2023). The spread of biological-infectious agents and the proliferation of harmful fauna directly affect the population living near illegal dumps (Mahmood *et al.*, 2023; Cusworth *et al.*, 2020).

Dumps not only cause environmental damage but also impose significant financial burdens on managers: in Australia, the management of more than nine thousand tons of illegally dumped waste in 2011-2012 cost more than 6 million dollars (Glanville and Chang, 2015); the Chartered Institution of Wastes Management estimated that the associated annual costs with the detection and remediation of illegal dumps in the United Kingdom were 100 to 150 million pounds in 2014 (Karimi and Richter, 2022) and the costs associated with their remediation in the USA were estimated at over $10 million dollars in 2016 (Quesada-Ruiz *et al.*, 2019). This underscores the urgent need for effective strategies to combat this issue.

Illegal dumps are not easily detected and frequently are in hidden locations or areas with low vehicular or pedestrian traffic. The detection of these sites is not only costly but also ineffective (Karimi and Richter, 2022). Identifying them is the crucial first step for their control and remediation, but the lack of public data in this regard presents a significant challenge (You *et al.*, 2020). Various methods exist to identify these dumps, with field surveys being the most effective but also the most expensive (Karimi and Richter, 2022; Angelino *et al.*, 2018). This complexity underscores the need for interdisciplinary collaboration and innovative solutions.

A relevant line of research consists of developing predictive models to identify probable areas of illegal dumps based on quantitative and qualitative information on the natural and social characteristics of the area (Matsumoto and Takeuchi, 2011). Glanville and Chang (2015) suggest a binary logistic regression model to identify explanatory variables, such as population density, land use, nearby waste facilities, and accessibility, that help predict the distribution of illegal dumps processed through geographic information systems (GIS). Additionally, methods that include social participation and the use of collective intelligence through social networks and applications, which provide georeferenced databases to influence authorities (Torres *et al.*, 2021), are explored. However, these approaches may have biases and inaccuracies, and their effectiveness on a large scale may be limited because they do not always reflect recent changes in landfills and the environment.

Satellite remote sensing, a precise and effective tool, is instrumental in detecting illegal dumps in large geographical areas (Shrivastava *et al.,* 2015; Niu *et al.,* 2023). It offers an alternative for analyzing relevant natural environmental features, which can be complemented with GIS (Ahmed *et al.,* 2006). This technology allows the spectral reflectance of topographic features to be recorded at visible and infrared wavelengths. The temperature difference at the earth's surface has been used to detect illegal waste dumping activities (Gill *et al.,* 2019; Yan, 2014) since areas where organic waste decomposes tend to have higher temperatures due to exothermic chemical reactions (Mahmood *et al.*, 2023). Gill *et al.* (2019) used Landsat images to identify regions most prone to waste discharge, with an overall accuracy of 72%. Mahmood *et al.* (2023) evaluated thermal energy release in FDSs through remote sensing, measuring thermal anomalies with Landsat 8 images.

Silvestri and Omri (2008) and Karimi and Richter (2022) used remote sensing with Landsat-8 and Suomi NPP satellite images to develop multi-criteria decision models to identify illegal dumps in large areas. Their model included spectral indices and surface temperatures as quantitative decision variables.

A common way to identify illegal dumps and monitor areas impacted by waste is by using spectral indices, although mainly as another decision variable (Mahmood *et al.*, 2023). One of the most useful indices is the Normalized Difference Vegetation Index (NDVI), which evaluates the health and vegetation cover, revealing areas affected by waste. Although effective, NDVI may be less reliable over large areas where diversified environmental factors influence the index (Niu et al., 2023). The Normalized Difference Water Index (NDWI) and its modified variant (MNDWI) are crucial for identifying water accumulations and wet areas impacted by dumps, which is critical to understanding environmental damage. However, these indices face challenges in urbanized areas where they can be confused with built infrastructure (Papale *et al.*, 2023; Xu, 2006).

The Normalized Difference Urbanized Index (NDBI) distinguishes urbanized areas with high-resolution data from the Sentinel-2A satellite, highlighting the need to carefully select indices and bands for accurate identifications (Vigneshwaran and Kumar, 2018; Xi *et al.*, 2019). The Bare Soil Index (BSI) is essential for detecting areas without vegetation cover using near-infrared and shortwave wavelengths, which can indicate the presence of dumps (Nguyen et al., 2021). Finally, the Soil Adjusted Vegetation Index (SAVI) adjusts soil brightness, improving the accuracy of vegetation measurement. Its use together with NDWI and MNDWI has proven effective in assessing erosion and deposition in places such as the Ping River in Thailand, highlighting its ability to improve environmental monitoring strategies (Laonamsai *et al*., 2023). The studies reviewed have shown the usefulness of spectral indices to detect illegal dumps; however, they have been used as another variable in decision models, which constitutes a limited application of them.

Artificial intelligence, a cutting-edge technology, has been effectively used to analyze satellite images of the earth's surface; one of its potential uses is the identification of illegal dumps. Niu *et al*. (2023) proposed a deep learning model for FDS mapping based on remote sensing of high spatial resolution images (0.5 meters) and a dual-stream deep neural network. Their model achieved high performance with an average accuracy of 90.62% and could be used to recognize landfills in large-scale regions. Dabholkar *et al*. (2017) also applied artificial intelligence, particularly neural networks, to analyze high-resolution satellite images to detect illegal dumps.

In addition to the neural networks used by Niu et al. (2023) and Dabholkar et al. (2017), supervised classification algorithms can be used for remote sensing of illegal dumping sites, as done by Perumal and Bhaskaran (2010) and Wang *et al.* (2024). Random forests is an artificial intelligence algorithm used for classifying remotely sensed images and outperforms individual decision tree classifiers (Zhou *et al.,* 2021). Also, it performs identically or better than several advanced pattern recognizers, such as artificial neural networks, support vector machines, and bagging and boosting methods (Shi and Yang, 2016; Zhou *et al.,* 2021).

Remote sensing can provide crucial information for identifying contaminated sites, but there are few rigorously validated approaches (Silvestri and Omri, 2008). This work addresses the use of remote sensing with multispectral images taken from the Sentinel-2 satellite to analyze the earth's surface and, using a supervised classification algorithm, identify probable illegal dumps of solid waste in hydrographic subbasins surrounding the metropolitan area of Oaxaca, Mexico. A classifier model was trained with random forests from a set of control polygons on known surfaces with waste in the open. The study area's image collection, processing, and supervised classification were done on the Google Earth Engine platform. The results were verified in the office using geographic information systems and, in the field, taking a sample of the probable illegal dumps identified in the office.

2. Materials and methods

This section describes the five stages used to identify probable illegal dumps in hydrographic subbasins bordering the Metropolitan Area of Oaxaca, Mexico. The diagram in Figure 1 shows the research process used in this work.

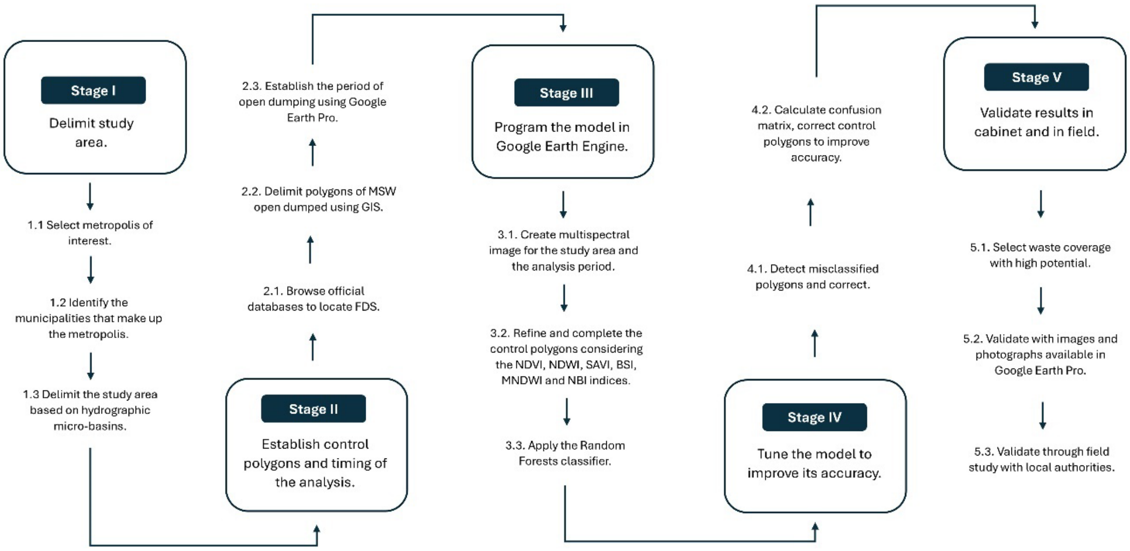

Figure 1. Research process to identify potential illegal dumps.

2.1. Stage I. Delimit study area

This research analyzed the metropolitan area of Oaxaca, which is one of the priority regions to address the problem of MSW since it represents the problem faced throughout Mexico. The area is one of the distributed MSW management systems studied by the National Research and Advocacy Project (Pronaii, by its acronym in Spanish) called "Transdisciplinary Strategy for Research and Resolution of the National Problem of Urban Solid Waste" in Mexico. This project aims to contribute to the transformation of solid waste management systems with a supportive, social, inclusive and circular economy approach.

The limit of the study area was carried out considering geohydrological and not administrative criteria. A delimitation of hydrographic micro-basins bordering the metropolitan area of Oaxaca was used. The hydrographic basin as a basic planning unit allows for the quantification of some natural and anthropic processes, such as the hydrological cycle, soil erosion processes, and dispersion of pollutants and nutrients, among others.

Figure 2 shows the three micro-basins surrounding the metropolitan area of Oaxaca that were analyzed to identify illegal dumps. The polygonal product of hydrographic basins used in this study is HydroBASINS (Lehner and Grill, 2013) on a global scale. This product was developed on behalf of the United States World Wildlife Fund with the support of international organizations and is part of the HydroSHEDS database (Hydrographic data and maps based on Shuttle Elevation Derivatives).

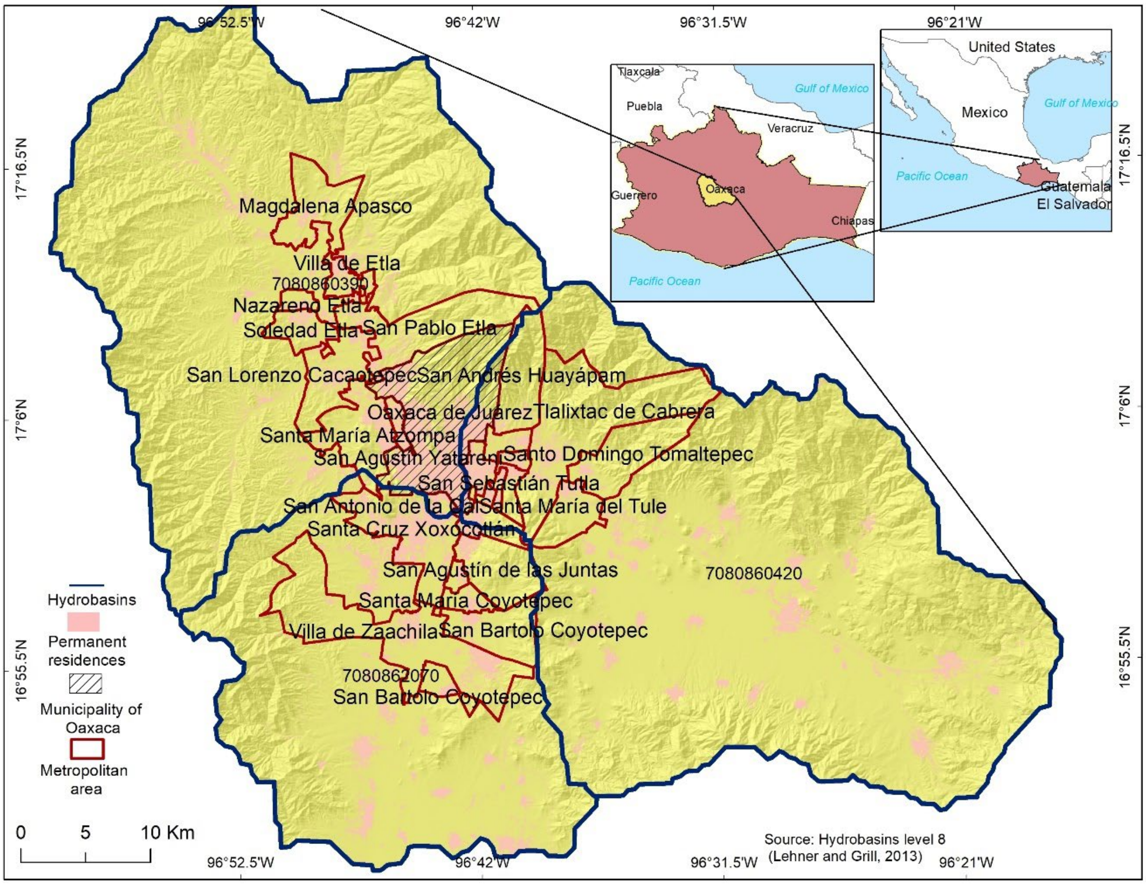

Figure 2. Microbasins of the metropolitan area of Oaxaca.

2.2. Stage II. Establish control polygons and timing of the analysis

The determination of official FDS that could contain surfaces covered by MSW deposited in the open was derived from the census of municipal governments and territorial demarcations of Mexico City in 2021 (INEGI, 2022). Thirty-nine FDS were identified, located, and analyzed using ArcGIS Pro to create the vector layer with the geolocation of the official FDS.

Photo interpretation work of the satellite images in the Google Earth Pro software identified surfaces covered with MSW deposited in the open between December 15, 2022, and January 15, 2023. These surfaces served as initial control polygons to train the supervised classification model and consisted of fifty-six control polygons for MSW coverage. This stage resulted in a vector layer of polygons covered by open-air MSW and a period of analysis.

2.3. Stage III. Program the model in Google Earth Engine

The supervised classification model of the multispectral images of the study area was programmed using Google Earth Engine. This online platform offers unparalleled access to an extensive catalog hosting millions of satellite images such as Sentinel-1, -2, -3, and -5P. This tool can also process large volumes of data efficiently and quickly in the cloud, which is of enormous value considering the size of the study area.

Figure 3 shows the process flow diagram of the model based on the methodologies presented in the End-to-End Google Earth Engine course (Gandhi, 2021) to identify probable illegal dumps used to process the images acquired from the Sentinel-2 MSI sensor, calculate spectral indices, and perform supervised classification. The first part of the model allows access to vector resources on the location of control polygons and delimitation of the study area. Next, multispectral images from the COPERNICUS_S2_SR collection of the Sentinel-2 sensor were acquired and preprocessed. Image preprocessing includes corrections to remove atmospheric effects, geometric and radiometric correction, and image harmonization to ensure consistency between images from different dates and paths. The choice of these images and their proper processing are essential to obtain reliable data.

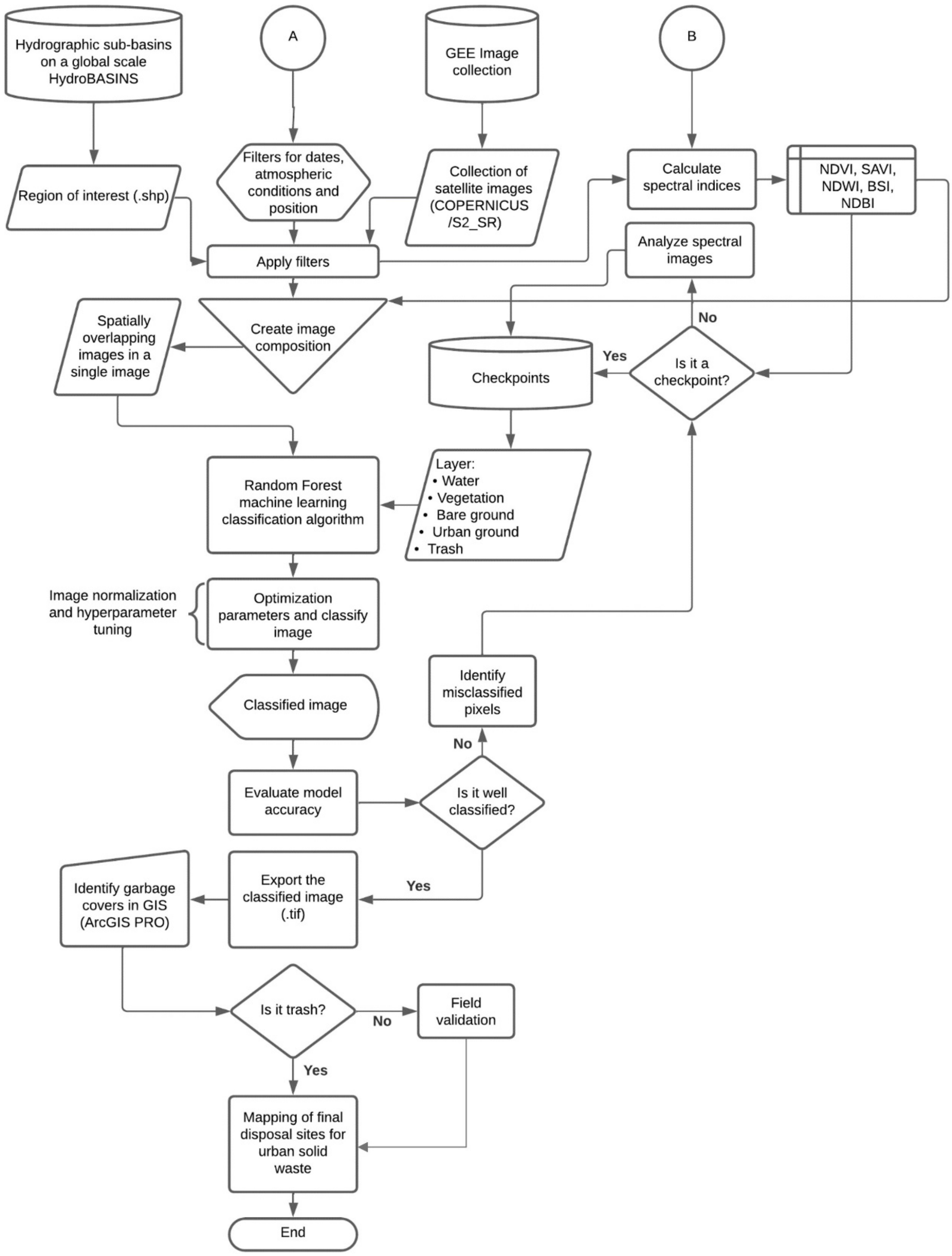

Figure 3. Flowchart of the supervised classification model in Google Earth Engine. Source: Own elaboration based on the supervised classification model proposed by Gandhi (2021).

The next block of the model uses spectral indices for the identification and differentiation of land cover. Land cover is a biophysical indicator that describes the materials that cover the territory, although it is understood differently depending on the discipline, the joint use of the concepts of land use and land cover is very common (Torres *et al.*, 2023). This study established five categories of land cover:

1) built structures or urban coverage; 2) bare soil or soil devoid of vegetation; 3) bodies of water; 4) surfaces with vegetation cover; and 5) surfaces covered with MSW.

In this part, the spectral indices NDVI, MNDWI, SAVI, NDBI, and BSI are calculated, with which the control polygons were precisely defined. Figure 4 shows the spectral image of the study area. Each of these indices provides critical information for the analysis, such as the amount and health of vegetation, the detection of the presence of water, the identification of areas with vegetation, and the characterization of areas without vegetation cover. The analysis of spectral indices is presented as an essential component of the procedure since it allows the precise delineation of specific characteristics on the Earth's surface, which is essential for the subsequent identification of probable illegal dumps.

The random forest classifier was the supervised machine learning algorithm used for classification and regression tasks in this work. This classifier integrated fifty decision trees into the forest. This classifier used 70% of the polygons for training and the remaining 30% for validation tests. The decision criteria were the values of the 13 bands of the optical images, the spectral indices NDVI, MNDWI, SAVI, NDBI, and BSI, and the slope and altitude of the surface.

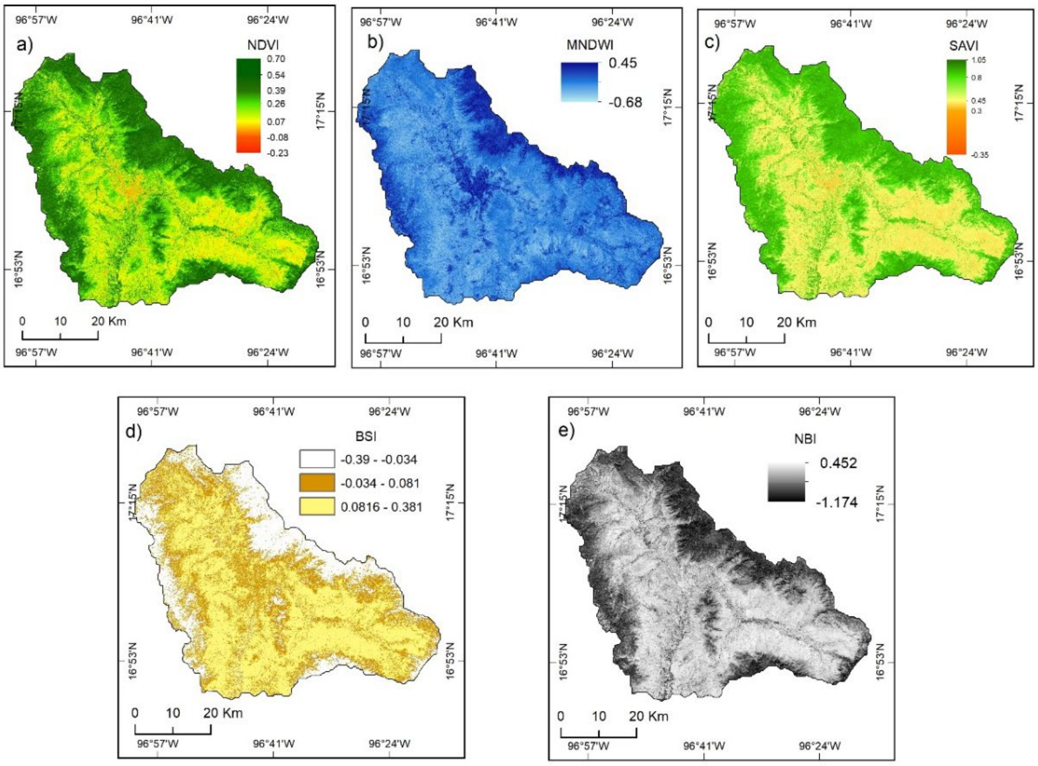

Figure 4. Images of the spectral indices in the study area: a) NDVI, b) MNDWI, c) SAVI, d) BSI and e) NBI.

2.4. Stage IV. Tune the model to improve its accuracy

After classification, model accuracy was evaluated using a confusion matrix, and the misclassified polygons were determined. The validation of supervised classification accuracy was executed through a Confusion Matrix, like an overall accuracy metric. Depending on the results obtained, the polygons were reclassified, eliminated, or others were included, using information from the spectral indices and satellite images available in Google Earth Engine. This run, fix, tune, and test loop consumed considerable research time. Due to this iterative process of improving accuracy, it was not considered necessary to use other metrics, such as ROC or the McNemar test, to validate the cartographic products.

2.5. Stage V. Validate results in cabinet and in field

The classification results were subjected to a rigorous validation process that included both cabinet photo interpretation techniques and field validation. The main objective of this validation was to ensure the accuracy of the classification. The validation in the cabinet was based on photo interpretation techniques to make a visual and detailed evaluation of the areas classified as illegal dumps. Figure 5a shows the supervised classification results on a property in the study area. Pixels classified as waste are shown in black color. Figure 5b shows the Google Earth Pro satellite image in which characteristic features of a surface covered with irregularly deposited waste are detected. Finally, Figure 5c shows a street-view photograph that seems to confirm the accumulation of waste in the area. The polygons with the greatest remote evidence were presented to the municipal authority of Oaxaca de Juárez for field validation.

From this stage, a map was obtained with the location of probable illegal dumps, with which the field visit was planned with the endorsement of the corresponding municipal and state authorities. The field validation consisted of visiting the polygons with the highest probability of being illegal dumps at the end of October 2023. To carry out this visit, an approach was made with the competent authorities, to whom the information obtained was presented and accompaniment was requested to tour their territory.

Figure 5. Validation process by photointerpretation.

3. Results and discussion

From a 5x5 confusion matrix shown in Table 1, it was possible to determine the performance of the supervised classification model. The matrix took the coverage established at the checkpoints and compared it to the classification assigned by Random Forests. The overall accuracy for the study area of Oaxaca was 0.9517, which indicates that the classifier's performance is good. Although the results in Table 1 show that the model is not capable of accurately distinguishing between urban soil, bare soil, and garbage, it does not mean that the classifier selected was the wrong one. Future works may use other classifiers to discriminate the classes, for example, Maximum Likelihood Estimation or Support Vector Machines, which were already used by Torres *et al.* (2023) in land cover classification. These classifiers could improve the performance of the current model. However, as Perumal and Bhaskaran (2010) point out, it is impossible to state only one classifier for all situations since the characteristics of each image and the circumstances of each study vary significantly (Table 1).

Table 1. Confusion matrix of the classification model.

	Urban	Bare ground	Water	Vegetal	Waste
Urban	354	8	0	6	0
Bare ground	73	950	0	1	8
Water	8	0	250	0	0
Vegetal	2	3	0	1320	0
Waste	10	20	0	2	68

The classifier's confusion between bare soil and vegetation can be explained by the seasonal changes in vegetation cover during the analysis period. The model's confusion in classifying urban coverage as water confirms what Xu (2006) points out: that the MNDWI works poorly in urbanized areas, confusing constructions with bodies of water.

More relevant to the objective of this work is the model's confusion when predicting the surfaces covered with waste. Waste is a mixture of various materials, so obtaining a spectral fingerprint of it is difficult. The spectral fingerprint depends on the composition of the waste at the selected control points. The most significant confusion in the model is between waste cover and bare ground. Despite the use of the BSI index as proposed by Nguyen *et al.* (2021), this classification error was not possible to avoid. Therefore, more research is necessary on the spectral fingerprint of urban solid waste.

The graph in Figure 6 shows the relative importance of the nineteen decision criteria for classifying land cover. As can be seen, the value of the NDVI and SAVI spectral indices, as well as the B1 spectral band of aerosols, is the most important characteristic when making the classification. The importance that the model gives to the spectral indices is consistent with what Mahmood *et al.* (2023) stated regarding their usefulness in detecting illegal dumps. The NDVI and SAVI indices had already been identified by Karimi and Richter (2022) as relevant metrics for remote sensing of these dumpsites.

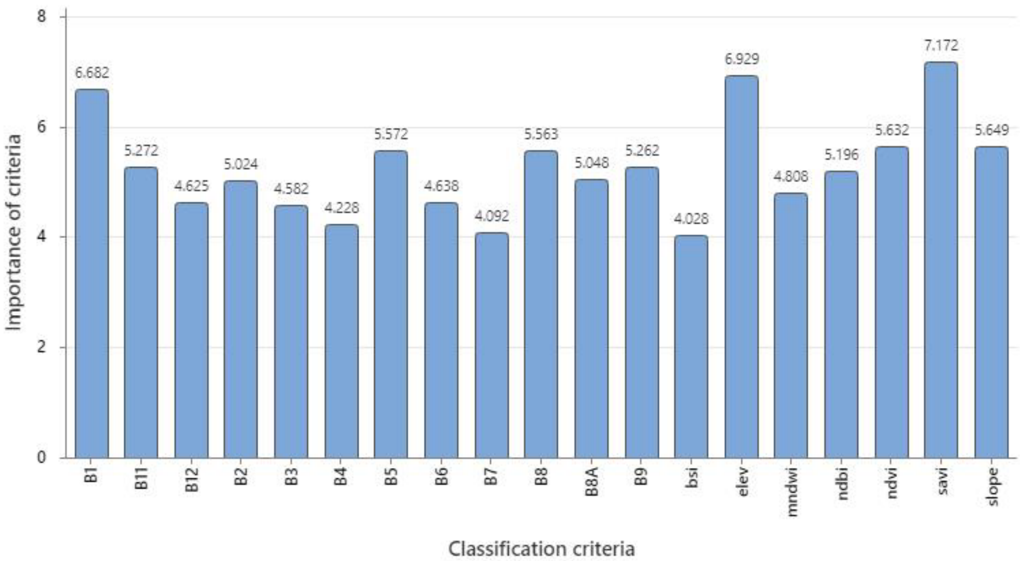

Figure 6. Importance of criteria in supervised classification.

Understandably, NDVI is a vital criterion for vegetal cover, bare ground, and even water bodies. Being a simple indicator of photosynthetically active biomass, the NDVI helps soil classification with high degrees of certainty, as Silvestri and Omri (2008) found. The SAVI helps mitigate the impact of ground shine, so its importance in classification is also consistent with the literature review. The SAVI index improves the precision of identifying areas affected by erosive processes, as Laonamsai *et al.* (2023) suggested.

The B1 band of aerosols stands out as a significant characteristic in the classification, and this result invites us to investigate the spectral fingerprint of solid waste further. Bands B4, B5, B6, and B7 were the least important in classifying. These bands are associated with red, infrared, and near-infrared, an optical characteristic of little relevance among the five coverages analyzed in this work. Future models could remove these bands.

The graph in Figure 7 shows the mean, maximum, and minimum values of the pixels classified based on the criteria in Figure 6. The graph 7a shows the spectral fingerprint of the urban coverage for 363 data points. The adjustment of the NDBI index for this type of coverage is confirmed, as is the usefulness of the NDVI and the SAVI, whose values show a dispersion of less than 5%. Graph 7b shows the spectral fingerprint for 1187 bare ground data points. The adjustment of the BSI, NDBI, NDVI, and SAVI indices for this coverage stands out.

Graph 7c shows the spectral footprint of residues for 94 data points. For this cover, the NDVI, NDBI and SAVI indices present more significant adjustments, along with the bands B03, B11 and B12. The adjustment of the NDVI and SAVI spectral indices shown in the figure confirms what Karimi and Richter (2022) point out regarding the reduction of vegetation cover due to large-scale soil disturbances caused by residues. As shown in Figure 7c, the values of these indices are lower.

The similarity between the spectral fingerprints of urban cover, bare ground, and waste should be appreciated, which partially explains the classifier's confusion and the difficulty of finding illegal dumps.

The graphs 7d and 7e show the spectral fingerprint of the vegetation cover and water bodies, respectively. The 1030 vegetative cover data points analysis permitted to represent the spectral fingerprint that fits well to the first six bands. In general, the model distinguished wellvegetation and bodies of water cover. The analysis of 258 data points predicted by the model built the clear spectral fingerprint of the water.

Graph 7c is an example of how remote sensing can be used as a discrimination tool to recognize the unique spectral reflectance of clandestine dumpsites. This graph is the first approximation of the spectral fingerprint that, according to Karimi and Richter (2022), is of capital importance to delve deeper into the topic of detection and control of illegal dumpsites. Graph 7c is the first approximation of the spectral fingerprint of waste, but future work should delve into the topic, as well as the construction of a spectral index of waste.

Figure 8 shows the supervised classification results in the study area: urban cover represents 2.7%, 31.5% is bare ground, and 65.4% is vegetal cover. Only 200 hectares in the territory could correspond to bodies of water, which is consistent with the climate of the study area. Around 288 thousand 100 hectares of land surface were analyzed, supporting what Shrivastava *et al.* (2015) and Niu *et al*. (2023) pointed out, that satellite remote sensing is a powerful tool for analyzing large geographic areas.

Regarding the surface classified as covered with waste, the model produced a result of 400 hectares, which represents only 0.14% of the surface. This area is equivalent to fifteen times the size of the Zaachila Landfill, which served the municipality of Oaxaca de Juárez. This area was explored in the cabinet through photo interpretation to discard those areas with low potential to be illegal dumps.

The photo interpretation work after supervised classification was arduous because the model yields a large surface with probable waste coverage. Of the total area classified with waste coverage, only 4% had sufficient evidence in satellite images to establish probable illegal dumps. At this stage, using predictive models could further reduce the areas to be explored, such as the one proposed in Glanville and Chang (2015), which considers variables such as population density, land use, nearby waste facilities, and access roads; this improvement would make the process of detecting illegal dumps more efficient.

Figure 7. Spectral footprint of the types of coverage: a) urban, b) bare ground, c) waste, d) vegetal and e) water.

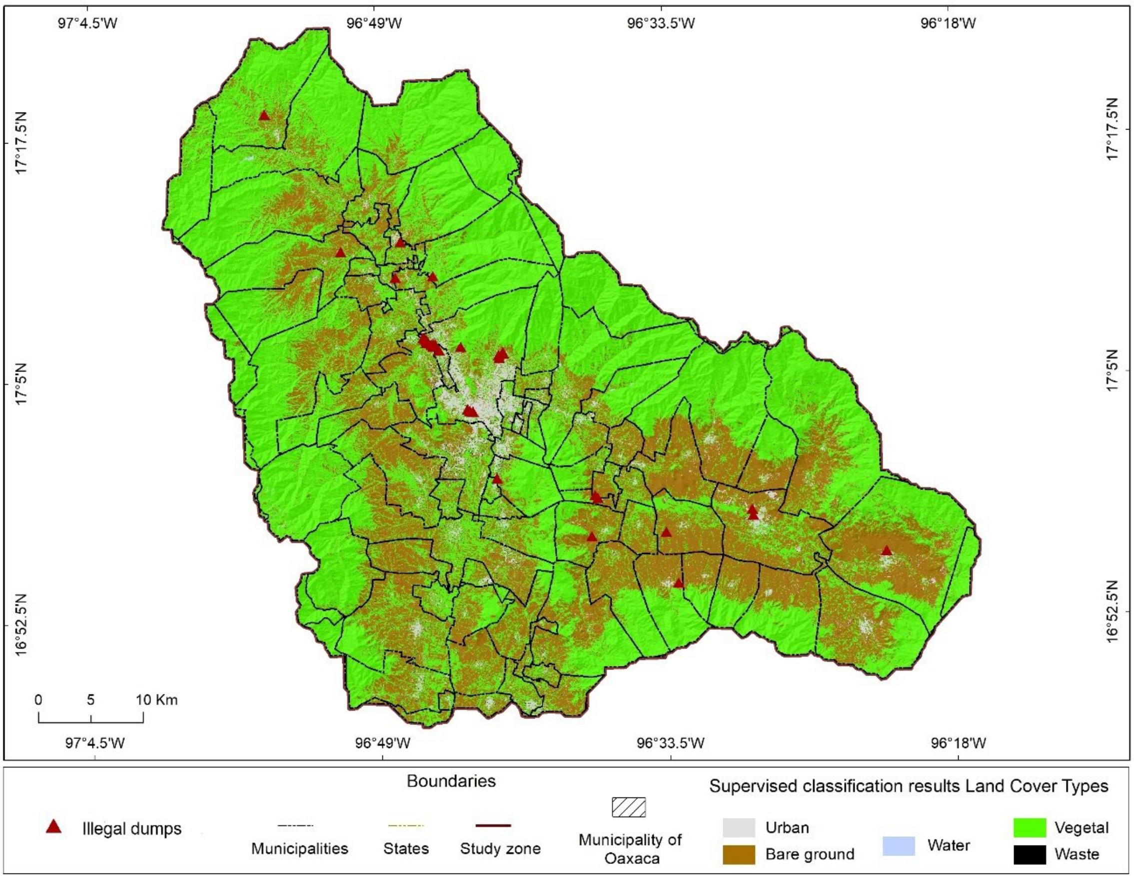

Figure 8. Results of supervised classification in the study area and location of probable illegal dumping sites.

In the end, 32 probable illegal dumps were detected, which cover approximately seventeen hectares (Fig. 8). Fifteen of these sites are in the municipality of Oaxaca de Juárez or its surroundings. This number is understandable, given that Oaxaca de Juárez has the largest population of the twenty-four municipalities that comprise the Metropolitan Zone of Oaxaca. Figures 9, 10, 11, and 12 show the result of supervised classification and the five land cover categories with the color indicated on the legend of each image. As a background, the available satellite image does not correspond to the model's analysis period. So, this produced several discrepancies between the coverage predicted by the model and the background image. Also, these figures illustrate the model confusion discussed in previous paragraphs.

The thirty-two probable illegal dumps could not be corroborated in the field since it was impossible to obtain the support of the municipal or state authorities of the region. Only the city council of Oaxaca de Juárez provided the conditions to tour its territory and carry out the field validation of seventeen sites. In this tour, the team discarded some probable illegal dumps; for example, sites 29 and 32 were banks of stone materials instead of waste, and their reflectance was like that of MSW. Site 30 contained construction and demolition waste, but this waste was on private property, so it cannot be considered a dump itself. At site 31, there was no trace of a waste deposit, and at site 28, only remains of burned waste were found.

The rocks identified at sites 29 and 32 showed hydrothermal alteration. During hydrothermalism, water reacts with the original rock-forming minerals, such as feldspars, pyroxenes, and amphiboles, to form micas and clays (Frost and Frost, 2019). Rocks display characteristic spectral signatures in multispectral and hyperspectral RS systems based on their mineralogy and texture (Girijaa and Mayappana, 2019). Girijaa and Mayappana (2019) report that it is possible to identify minerals formed by hydrothermal alteration from band ratios (BR) and Principal component analysis (PCA). For the same purpose, Abdel-Rahman (2023) used selective band ratios (BR), directed principal component analysis (DPCA), feature-oriented false color composites (FFCC), and constrained Energy Minimization (CEM) using ASTER and Sentinel 2 data. Using BR and PCA can reduce the confusion model between residues and altered rocks in future studies.

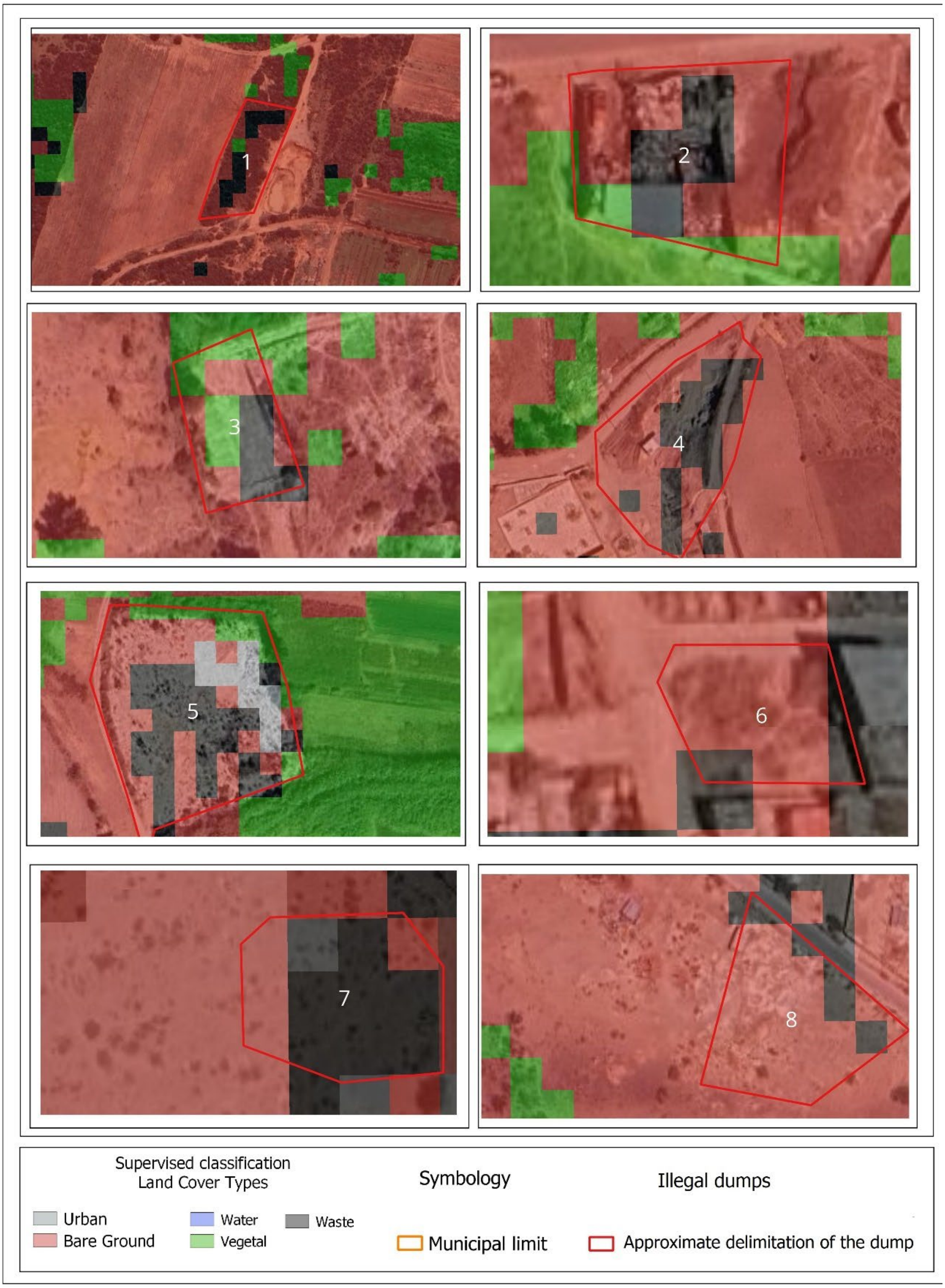

Figure 9. Probable illegal dumps 1-8.

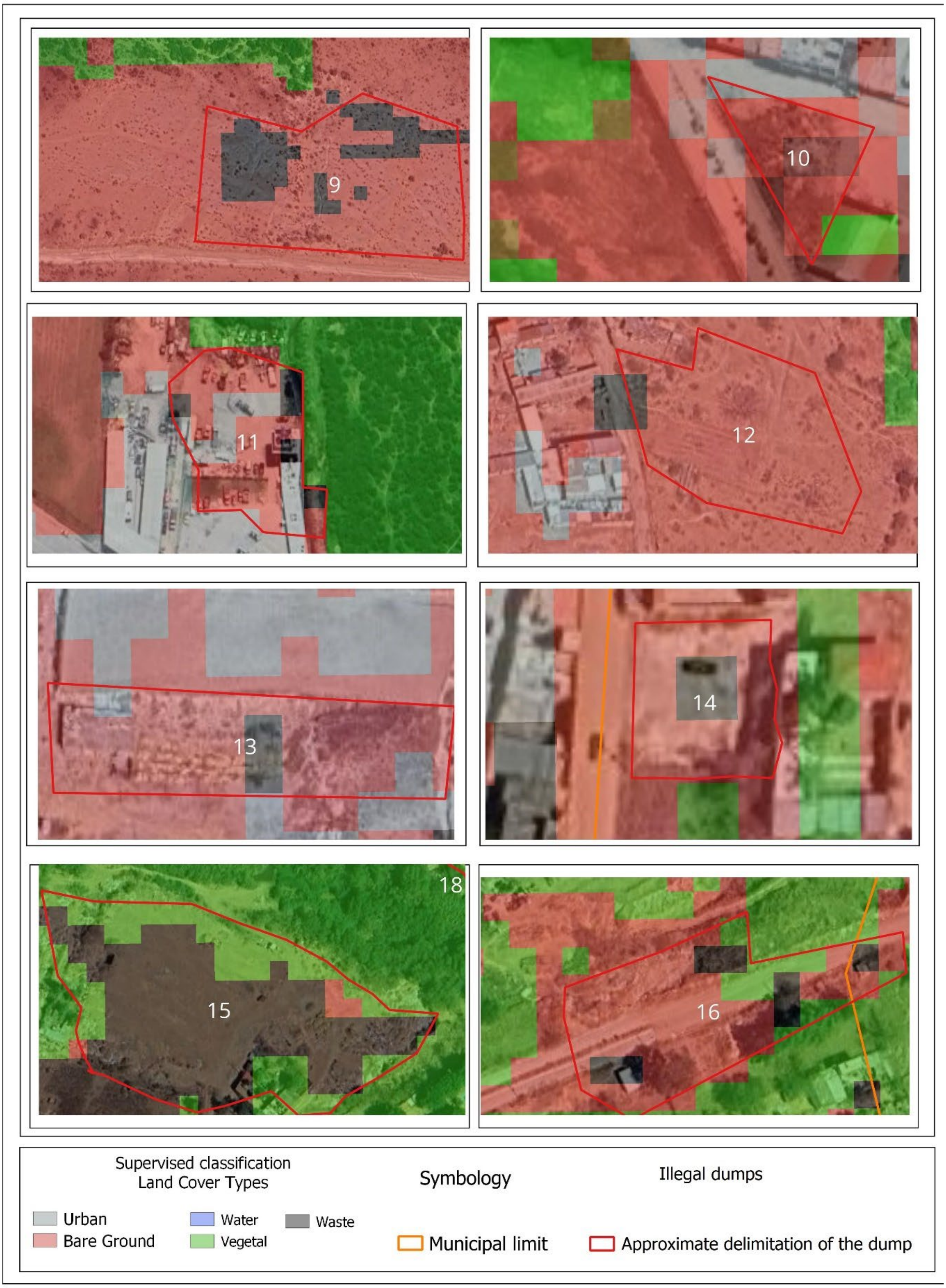

Figure 10. Probable illegal dumps 9-16.

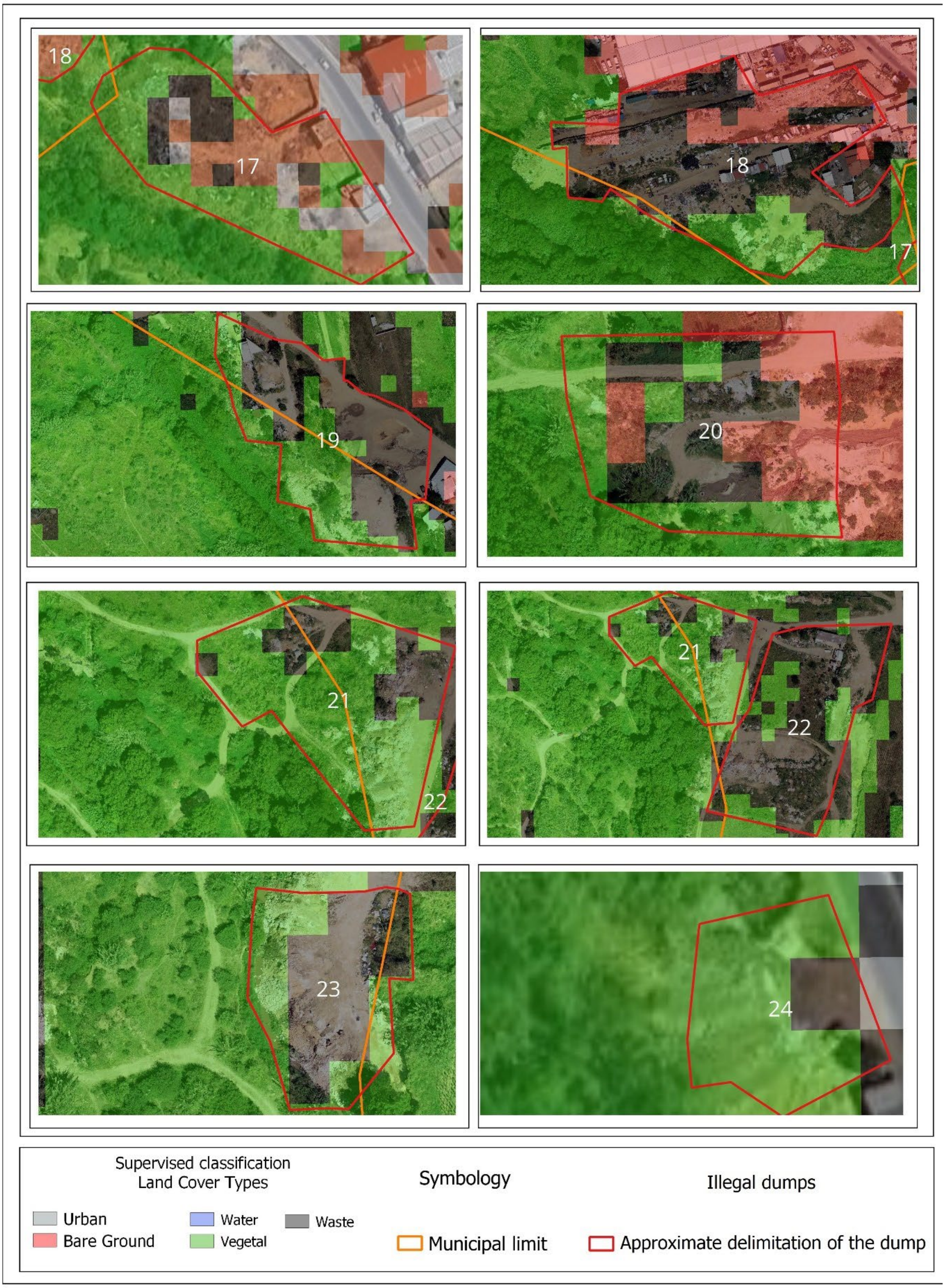

Figure 11. Probable illegal dumps 17-24.

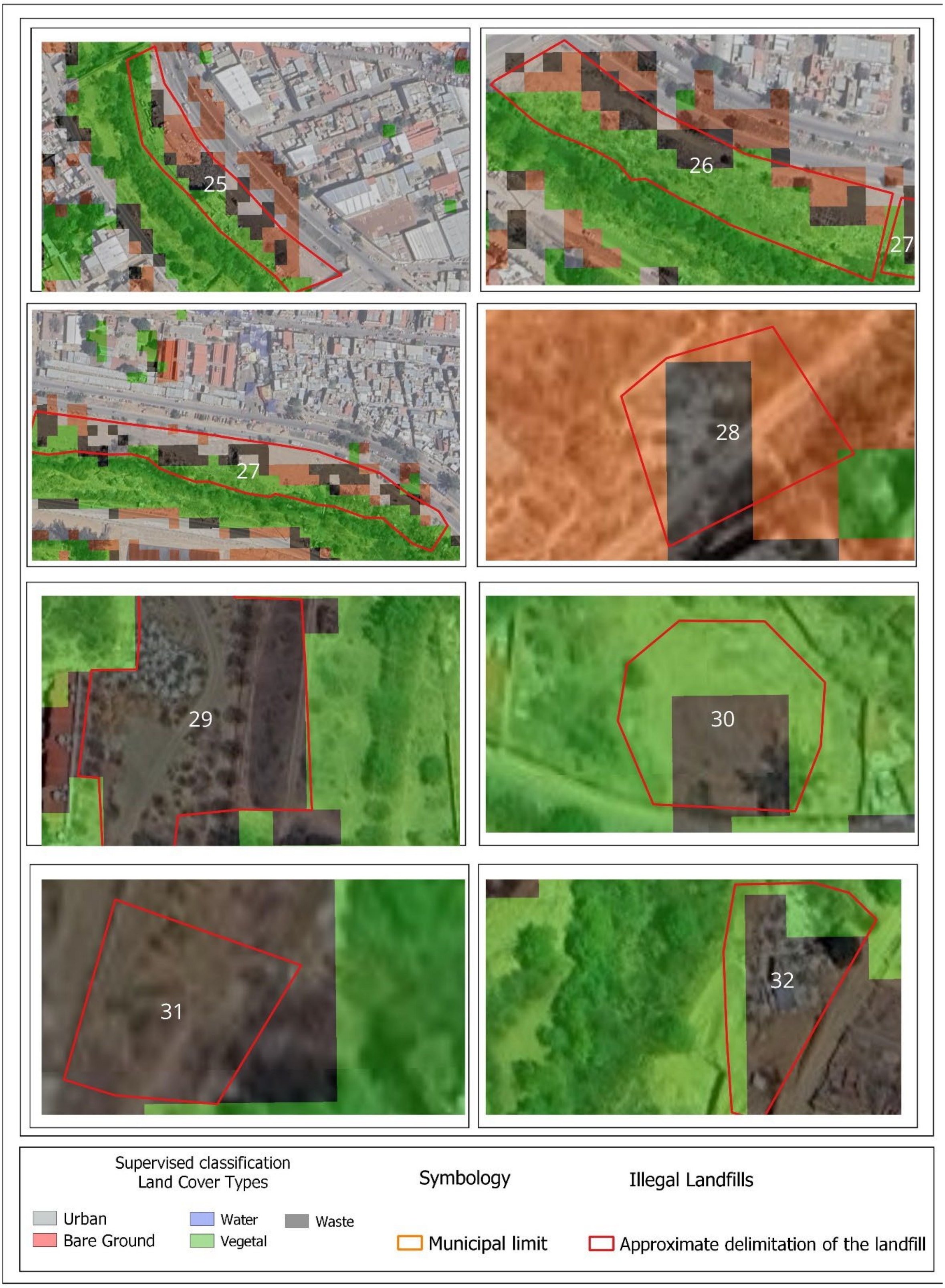

Figure 12. Probable illegal dumps 25-32.

In twelve properties identified as probable illegal dumps, there was evidence of irregular solid waste deposits. Fieldwork confirmed the possible illegal dumps 15, 16, 17, 18, 19, 20, 21, 22,23, 25, 26, and 27. Some of the background images of these sites correspond to an orthophoto obtained with a drone flight during the field validation carried out between October 24 and 26, 2023. The area of validation was the bank of the Atoyac River was poses a high risk of contamination for this body of water. People regularly and clandestinely dump urban solids and special handling waste in this area.

Future research should review the impacts of irregular discharge on the banks of the Atoyac River in depth. Unfortunately, the lack of resources allocated to monitoring and controlling waste dumping causes clandestine practices with strong environmental impacts. The results of this work required arduous validation work in the office but less than traditional classification methods, which require expensive and time-consuming field studies, as Perumal and Bhaskaran (2010) pointed out. Using collective intelligence and social participation can improve the validation process in the field (Torres *et al.*, 2021).

This work improved classification models by using spectral indices as relevant characteristics. Although Niu *et al.* (2023) pointed out that the effectiveness of these indices decreases in large territories, the results of this work show that the post-processing techniques can correct the effectiveness of discarding poorly classified areas.

The model presented in this investigation does not reach the level of precision of other models such as those developed by Gill *et al.* (2019) and Niu *et al.* (2023). However, one of its benefits is the possibility of exploring large geographical areas, and that in combination with photointerpretation and GIS techniques, it can successfully detect illegal dumps.

It is important to mention that other artificial intelligence algorithms, such as artificial neural networks, have been used, obtaining greater precision (Gill *et al.*, 2019; Niu *et al*., 2023; Dabholkar *et al*., 2017; Zhou *et al.,* 2021; Wang *et al.* (2024)). However, these models require high-resolution images, which limits their large-scale applicability.

An important line of research is to develop predictive models to identify probable areas of illegal dumping sites. Although this work provides relevant information for this, a model of this type usually ignores spatial heterogeneity and local contamination characteristics, so this work did not follow this line.

4. Conclusions

The traditional approach to garbage management in Mexico is inefficient and fails to fully protect the environment and the health of the population. Proof of this are the dozens of illegal dumps found in the study area. Technological tools such as the presented in this work can help managers and decision makers develop actions to control and mitigate pollution in these sites. The development of these models requires technical skills that could be beyond the reach of local authorities; however, collaboration between the academic sector and the authorities, as exemplified by the Pronaii in Mexico, can bridge this gap.

This work has shown how artificial intelligence, satellite remote sensing, and geographic information systems constitute powerful tools for identifying illegal dumps. The availability of satellite images over increasingly shorter periods and tools to process them, such as Google Earth Engine, can contribute to the development of remote sensing in solid waste management. This study is the one first approach to regional studies for detecting illegal dumps, through of the solid methodology that could replicated in other regions in Mexico.

Although the model presented does not achieve the level of precision of other models, such as those developed by Gill *et al.* (2019) and Niu *et al.* (2023), one of the benefits is that it can be expanded to explore large geographical areas and, combined with photointerpretation and GIS techniques, it can be successful in detecting illegal dumps. It should also be noted that other artificial intelligence algorithms such as artificial neural networks have been used, obtaining greater precision (Gill *et al.,* 2019; Niu *et al.,* 2023; Dabholkar *et al.,* 2017); however, these models require high-resolution images, which limits their large-scale applicability.

The surfaces covered with solid waste occupied as control polygons allowed us to explore the spectral fingerprint of the solid waste and identify the bands and indices that are characteristic of it. This is a contribution to the remote sensing of illegal dumps. Future work will be able to specify the spectral fingerprint of the waste, and even construct a spectral index for the detection of solid waste on a large scale.

Others research should focus on improving spectral techniques for waste and integrating social participation mechanisms to enhance the detection and management of illegal dumps.

Acknowledgements

Financial support for this study was provided by the National Council of Humanities, Sciences and Technologies (Conahcyt) of Mexico, through the National Research and Advocacy Project (Pronaii) with number 321379, entitled like "Transdisciplinary Strategy for Research and Resolution of the National Problem of Urban Solid Waste." Thanks to the efforts of the Center for Research and higher Studies in Social Anthropology (Ciesas) Golfo, the resources for the development of this research were made available. The authors thank Vicente Torres Rodríguez and Miguel Ángel Blancas Reza for their advice on technical and methodological aspects. Finally, we would like to thank the authorities of the municipality of Oaxaca de Juárez for their support.

References

Ahmed, S. M., Muhammad, H., Sivertun, A., 2006, September. Solid waste management planning using GIS and remote sensing technologies case study Aurangabad City, India. In *2006 International Conference on Advances in Space Technologies* (pp. 196-200). IEEE. https://doi.org/10.1109/ICAST.2006.313826

Angelino, C. V., Focareta, M., Parrilli, S., Cicala, L., Piacquadio, G., Meoli, G., De Mizio, M., 2018. A case study on the detection of illegal dumps with GIS and remote sensing images. In *Earth Resources and Environmental Remote Sensing/GIS Applications IX* (Vol. 10790, pp. 165-171). SPIE. https://doi.org/10.1117/12.2325557

Cusworth, D. H., Duren, R. M., Thorpe, A. K., Tseng, E., Thompson, D., Guha, A., Newman, S., Koster, K.T., Miller, C. E., 2020. Using remote sensing to detect, validate, and quantify methane emissions from California solid waste operations. *Environmental Research Letters* 15(5), 054012. https://doi.org/10.1088/1748-9326/ab7b99

Dabholkar, A., Muthiyan, B., Srinivasan, S., Ravi, S., Jeon, H., Gao, J., 2017. Smart illegal dumping detection. In *2017 IEEE Third International Conference on Big Data Computing Service and Applications (BigDataService)* (pp. 255-260). IEEE. https://doi.org/10.1109/BigDataService.2017.51

Ferronato, N., Portugal Alarcón, G.P., Guisbert Lizarazu, E.G., Torretta, V., 2021. Assessment of municipal solid waste collection in Bolivia: Perspectives for avoiding uncontrolled disposal and boosting waste recycling options. *Resources, Conservation and Recycling* 167, 105234. 2021. https://doi.org/10.1016/j.resconrec.2020.105234

Frost, B.R. Frost, C.D., 2019. *Essentials of Igneous and Metamorphic Petrology.* Second edition. Cambridge University Press, New York, 297 pages.

Gandhi, U. 2021. End-to-End Google Earth Engine Course. Spatial Thoughts. Available: https://courses.spatialthoughts.com/end-to-end-gee.html

Glanville, K., Chang, H. C., 2015. Mapping illegal domestic waste disposal potential to support waste management efforts in Queensland, Australia. *International Journal of Geographical Information Science*, 29(6), 1042-1058. https://doi.org/10.1080/13658816.2015.1008002

Gill, J., Faisal, K., Shaker, A., Yan, W. Y., 2019. Detection of waste dumping locations in landfill using multi-temporal Landsat thermal images. *Waste Management and Research*, 37(4), 386-393. https://doi.org/10.1177/0734242X18821808

Instituto Nacional de Estadística y Geografía (INEGI), 2021. *Censo Nacional de Gobiernos Municipales y Demarcaciones Territoriales/Delegacionales, 2021.* https://www.inegi.org.mx/programas/cngmd/2021/

Karimi, N., Ng, K. T. W., Richter, A., 2022. Development and application of an analytical framework for mapping probable illegal dumping sites using nighttime light imagery and various remote sensing indices. *Waste Management*, 143, 195-205. https://doi.org/10.1016/j.wasman.2022.02.031

Kaza, S., Yao, L., Bhada-Tata P., Van Woerden, F., 2018. *What a waste 2.0: a global snapshot of solid waste management to 2050.* Urban Development Series. World Bank Group, Washington, D.C. https://doi.org/10.1596/978-1-4648-1329-0

Laonamsai, J., Julphunthong, P., Saprathet, T., Kimmany, B., Ganchanasuragit, T., Chomcheawchan, P., Tomun, N., 2023. Utilizing NDWI, MNDWI, SAVI, WRI, and AWEI for Estimating Erosion and Deposition in Ping River in Thailand. *Hydrology* 10(3), 70. https://doi.org/10.3390/hydrology10030070

Lehner, B., Grill G. 2013. Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems. *Hydrological Processes* 27(15), 2171–2186. https://doi.org/10.1002/hyp.9740

Mahmood, K., Iftikhar, W., Faizi, F., 2023. Geospatial indices as an alternative for environmental impact assessment of dumped waste. *Acta Geophysica* 71(1), 309-322. https://doi.org/10.1007/s11600-022-00974-6

Matsumoto, S., Takeuchi, K., 2011. The effect of community characteristics on the frequency of illegal dumping. *Environmental Economics and Policy Studies* 13 (3), 177–193. https://doi.org/10.1007/s10018-011-0011-5

Nguyen, C. T., Chidthaisong, A., Kieu Diem, P., Huo, L. Z., 2021. A Modified Bare Soil Index to Identify Bare Land Features during Agricultural Fallow-Period in Southeast Asia Using Landsat 8. *Land* 10(3), 231. https://doi.org/10.3390/land10030231

Niu, B., Feng, Q., Yang, J., Chen, B., Gao, B., Liu, J., Li, Y., Gong, J., 2023. Solid waste mapping based on very high-resolution remote sensing imagery and a novel deep learning approach. *Geocarto International* 38(1), 2164361. https://doi.org/10.1080/10106049.2022.2164361

Papale, L. G., Guerrisi, G., De Santis, D., Schiavon, G., Del Frate, F., 2023. Satellite Data Potentialities in Solid Waste Landfill Monitoring: Review and Case Studies. *Sensors*, 23(8), 3917. https://doi.org/10.3390/s23083917

Padubidri, C., Kamilaris, A., Karatsiolis, S., 2022. Accurate Detection of Illegal Dumping Sites Using High Resolution Aerial Photography and Deep Learning. *2022 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops)*, Pisa, Italy, pp. 451-456. https://doi.org/10.1109/PerComWorkshops53856.2022.9767451

Perumal, K., Bhaskaran, R., 2010. Supervised classification performance of multispectral images. *arXiv preprint arXiv:1002.4046*. https://doi.org/10.48550/arXiv.1002.4046

Quesada-Ruiz, L. C., Rodriguez-Galiano, V., Jordá-Borrell, R., 2019. Characterization and mapping of illegal landfill potential occurrence in the Canary Islands. *Waste Management* 85, 506-518. https://doi.org/10.1016/j.wasman.2019.01.015

Girija, R. R., Mayappan, S., 2019. Mapping of mineral resources and lithological units: a review of remote sensing techniques. *International Journal of Image and Data Fusion* 10, 79–106. https://doi.org/10.1080/19479832.2019.1589585

Rodríguez, A., 2023. Edoméx cuenta con al menos 50 tiraderos clandestinos. *Capital Media*. https://www.capitaledomex.com.mx/local/edomex-cuenta-con-al-menos-50-tiraderos-clandestinos/ (último acceso: 28/05/2024)

Secretaría de Medio Ambiente y Recursos Naturales (Semarnat), 2020. *Diagnóstico básico para la gestión integral de los residuos. 2020.* https://www.gob.mx/cms/uploads/attachment/file/554385/DBGIR-15-mayo-2020.pdf

Secretaría del Medio Ambiente de la Ciudad de México (Sedema), 2021. *Inventario de Residuos Sólidos de la Ciudad de México 2021.* https://www.sedema.cdmx.gob.mx/storage/app/media/DGCPCA/residuos/InventariodeResiduosSolidos2021.pdf

Shi, D., Yang, X., 2016. An assessment of algorithmic parameters affecting image classification accuracy by random forests. *Photogrammetric Engineering and Remote Sensing* 82(6), 407-417. https://doi.org/10.14358/PERS.82.6.407

Shrivastava, P., Mishra, S., Katiyar, S. K., 2015. A review of solid waste management techniques using GIS and other technologies. In *2015 International conference on computational intelligence and communication networks (CICN)* (pp. 1456-1459). IEEE. https://doi.org/10.1109/CICN.2015.281

Silvestri, S., Omri, M., 2008. A method for the remote sensing identification of uncontrolled landfills: formulation and validation. *International Journal of Remote Sensing* 29 (4), 975-989. https://doi.org/10.1080/01431160701311317

Torres, R. G., Bezada, M., Santalla, I. R., 2023. Cartografía de cobertura del suelo mediante datos de teledetección en la planicie de desborde del río Apure (Venezuela). *Cuadernos de Investigación Geográfica (Geographical Research Letters)* 49(1), 113-137. http://doi.org/10.18172/cig.5607

Torres, R. N., Fraternali, P., Biscontini, A., 2021. On the Use of Class Activation Maps in Remote Sensing: the case of Illegal Landfills. *2021 IEEE 8th International Conference on Data Science and Advanced Analytics (DSAA).* Porto, Portugal, 2021, pp. 1-10, https://doi.org/10.1109/DSAA53316.2021.9564243

Vigneshwaran, S., Vasantha Kumar, S., 2018. Extraction of built-up area using high resolution sentinel-2a and Google satellite imagery. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-4/W9, 165–169. https://doi.org/10.5194/isprs-archives-xlii-4-w9-165-2018 .

Vu, H.L., Bolingbroke, D., Ng, K.T.W., Fallah, B., 2019. Assessment of waste characteristics and their impact on GIS vehicle collection route optimization using ANN waste forecasts. *Waste Management* 88, 118-130. https://doi.org/10.1016/j.wasman.2019.03.037

Wang, Y., Zou, B., Zuo, X., Zou, H., Zhang, B., Tian, R., Feng, H., 2024. A remote sensing analysis method for soil heavy metal pollution sources at site scale considering source-sink relationships. *Science of The Total Environment* 946, 174021. https://doi.org/10.1016/j.scitotenv.2024.174021

Xi, Y., Thinh, N. X., Li, C., 2019. Preliminary comparative assessment of various spectral indices for built-up land derived from Landsat-8 OLI and Sentinel-2A MSI imageries. *European Journal of Remote Sensing* 52(1), 240–252. https://doi.org/10.1080/22797254.2019.1584737

Xu, H., 2006. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International Journal of Remote Sensing*, 27(14), 3025–3033. https://doi.org/10.1080/01431160600589179

Yan, W. Y., Mahendrarajah, P., Shaker, A., Faisal, K., Luong, R., Al-Ahmad, M., 2014. Analysis of multi-temporal Landsat satellite images for monitoring land surface temperature of municipal solid waste disposal sites. *Environmental monitoring and assessment* 186, 8161-8173. https://doi.org/10.1007/s10661-014-3995-z

You, T., Chen, W., Wang, H., Yang, Y., Liu, X., 2020. Automatic Garbage Scattered Area Detection with Data Augmentation and Transfer Learning in SUAV Low-Altitude Remote Sensing Images. *Mathematical Problems in Engineering* 2020, 1-13. https://doi.org/10.1155/2020/7307629

Zhou, L., Luo, T., Du, M., Chen, Q., Liu, Y., Zhu, Y., He, C., Wang, S., Yang, K., 2021. Machine learning comparison and parameter setting methods for the detection of dump sites for construction and demolition waste using the google earth engine. *Remote Sensing* 13(4), 787. https://doi.org/10.3390/rs13040787