



Article

Enhanced Detection of Artisanal Small-Scale Mining with Spectral and Textural Segmentation of Landsat Time Series

Alejandro Fonseca ¹, Michael Thomas Marshall ^{2,*} and Suhyb Salama ³¹ IABG Industrieanlagen-Betriebsgesellschaft, Hermann-Reichelt-Str. 3, 01109 Dresden, Germany² Department of Natural Resources, Faculty of Geo-Information Science and Earth Observation, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands³ Department of Water Resources, Faculty of Geo-Information Science and Earth Observation, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands

* Correspondence: m.t.marshall@utwente.nl

Abstract: Artisanal small-scale mines (ASMs) in the Amazon Rainforest are an important cause of deforestation, forest degradation, biodiversity loss, sedimentation in rivers, and mercury emissions. Satellite image data are widely used in environmental decision-making to monitor changes in the land surface, but ASMs are difficult to map from space. ASMs are small, irregularly shaped, unevenly distributed, and confused (spectrally) with other land clearance types. To address this issue, we developed a reliable and efficient ASM detection method for the Tapajós River Basin of Brazil—an important gold mining region of the Amazon Rainforest. We enhanced detection in three key ways. First, we used the time-series segmentation (LandTrendr) Google Earth Engine (GEE) Application Programming Interface to map the pixel-wise trajectory of natural vegetation disturbance and recovery on an annual basis with a 2000 to 2019 Landsat image time series. Second, we segmented 26 textural features in addition to 5 spectral features to account for the high spatial heterogeneity in ASM pixels. Third, we trained and tested a Random Forest model to detect ASMs after eliminating irrelevant and redundant features with the Variable Selection Using Random Forests “ensemble of ensembles” technique. The out-of-bag error and overall accuracy of the final Random Forest was 3.73 and 92.6%, which are comparable to studies mapping large industrial mines with the normalized difference vegetation index (NDVI) and LandTrendr. The most important feature in our study was NDVI, followed by textural features in the near and shortwave infrared. Our work paves the way for future ASM regulation through large area monitoring from space with free and open-source GEE and operational satellites. Studies with sufficient computational resources can improve ASM monitoring with advanced sensors consisting of spectral narrow bands (Sentinel-2, Environmental Mapping and Analysis Program, PRecurso IperSpettrale della Missione Applicativa) and deep learning.

Keywords: deforestation; forest degradation; Amazon; remote sensing; machine learning; LandTrendr; Google Earth Engine



Citation: Fonseca, A.; Marshall, M.T.; Salama, S. Enhanced Detection of Artisanal Small-Scale Mining with Spectral and Textural Segmentation of Landsat Time Series. *Remote Sens.* **2024**, *16*, 1749. <https://doi.org/10.3390/rs16101749>

Academic Editors: Meisam Amani, Arsalan Ghorbanian, Sadegh Jamali, Feng Tian, Per-Ola Olsson and Torbern Tagesson

Received: 4 March 2024

Revised: 24 April 2024

Accepted: 26 April 2024

Published: 15 May 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Global deforestation has slowed in the past few decades because of widespread afforestation campaigns in high-income countries outside the tropics [1]. Forest loss in low-income tropical countries offsets the gains. Countries in South America have experienced the largest declines despite improved forest management and the expansion of protected areas [2]. Gold mining near protected areas caused much of this loss because of a surge in gold prices leading up to the global economic crisis of 2008 [3]. The analysis of satellite image data plays an integral role in monitoring industrial mines because it captures surface conditions consistently and frequently over large areas through time [4].

The detection of artisanal small-scale mines (ASMs) with satellite image data is challenging for three main reasons [5]. First, pixels are mixed at the ground sampling distance

(GSD) of moderate-resolution (~30 m) monitoring satellites. Second, ASMs are irregularly shaped and unevenly distributed. Third, the spectral response is like that of other land clearance types.

Studies involving the pixel-wise detection of ASMs with satellite image data typically consist of two main phases: (i) the transformation of spectral information into meaningful environmental indicators and (ii) classification or probabilistic mapping derived from these indicators with simple decision trees or more advanced data mining techniques. Image transformation compensates for mixed pixel effects and the irregular shape of the ASMs. The most common transformations are vegetation indices (VIs) such as the normalized difference vegetation index (NDVI) [6]. Other transformations include orthogonal rotation with principal components analysis [7] or spectral mixture analysis [3,8]. Isidro et al. [5] fused higher spatial resolution panchromatic imagery with coarser resolution multispectral broadband imagery (i.e., pan-sharpening) to improve ASM detection. Forkuor et al. [9] developed simple thresholds of ASM occurrence from summary statistics of Sentinel-1 backscatter metrics. Spatial image segmentation has also been widely used to extract ASM boundaries from spectral information [10]. More advanced data mining techniques determine ASM occurrence without or in concert with spatial image segmentation. These include support vector machines [11,12], classification and regression trees [13], canonical correlation analysis [14], and deep convolutional neural networks [15]. Distance to river networks, soil properties, and topographic features are sometimes included as covariates because of the close proximity of ASMs to water bodies [16].

The footprints of ASMs extend beyond a single pixel and consist of vegetation, pits, mounds of deposited soil, and standing pools of water, mercury, and amalgamated gold [3]. This makes them spatially heterogeneous. The pixel-wise detection of ASMs with spectral information could therefore be improved by considering spatial dependencies in neighboring pixels [17] through texture analysis. Textural metrics have been used alongside spectral metrics to segment single-date images to identify ASMs [5]. They have not been used for ASM detection in satellite image time series (SITS).

The pixel-wise segmentation of time series (i.e., time-series segmentation) is increasingly used to detect abrupt and gradual changes in vegetation [18]. This is due in large part to the availability of dense archives of free and open-access Landsat SITS and cloud-computing platforms such as Google Earth Engine (GEE) [19]. Time-series segmentation has not been applied to ASM detection, but results are promising from mapping industrial mining and mining impacts. Time-series segmentation of spectral information yielded accuracies above 80% for the detection of large open-cast (coal, mineral) mines and soil moisture declines caused by mining [20–22]. Temporally segmented spectral metrics derived from SITS better capture changes in forest cover through time when textural metrics are included as covariates [23]. The integration of temporally segmented spectral and textural features could therefore enhance ASM detection, but such an evaluation has not been previously reported.

We present a reliable and efficient ASM mapping method with the time-series segmentation of spectral and textural features derived from Landsat SITS. The method was tested in the most important gold mining region of the Brazilian portion of the Amazon Rainforest. The method involved building a Random Forest (RF) with Landsat-based detection of Trends in Disturbance and Recovery (LandTrendr) [18] segmented spectral and textural metrics, as well as proximity and topographic features. Our study utilized the Google Earth Engine (GEE) LandTrendr Application Programming Interface (API), which we modified for texture analysis. The study identified the most important spectral, textural, proximity, and topographic features and their functional relationships with the probability of ASM occurrence. We evaluated the model with ASM polygons derived from ground-based surveys and digitized manually from high spatial resolution imagery in the study area.

2. Study Area

The study area was in the Amazon Rainforest in the southeast corner of the Pará state of Brazil (Figure 1). The area spans 370 km² and falls within Landsat scene path 228, row 64. It is part of the Tapajós-Xingu moist forest ecoregion and includes the downstream contributing area of the Tapajós River, which is among the largest tributaries of the Amazon River. The Tapajós River Basin is predominantly covered by dense tropical rainforest, which serves as a major global carbon sink and critical habitat for numerous plant and animal species. The elevation varies widely, ranging from low-lying floodplains to mountainous terrain. The basin encompasses parts of the Amazon rainforest and the Brazilian Shield, with elevations ranging from sea level to over 2000 m. The climate in the Tapajós River Basin is typical tropical rainforest, which is characterized by high temperatures and humidity throughout the year. Precipitation is abundant, with a distinct wet season from December to May. The area undergoes a major flood pulse from February to May. The dry season occurs from June to September.

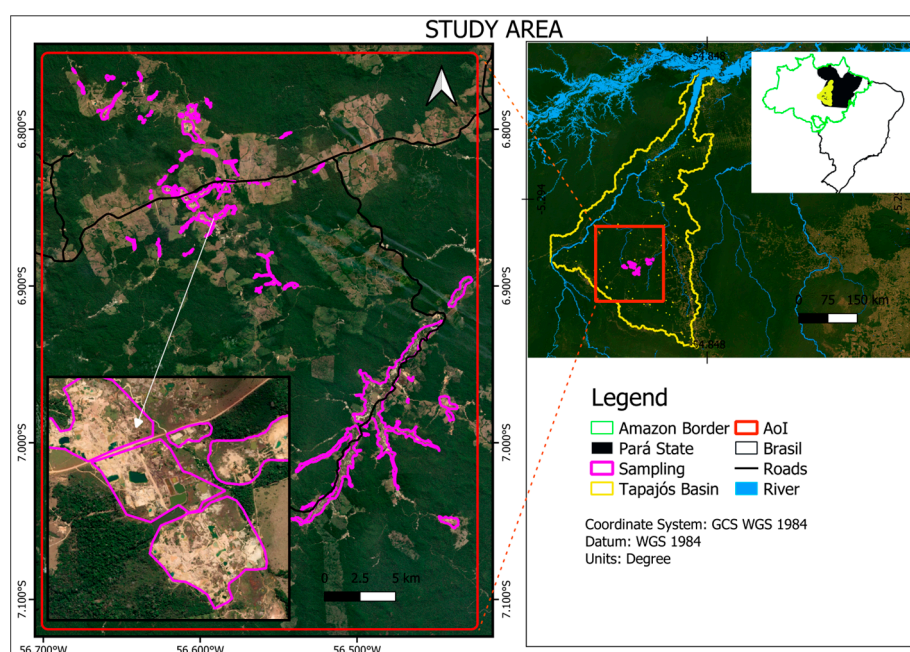


Figure 1. The Pará state and Tapajós River basin. The background of the area of interest (AOI) is a true color Landsat 8 image composite captured on 27 October 2020. The inset of the AOI b highlights ASM polygons in one of the three tiles sampled for model training and validation.

ASM activities in Pará state have significantly altered the landscape over the past 60+ years [24]. The first major gold rush began in 1958 and quickly made Pará state the principal gold producer of Brazil [25]. A new gold rush began in the early 1980s, which led the Brazilian government to create a gold mining reserve to support local miners and slow the impact of mining on the environment [26]. Gold mining not only causes forest loss. It is traditionally performed by removing topsoil near rivers or dragging out the sediments from river bottoms using suction and separating the gold by gravity [27]. Both techniques discharge sediments composed chiefly of fine organic particles and trace amounts of mercury used to amalgamate gold [28]. Mercury is also carried by air and deposited in waters further downstream. In water, mercury is methylated by microorganisms, which bioaccumulates in fish, humans, and other animals. It is an endocrine disruptor that interferes with genetic and enzyme systems and damages the nervous system and developing embryos.

CrepORIZINHO and Creporizão are the two main settlements in the study area. Both are located close to the Transgarimpeira road (06°50'14.1''S–56°35'00.0''W). The establishment of Creporizão is directly linked to the construction of the Transgarimpeira in the late 1980s. It is the local center of commerce because it connects the road to a landing on

the river. Creporizinho, on the other hand, is a typical gold-mining village. Creporizinho emerged in the early 1960s with the first artisanal extractions in the gold-mining district and increased in the 1980s when miners started to work with some machinery for processing alluvial–colluvial terraces.

3. Material and Methods

The technical workflow consisted of five main steps: (i) data acquisition from GEE and other online sources; (ii) pre-processing and transformation of spectral, textural, proximity, and topographic features; (iii) temporal segmentation with LandTrendr; (iv) feature selection with the Variable Selection Using Random Forests (VSURF) algorithm [29]; and (v) RF classification [30] of occurrence and non-occurrence (Figure 2).

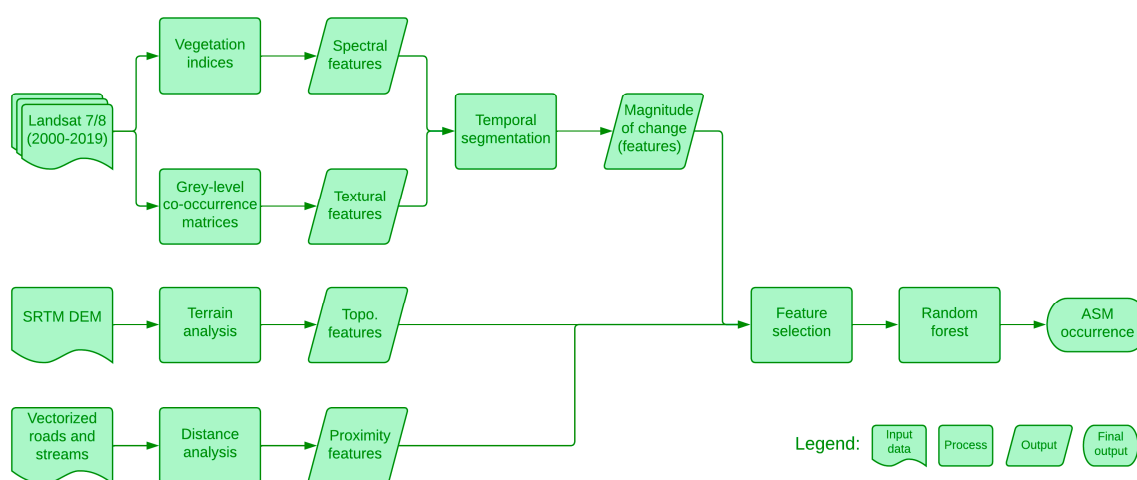


Figure 2. Technical workflow.

3.1. Geospatial Data Acquisition and Processing

We accessed Landsat 7 Enhanced Thematic Mapper (ETM+) and Landsat 8 Operational Land Imager (OLI) surface reflectance (Level-2) images, WGS-84 in GEE, from 2000 to 2019 to develop annual spectral and textural features for temporal segmentation. Landsat has a 16-day return frequency and a GSD of 30 m. Compositing was performed for the low-water period corresponding to the dry season from June to September when ASM alluvium is more easily detected remotely [16]. Six Landsat bands (Table 1) were considered for the analysis. All pre-processing steps were performed in GEE. This included Landsat 7 and 8 image harmonization; the removal of pixels with cloud, cloud shadow, and other poor-quality pixels; and compositing. Coefficients taken from Roy et al. [31] linearly transformed Landsat 7 spectral bands to minimize radiometric inconsistencies with Landsat 8. Cloud and cloud shadow were flagged and masked with the CFmask algorithm [32]. The unmasked pixels were composited on an annual basis using the medoid method [33]. The medoid method is commonly used to generate time series for LandTrendr [34]. The composites were filtered for pixels containing more than 50% cloud or cloud shadow.

Table 1. Description of Landsat 7 and 8 spectral bands.

Sensor	Designation	Spectral Range (nm)
Landsat 7 (8)	BLUE	450–520 (450–510)
	GREEN	520–600 (530–590)
	RED	630–690 (640–670)
	Near-infrared/NIR	770–900 (850–880)
	Shortwave infrared 1/SWIR 1	1550–1750 (1570–1650)
	Shortwave infrared 2/SWIR 2	2090–2350 (2110–2290)

We downloaded the National Aeronautics and Space Administration (NASA) Jet Propulsion Laboratory Shuttle Radar Topography Mission V3 digital elevation model (DEM) at a resolution of 1 arcsecond (approximately 30 m on the equator) in WGS-84 [35]. This product includes derived products from the DEM—aspect, elevation, hill shade, and slope.

The input data for the distances to roads and streams were derived from the 2019 annual Landsat image composite. The Brazilian geo web service at the Institute of Geography and Statistics Monitroamento da Cobertura e Uso da Terra do Brasil [36] roads and water bodies datasets aided the interpretation of the composite. Euclidean distance was used for proximity, which was defined as the nearest orthogonal distance from each ASM site to either roads or streams.

3.2. Spectral Transformations

We used eight spectral vegetation indices (VIs) for model building (Table 2). VIs are more robust than raw spectral bands for changing atmospheric and surface conditions [37]. Three of the indices (NBR, NDMI, NDVI) are standard functions in the LandTrendr API. We also segmented EVI, LSWI, MNDWI, NDPI, and NDWI because of their sensitivity to changes in vegetation growth and development. EVI uses coefficients and the blue band to reduce the effects of varying soil backgrounds and atmospheric constituents that adversely affect NDVI. LSWI includes a SWIR band, which is sensitive to the moisture content of canopies and soil background. The NIR band in LSWI increases as the abundance and complexity of vegetation increases. The NIR band also acts to normalize the sensitivity of the SWIR band to noise. The NDWI and MNDWI indices both enhance water features using the visible green band. Water reflects more strongly in the visible green compared to longer wavelengths. MNDWI normalizes the green band with an SWIR band. NDPI employs red, NIR, and SWIR bands. Chlorophyll in vegetation absorbs strongly in the red [38]. The red band, however, is sensitive to soil moisture, so the index includes an SWIR band to correct this effect. NDPI is scalable and able to spectrally separate water bodies from other surfaces.

Table 2. Spectral vegetation indices for temporal segmentation.

Spectral Vegetation Index	Formula	Source
Enhanced Vegetation Index	$2.5 * \frac{EVI = (NIR - RED)}{(NIR + 6 * RED - 7.5 * BLUE + 1)}$	[37]
Land Surface Water Index	$LSWI = \frac{(SWIR 1 - NIR)}{(SWIR 1 + NIR)}$	[39]
Modified Normalized Difference Water Index	$MNDWI = \frac{(GREEN - SWIR 1)}{(GREEN + SWIR 1)}$	[40]
Normalized Burn Ratio	$NBR = \frac{(NIR - SWIR 2)}{(NIR + SWIR 2)}$	[41]
Normalized Difference Moisture Index	$NDMI = \frac{(NIR - SWIR 1)}{(NIR + SWIR 1)}$	[42]
Normalized Difference Phenology Index	$NDPI = \frac{(NIR - SWIR 1)}{(NIR + SWIR 1)}$	[43]
Normalized Difference Vegetation Index	$NDVI = \frac{(NIR - RED)}{(NIR + RED)}$	[6]
Normalized Difference Water Index	$NDWI = \frac{(GREEN - NIR)}{(GREEN + NIR)}$	[44]

3.3. Textural Transformations

We used twelve grey-level co-occurrence matrix (GLCM) textural metrics [45,46] for model building (Table 3). The metrics were calculated for each Landsat band with GLCM, yielding a total of 72 textural features. GLCMs establish the frequency of occurrence of angular relationships and spatial dependence between image pixel pairs. The features are essentially statistical inferences of the frequency distributions. The features were calculated in GEE with a directional averaging kernel of 3×3 (90×90 m²). Kernel size was based on the smallest patch size of an ASM in the training dataset (~1 ha) and the GSD of the Landsat imagery. Angular second momentum is an indication of textural uniformity in the grey levels [47]. Contrast shows local differences in high and low pixel values. The higher the

contrast, the higher the difference in these values. Correlation measures the linear strength of the relationship between a given pixel with a point in the image. Variance measures the overall deviation of gray levels from the mean. High variance tends to correspond to high contrast in grey levels. Cluster prominence is a measure of asymmetry. Inertia is indicative of the contrast in intensity between a pixel and its neighbors. Cluster shade accounts for grayscale matrix skewness and uniformity [48].

Table 3. Textural metrics for temporal segmentation.

Texture Metric	Designation	Description
Angular Second Moment	Asm	Measures the number of repeated pairs
Contrast	Con	Measures the local variability of an image
Correlation	Corr	Measures the linear dependency between pixel pairs
Variance	Var	Measures the spread of the grey-level distribution
Inverse Difference Moment	IDM	Measures the homogeneity
Entropy	Ent	Measures the randomness of the gray-level distribution
Difference Variance	Dvar	Measures the variance of the gray-level distribution
Difference Entropy	DEN	Measures the difference in the randomness of the gray-level distribution
Cluster Prominence	Prom	Measures clusters by the gray-level occurrence
Dissimilarity	Diss	Measures the variation between pairs of pixels
Inertia	Iner	Measures the intensity between a pixel and its neighborhood
Shade	Shade	Measures the cluster shade of gray-level distribution

3.4. Temporal Segmentation with LandTrendr

We used the LandTrendr version and procedure described by Kennedy et al. [49] to segment the spectral and textural annual Landsat time series on a pixel basis. LandTrendr detects abrupt or gradual pixel-wise changes according to a set of breakpoints and line segments defined by the time series. The curve shown in Figure 3 represents the trajectory and LandTrendr segmentation of a 20-year NDVI time series representing a pixel that likely experienced ASM encroachment in 2008. The figure illustrates LandTrendr's ability to detect interannual noise as well as the timing and magnitude of vegetation disturbance and recovery. LandTrendr employs piecewise linear regression to determine the best-fitting trajectory along breakpoints or vertices in the time series. Like other time series analytical techniques, model tuning is required, which can lead to under- or over-fitting. Eight control parameters are required for LandTrendr to define the breakpoints and for line segment fitting. These include the maximum number of segments, spike threshold, vertex count overshoot, prevent one-year recovery (true/false), recovery threshold, p -value threshold, best model proportion, and minimum observations needed. LandTrendr uses root mean square error (RMSE) to identify the "best" model fit. Multiple combinations and different values for the eight control parameters were tested, and the set of eight parameters with the lowest RMSE was chosen for the time series analysis.

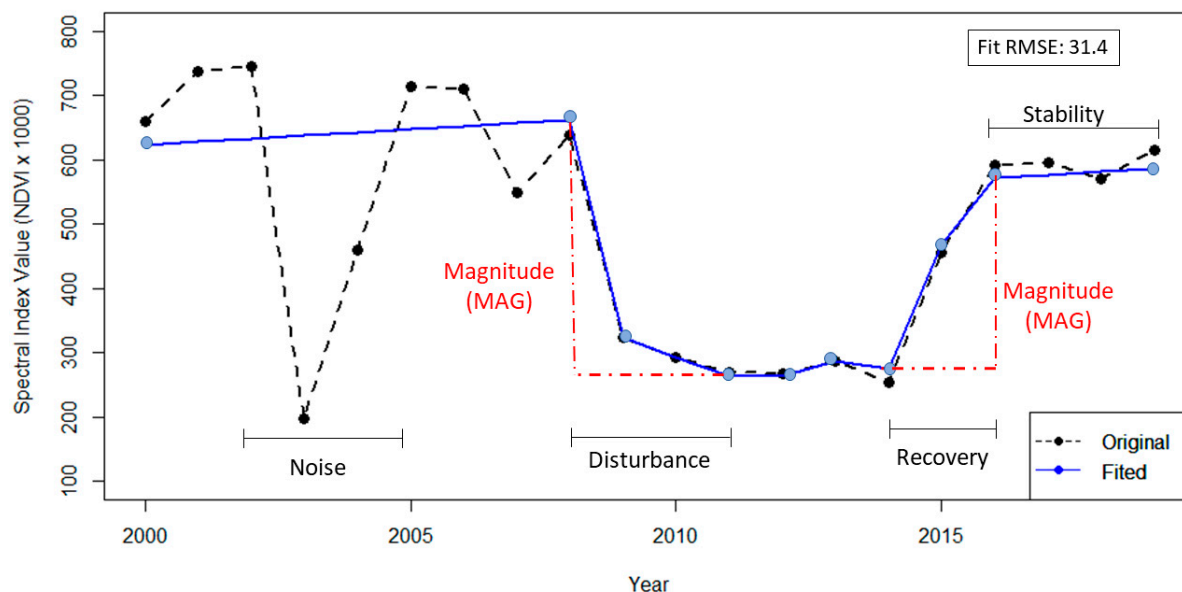


Figure 3. Temporal segmentation of a 20-year NDVI time series from an ASM disturbed pixel ($6^{\circ}50'35.2''S$ $56^{\circ}36'40.7''W$) using LandTrendr. The dotted dashed red lines represent the delta by loss drop in magnitude and disturbance duration.

Only the magnitude (MAG) of spectral and textural features was considered for Random Forest. MAG in LandTrendr is calculated as the greatest delta by loss drop in magnitude and disturbance duration, representing the largest rate of change over the time series. Ultimately, the final output included 80 MAG features (12 textures per 6 raw spectral bands plus 8 vegetation indices).

3.5. Ancillary Geospatial Information

Six ancillary data predictors were considered for model building to boost the predictive power of the spectral and textural features. These included the Euclidean distance to roads and streams, aspect, elevation, slope, and the topographic wetness index (TWI) [50]. TWI is indicative of the upstream contributing area. A high (low) TWI corresponds to stream beds (hillslopes) where water accumulation is high (low).

3.6. Feature Selection with VSURF and Random Forest Classification

RF is an ensemble machine learning technique widely used in satellite image classification [30]. RF takes the ensemble mean of a number of “weak” bootstrapped independent decision tree classifications, which tends to yield higher accuracies for land cover change detection than other classification methods [51]. RF is prone to over-fitting, however, when the number of features is high [52]. We therefore removed redundant and irrelevant features with the VSURF package in the R software version 4 environment [29] before model building. VSURF is essentially an “ensemble of ensembles” technique. It consists of three main steps: (i) thresholding; (ii) interpretation; and (iii) prediction. The first stage uses 50 iterations of RF to determine a cut-off threshold for features based on the standard deviation of feature importance. Highly irrelevant features are eliminated because they have small standard deviations. In the second stage, 25 iterations of RF are performed to identify important explanatory variables that yield a parsimonious model with low error. Some redundancy in features can remain after the interpretation stage, so the prediction stage removes any remaining redundancies in a stepwise fashion with 25 iterations of RF. Each RF consisted of 500 trees and the number of variables per split was set to the square root of the input bands as recommended by [53]. Partial dependence plots (PDPs: [54]) were extracted from the Random Forest with the PDP package in the R software version

4 environment [55]. PDPs express the relationship between ASM occurrence and a target feature, irrespective of the other features.

3.7. Random Forest Model Training and Testing

We acquired the ground reference data for RF modeling from the Rede Amazônica de Informação Socioambiental (RAISG) [56]. RAISG consists of civil societies in countries of the Amazon concerned with the health and sustainability of the Amazon rainforest. It produces vectorized data on the geolocation of mining in the Amazon basin. The dataset is a compilation of more than two decades of multidisciplinary fieldwork. We subset the dataset for mine type (gold and tin) from 2000 to 2019 to coincide with the SITS. The subset consisted of 826 “point” samples and 182 polygons. We utilized very high-resolution imagery from Google Earth to identify and digitize mine polygons based on the point data. The combination of hand-digitized polygons and RAISG polygons resulted in more than 1000 polygons in total for the analysis. We employed a pixel-stratified sampling framework to ensure spatial representativeness and prevent autocorrelation. The spatial disjoining involved dividing the entire sample set into three main tiles according to their cardinal direction in the area of interest (east, west, south). Two of the tiles were used for training and one was used for validation. Considering the nature of the problem and limitations with ground truth data, we addressed class imbalance and overfitting by employing a dual strategy [57]. All the occurrence pixels from mine polygons were used for training, which was balanced by an under-sampling of the non-occurrence class.

We used the out-of-bag (OOB) error and confusion matrix to assess model performance. RF includes an internal validation. At each iteration, RF uses a portion of the data for calibration. The unused portion is the out-of-bag sample. The OOB is the prediction error of approximately 1/3 of the samples not used in each iteration to build a Random Forest. The OOB is averaged over all iterations to give an unbiased estimate of model performance [58]. The confusion matrix expresses the frequency of correctly and misclassified samples. The diagonal of the matrix indicates the number of correctly calculated samples. Overall accuracy is calculated by summing the diagonal and dividing it by the total number of samples classified. The off-diagonals report the error of omission (Type 1 error) and error of commission (Type II error).

We converted ASM occurrence to the probability of occurrence to aid decision-makers who may need to focus resources in “hot spots” (i.e., areas of high probability of ASM occurrence). In RF, the probability of occurrence is calculated with logistic regression, probability machines, or vote counting [59]. We selected vote counting. The probability is interpreted as the proportion of votes in favor of the majority class (i.e., non-occurrence), which tends to reduce noise and increases recognition accuracy for binary classification [60].

4. Results

4.1. Temporal Segmentation with LandTrendr

LandTrendr produced maps that integrated the full temporal range of ASM disturbances over the study period (2001–2019). This is illustrated for NDVI MAG and occurrence year (YoD) in Figure 4. The occurrence year represented the year in which the temporal segment showed the first significant drop in NDVI. For example, we see near the urban areas of Creporizao and Creporizinho the expansion of the ASMs, evident by large declines in NDVI. Areas representing agricultural expansion and other forms of land clearance (e.g., logging, topsoil removal, bushfires) experienced much smaller and/or earlier (pre-2000) declines in NDVI. Other forested areas remained largely intact as they showed no change from 2000 to 2019. The occurrence of ASM expansion varied but mainly followed the global economic crisis of 2008, as can be seen in Figure 5. The figure depicts the temporal segmentation according to a red–green–blue (RGB) color composite to summarize spatial patterns. In this representation, red indicates changes occurring in the year 2000, green represents changes in 2010, and blue signifies changes in 2019. Consequently, pixels appearing yellow denote changes within the initial decade of the timeframe and purple denotes the last ten

years. The inset at the bottom of the figure depicts the temporal segmentation of an ASM pixel. It was generated from NDVI in LandTrendr to improve the interpretation of changes when they occurred according to the RGB composite. Thirty-five percent of the ASM pixels occurred before 2008. From 2008 to 2010, detected ASM pixels dropped to 25%. This figure increased to 40% after 2010.

We evaluated the performance of 86 proximity, spectral, textural, and topographic features with VSURF. Thresholding (phase 1) and interpretation (phase 2) eliminated more than half of the predictors (49 out of 86). The model at the end of these phases had an average OBB error and overall accuracy of 3.76% and 92.6%. The prediction phase reduced the number of features to 33. The overall accuracy increased by 3.33% and the OBB error decreased by 0.03%, even though an additional 16 features were removed from the model.

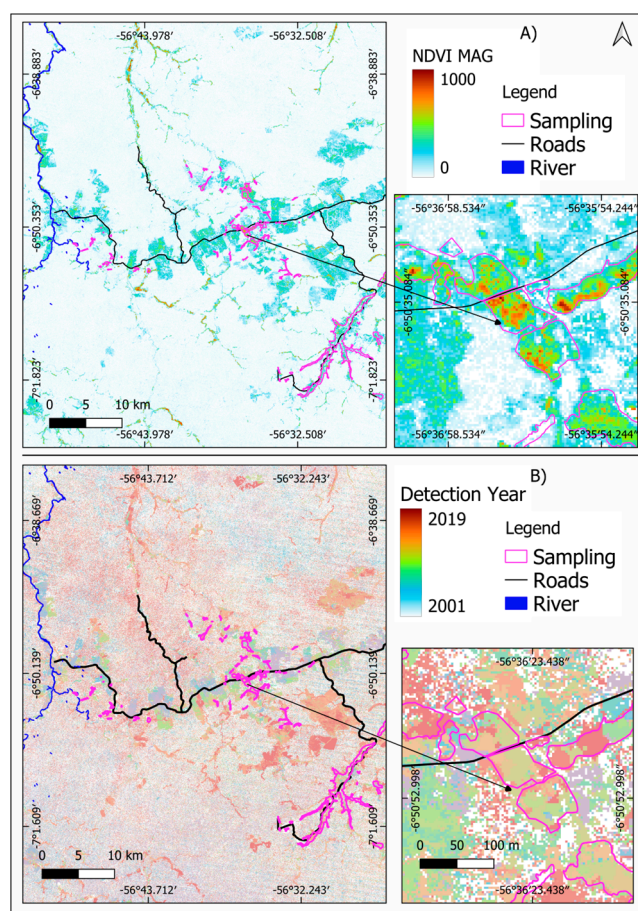


Figure 4. NDVI MAG (A) and YoD (B) calculated with LandTrendr. NDVI was expressed as an integer ($\times 1000$) in GEE LandTrendr to reduce computational demand. The insets illustrate ASM activity around the towns of Creporizão and Creporizinho.

Figure 6 shows the final subset of features we used for model building. We display only the 18 most important features for visualization purposes. NDVI ranked first. Removing it from the model led to an increase in OBB MSE of 0.086. Textural features were the next most important. If they were removed from the model, textural features led to an increase in OBB MSE of 0.066. The features included textural properties of the NIR and SWIR bands (B7 Var, B5 Prom, B5 Shade, B6 Shade, B6 contrast).

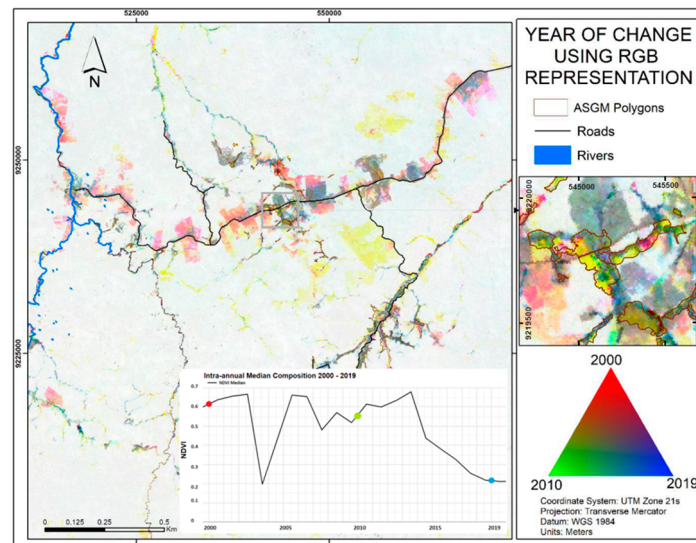


Figure 5. NDVI year of occurrence using a red–green–blue (RGB) 2000–2010–2019 composite. The inset shows the trajectory of a pixel experiencing change due to ASM activity with respect to the RGB coloring scheme. The disturbance occurred between 2010 and 2019, so the pixel appears purple in the map.

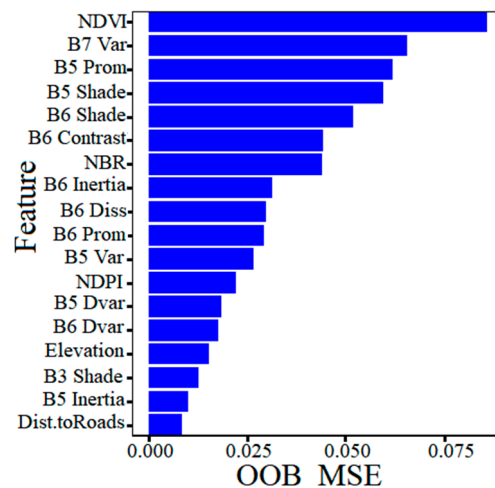


Figure 6. Features in descending order of importance versus the change in MSE that occurred if the feature was eliminated from the model.

4.2. Random Forest Classification

The overall accuracy of the Random Forest for the hold-out sample set was 85.2% for the binary classification (Table 4). Type I errors were higher in occurrence (6.4%) than non-occurrence (2.0%). Type II errors, on the other hand, were higher for non-occurrence (5.2%) than occurrence (2.4%). The distribution of the categories was approximately balanced as occurrence and non-occurrence represented 40.6% and 59.3% of the hold-out sample set, respectively. Type I errors for occurrence were relatively high because ASMs exhibited similar spectral and textural characteristics of other land clearance types. The binary classification was able to distinguish ASMs mainly along streams and roads that connect informal settlements. ASMs cluster along streams because it is relatively easy to develop mining ponds for mineral treatment and extraction. ASMs are clustered along roads to facilitate storage and transportation in a strategy locally known as “garimperios.”

Table 4. Confusion matrix (classified versus reference samples) reported for the hold-out sample set.

Classification	Reference			Total
	Class	Occurrence	Non-occurrence	
Occurrence		76%	24%	6402
Non-occurrence		5%	95%	6367
Total		5194	7575	12,769

4.3. Probabilistic Classification

The probability of occurrence provides a more nuanced interpretation of the binary classification (Figure 7). In the figure insets, it appears only the core areas of the ASM clusters especially along stream networks are likely mines (probability $\geq 70\%$). Probabilities tend to drop off sharply away from the core areas. Forest disturbances and other land clearance types unrelated to ASM activity correspond to probabilities below 50%. Some of these areas were misclassified as ASMs in the binary classification.

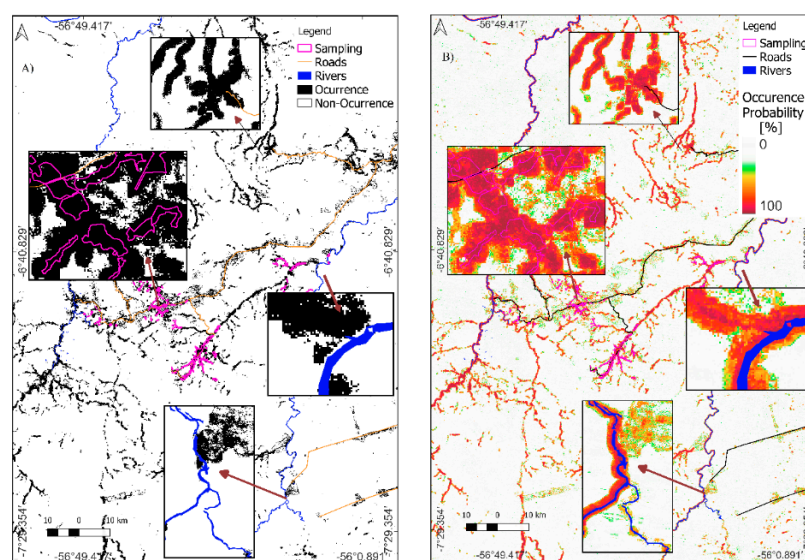


Figure 7. ASM binary (A) and probabilistic (B) classification. The three insets illustrate areas with a high frequency of ASM occurrence along streams.

4.4. Partial Dependence of Important Features

The relationships between the six most important features and the probability of ASM occurrence are shown in Figure 8. We only show six because the PDPs became less meaningful as the contribution of a feature to the predictive power of the RF declined. The probability of occurrence steadily rose as NDVI increased from 0.10 to 0.45. Values below 0.1 represent bare soils, burned areas, or some other type of non-vegetative surface. The probability gradually declined as pixels became more vegetated and NDVI increased beyond 0.5. Similarly, the likelihood of an ASM increased as differences in pixel grey levels in the SWIR increased. At extreme differences in grey levels (>500), however, ASM likelihood declined. ASMs tended to occur where NIR and SWIR shade were low (100–200 MAG). Shade is a skewness indicator; lower values represent a lack of spatial symmetry, while higher values represent spatial uniformity. On the other hand, high cluster prominence (B5 Prom) indicates high variation (contrast) in grayscale levels according to the peak of the mean in the GCLM matrix. Local (i.e., within the processing window) variations in the SWIR increased the likelihood of an ASM until around 100. Beyond 400, the probability of an ASM declined.

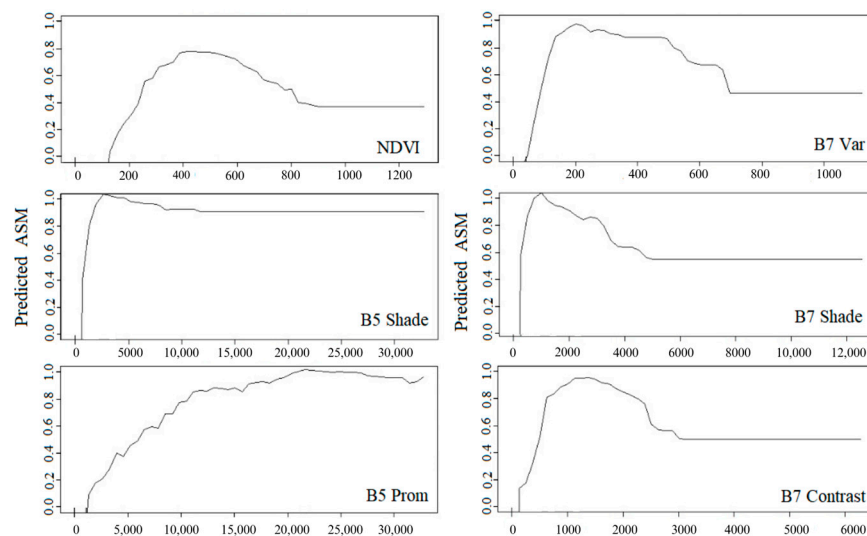


Figure 8. Partial dependence plots. Features are expressed as an integer ($\times 1000$).

5. Discussion

The results showed that the ASMs can be mapped with high accuracy at 30 m resolution with temporally segmented Landsat spectral and textural features using free and open source GEE. The analysis made three important observations. First, temporal segmentation confirmed large and widespread increases in ASM-driven natural vegetation disturbance in the study area following the global financial crisis of 2008. Second, NDVI, the traditional LandTrendr input, was the most important feature in ASM detection. However, textural features in the NIR and SWIR boosted model performance. Other spectral, topographic, and proximity features, which are commonly used to map various land cover types, were not as important. Lastly, feature selection with VSURF and balanced sampling increased the robustness of the RF model.

The performance of our RF model classification was quite high considering the small footprint of ASMs and the mixed pixel effect. Studies in inner Mongolia, China [21] and KwaZulu-Natal, South Africa [20], employed LandTrendr and Landsat SITS to interrogate disturbance caused by large open cast (coal, mineral) industrial mines. Both studies used NDVI as a model input only for RF. The models were assessed with high spatial resolution image interpretation in Google Earth. The former yielded an overall accuracy of 86.5%, while the latter yielded an overall accuracy of 99.0%. NDVI is effective at separating, spectrally, vegetative from non-vegetative surfaces. It is therefore widely used to monitor land degradation [61] caused by droughts [62], soil erosion [63], deforestation [64], and prescribed burning [65]. In this study, the PDPs revealed a proportional response of NDVI to ASM occurrence until 0.45. NDVI beyond this threshold was indicative of dense vegetation such as forest. ASMs, unlike other forms of land degradation, however, do not have a consistent spatial distribution or well-defined shape. Textural information, which characterizes the spatial structure of a landscape, can enhance the predictive power of spectral information. Meng et al. [23] used textural metrics derived from NBR to map the restoration of forests in China from orchards and other land cover types. NBR, like NDVI, is commonly used to monitor forest disturbance and recovery. The accuracy of the model improved from 20% to 63% when the temporally segmented GLCM-based texture metrics were included. We analyzed the raw spectral bands only. NBR is a ratio-based index with NIR and SWIR as input. NIR and SWIR bands were important texture metrics in our study. NIR is sensitive to changes in canopy structure, while SWIR is sensitive to changes in soil/canopy moisture. The high variance and cluster prominence together with low shade corresponded to increased ASM occurrence. This is logical since increases in these metrics are indicative of high spatial heterogeneity in terms of vegetation and moisture. Other spectral, topographic, and proximity features were likely not as critical to

ASM detection because they exhibited strong collinearity with NDVI and textural NIR and SWIR metrics. These variables were either removed during rigorous feature selection or only accounted for a small portion of the feature space after more important variables were added to the model.

Cluster Prominence, Cluster Shade, Variance, and Contrast were the most important textural features for ASM detection. They have large separations and small standard deviations compared to other metrics [48]. Cluster shade and prominence are indicative of local spectral asymmetry/heterogeneity. As both increase, the likelihood that pixels represent multiple land uses over small areas increases [66]. The second most important feature was variance in the SWIR-2 band (B7 Var). Var is a measure of the deviation of the values around the mean, indicating how spread out the sum of the gray level of a pixel pair is. Var accounts for the local gray-level variations (smoothness) of an image. Smoother images yield lower Var. Variance measures the variability by the sum of the squares of the differences between the intensity of the central pixel and its neighbors. Areas with high variance are prone to be high in contrast and thus are direct indicators of irregular shapes and image heterogeneity. Cluster Prominence texture (B5 Prom) was the third most important feature. This texture is an asymmetry indicator. SWIR-1 shade (B5 Shade) and SWIR-2 shade (B6 Shade) were the fourth and fifth most important features. Contrast (ranked sixth most important) identified local variations while calculating the differences between a pixel and neighboring pixels over an entire image.

The spatial distribution of the ASMs in the study area was dynamic and multifaceted, influenced by various factors that reflected the interplay of environmental, socioeconomic, and regulatory factors shaping where and when mining activities occur. ASMs tend to be concentrated in areas where geological formations indicate the presence of gold-bearing deposits. In our study, miners clearly targeted riverbeds, floodplains, and areas with visible signs of mineralization, such as quartz veins or altered rocks. The accessibility of mining sites plays a crucial role since miners may prefer areas with relatively easy access, such as along rivers or existing road networks. To a much lesser extent, ASM operations were in remote and difficult-to-reach areas. Miners face logistical challenges in these areas but also benefit from reduced competition and regulatory oversight. Many current mining sites are located near historical gold rushes, reflecting the continuity of mining traditions and the persistence of gold-bearing deposits.

Commission errors in the classification could be reduced by following two approaches: (i) applying strict data sampling criteria so that only the core ($\geq 70\%$ probability) mining pixels are used for training the RF classifier and (ii) eliminating uncertain samples based on local probabilities of class membership (i.e., proximity). A similar approach using TM/ETM+ images for mapping burned areas is described in Bastarrika et al. [67].

The integration of other sensors and analytical approaches into the technical workflow could improve the accuracy of ASM detection but would incur a loss of efficiency. We used Landsat 5 TM and 8 OLI whenever possible, but some Landsat 7 ETM+ image data were used to fill gaps in the time series. The SLC-off (i.e., null data) issue with Landsat 7 ETM+ was still noticeable in the final LandTrendr outputs even though LandTrendr applies an interpolation routine to correct for SLC-off. Sentinel-2 through the Harmonized Landsat and Sentinel-2 project [68] or new experimental (ENMAP, PRISMA) and planned (ESA CHIME, NASA SBG) hyperspectral missions could be added to overcome this and other spatial, spectral, and temporal limitations of Landsat. The Sentinel-2 constellation (a+b) has been acquiring images since 2015 with a higher spatial resolution (20 m) and return frequency (5 days) than Landsat [69]. It includes four additional narrow (≤ 20 nm) bands in the red edge and NIR. These bands are particularly sensitive to changes in vegetation growth and development. SWIR was an important region for ASM detection. Hyperspectral missions retrieve spectral information at ≤ 10 nm intervals in the SWIR, which is much higher than Landsat or Sentinel-2. The tropics experience persistent cloud cover, which limits the application of optical sensors such as these, particularly during the growing season. Polarimetric indices derived from cloud-penetrating Sentinel-1 or other synthetic aperture

radar overcome this challenge. This could further improve ASM detection by providing additional spectral information during the growing season when spectral differences between vegetation and other land cover types are most striking [70]. The analytical approach could be re-oriented toward semantic segmentation with convolutional neural networks (i.e., deep learning), given the importance of both spectral and textural features as well as the abundance of reference data. These approaches were recently demonstrated for ASM detection in Ghana [15]. The temporal dimension was not addressed in the Ghanaian study, possibly due to the computational demand of deep learning. A deep learning approach that more efficiently integrates time series, such as long short-term memory networks, could be a powerful new direction for monitoring ASMs.

The accurate and efficient detection of ASMs via a monitoring system could help government agencies, non-governmental organizations, and other decision-making bodies to intervene to achieve the following positive outcomes:

- **Environmental Protection:** ASMs often involve harmful practices such as deforestation, mercury pollution, and habitat destruction.
- **Regulatory Enforcement:** ASMs typically do not comply with environmental and land use laws, which leads to the illegal exploitation of natural resources.
- **Community Health and Safety:** ASMs have adverse effects on nearby communities, including health risks from exposure to toxic substances such as mercury as well as safety hazards from unstable mining structures.
- **Human Rights:** ASMs are often associated with human rights abuses such as forced labor, child labor, and exploitation.
- **Conflict Prevention:** In some regions, ASMs fuel conflicts and contribute to instability.

6. Conclusions

Mapping ASMs is challenging because they are small, spatially heterogeneous, and difficult to separate spectrally from other land cover/use types. We demonstrate a reliable and efficient method for ASM binary and probabilistic classification over a 20-year period (2000–2019) in the Pará state of Brazil—an important gold mining region of the Amazon River Basin. Previous studies mostly provide image snapshots of large industrial mines using the spatial segmentation of spectral and textural features with or without machine learning. More recent studies used time-series segmentation of NDVI derived from Landsat SITS to detect large industrial mines. Our study integrated the time-series segmentation of several Landsat spectral and textural features into an RF to detect ASMs. More than 1000 polygons were vectorized from a series of RAISG field surveys collected in the study area from 2000 to 2019. It employed rigorous feature selection with VSURF—an ensemble of ensembles technique—to eliminate all but 33 (38%) of the most important features. Feature selection increased the overall accuracy by 3.33% and decreased the OBB error by 0.03%. The final model produced an overall accuracy of 92.6%, which was 6.1% higher than a study mapping large industrial mines with the time-series segmentation of NDVI. Like other mining studies, NDVI was the most important feature in our study. However, textural features (variance, prominence, shade, contrast) in the NIR and SWIR regions of the electromagnetic spectrum enhanced model performance. The PDPs of the RF showed that as vegetation abundance (NDVI) decreased and the spatial heterogeneity in NIR and SWIR bands within ASMs and between ASMs and other land use types increased, the likelihood of a ASM activity increased.

Author Contributions: Conceptualization, A.F., M.T.M. and S.S.; Methodology, A.F., M.T.M. and S.S.; Software, A.F.; Validation, A.F.; Formal analysis, A.F.; Investigation, A.F.; Resources, A.F.; Writing—original draft, A.F., M.T.M. and S.S.; Writing—review & editing, A.F., M.T.M. and S.S.; Visualization, A.F.; Supervision, M.T.M. and S.S. 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 the study are included in the article, further inquiries can be directed to the corresponding author.

Acknowledgments: This publication was largely based on the MSc thesis of the lead author: Detecting Artisanal Small-Scale Gold mines with LandTrendr multispectral and textural features at the Tapajós river basin, Brazil. We would like to thank Roshanak Darvish and Harald van der Werff at ITC—University of Twente for their contributions to the MSc thesis.

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

References

- Sloan, S.; Sayer, J.A. Forest Resources Assessment of 2015 shows positive global trends but forest loss and degradation persist in poor tropical countries. *For. Ecol. Manag.* **2015**, *352*, 134–145. [[CrossRef](#)]
- Keenan, R.J.; Reams, G.A.; Achard, F.; de Freitas, J.V.; Grainger, A.; Lindquist, E. Dynamics of global forest area: Results from the FAO Global Forest Resources Assessment 2015. *For. Ecol. Manag.* **2015**, *352*, 9–20. [[CrossRef](#)]
- Asner, G.P.; Llactayo, W.; Tupayachi, R.; Luna, E.R. Elevated rates of gold mining in the Amazon revealed through high-resolution monitoring. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 18454–18459. [[CrossRef](#)] [[PubMed](#)]
- Werner, T.T.; Mudd, G.M.; Schipper, A.M.; Huijbregts, M.A.J.; Taneja, L.; Northey, S.A. Global-scale remote sensing of mine areas and analysis of factors explaining their extent. *Glob. Environ. Chang.* **2020**, *60*, 102007. [[CrossRef](#)]
- 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*, 945. [[CrossRef](#)]
- Rouse, J.W.; Haas, R.H.; Schell, J.A.; Deering, D.W. *Monitoring Vegetation Systems in the Great Plains with ERTS*; Scientific and Technical Information Office, National Aeronautics and Space Administration: Washington, DC, USA, 1974; Volume 351, p. 309.
- Almeida-Filho, R.; Shimabukuro, Y.E. Digital processing of a Landsat-TM time series for mapping and monitoring degraded areas caused by independent gold miners, Roraima State, Brazilian Amazon. *Remote Sens. Environ.* **2002**, *79*, 42–50. [[CrossRef](#)]
- Caballero Espejo, J.; Messinger, M.; Román-Dañobeytia, F.; Ascorra, C.; Fernandez, L.E.; Silman, M. Deforestation and Forest Degradation Due to Gold Mining in the Peruvian Amazon: A 34-Year Perspective. *Remote Sens.* **2018**, *10*, 1903. [[CrossRef](#)]
- Forkuor, G.; Ullmann, T.; Griesbeck, M. Mapping and Monitoring Small-Scale Mining Activities in Ghana using Sentinel-1 Time Series (2015–2019). *Remote Sens.* **2020**, *12*, 911. [[CrossRef](#)]
- Simionato, J.; Bertani, G.; Osako, L.S. Identification of artisanal mining sites in the Amazon Rainforest using Geographic Object-Based Image Analysis (GEOBIA) and Data Mining techniques. *Remote Sens. Appl. Soc. Environ.* **2021**, *24*, 100633. [[CrossRef](#)]
- Ibrahim, E.; Lema, L.; Barnabe, P.; Lacroix, P.; Pirard, E. Small-scale surface mining of gold placers: Detection, mapping, and temporal analysis through the use of free satellite imagery. *Int. J. Appl. Earth Obs. Geoinf.* **2020**, *93*, 102194. [[CrossRef](#)]
- Ngom, N.M.; Mbaye, M.; Baratoux, D.; Baratoux, L.; Ahoussi, K.E.; Kouame, J.K.; Faye, G.; Sow, E.H. Recent expansion of artisanal gold mining along the Bandama River (Côte d’Ivoire). *Int. J. Appl. Earth Obs. Geoinf.* **2020**, *112*, 102873. [[CrossRef](#)]
- Lobo, F.D.L.; Souza-Filho, P.W.M.; de Moreas Novo, E.M.L.; Carlos, F.M.; Barbosa, C.C.F. Mapping Mining Areas in the Brazilian Amazon Using MSI/Sentinel-2 Imagery (2017). *Remote Sens.* **2018**, *10*, 1178. [[CrossRef](#)]
- Snafir, B.; Simms, D.M.; Waive, T.W. Mapping the expansion of galamsey gold mines in the cocoa growing area of Ghana using optical remote sensing. *Int. J. Appl. Earth Obs. Geoinf.* **2017**, *58*, 225–233. [[CrossRef](#)]
- Gallwey, J.; Robiati, C.; Coggan, J.; Vogt, D.; Eyre, M. A Sentinel-2 based multispectral convolutional neural network for detecting artisanal small-scale mining in Ghana: Applying deep learning to shallow mining. *Remote Sens. Environ.* **2020**, *248*, 111970. [[CrossRef](#)]
- Lobo, F.; Costa, M.; Novo, E.; Telmer, K. Distribution of Artisanal and Small-Scale Gold Mining in the Tapajós River Basin (Brazilian Amazon) over the Past 40 Years and Relationship with Water Siltation. *Remote Sens.* **2016**, *8*, 579. [[CrossRef](#)]
- Li, M.; Zang, S.; Zhang, B.; Li, S.; Wu, C. A Review of Remote Sensing Image Classification Techniques: The Role of Spatio-contextual Information. *Eur. J. Remote Sens.* **2014**, *47*, 2279–7254. [[CrossRef](#)]
- Kennedy, R.E.; Yang, Z.; Cohen, W.B. Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr—Temporal segmentation algorithms. *Remote Sens. Environ.* **2010**, *114*, 2897–2910. [[CrossRef](#)]
- Vogelmann, J.E.; Gallant, A.L.; Shi, H.; Zhu, Z. Perspectives on monitoring gradual change across the continuity of Landsat sensors using time-series data. *Remote Sens. Environ.* **2016**, *185*, 258–270. [[CrossRef](#)]
- Dlamini, L.Z.D.; Xulu, S. Monitoring Mining Disturbance and Restoration over RBM Site in South Africa Using LandTrendr Algorithm and Landsat Data. *Sustainability* **2019**, *11*, 6916. [[CrossRef](#)]
- Xiao, W.; Deng, X.; He, T.; Chen, W. Mapping Annual Land Disturbance and Reclamation in a Surface Coal Mining Region Using Google Earth Engine and the LandTrendr Algorithm: A Case Study of the Shengli Coalfield in Inner Mongolia, China. *Remote Sens.* **2020**, *12*, 1612. [[CrossRef](#)]
- Yi, Z.; Liu, M.; Liu, X.; Wang, Y.; Wu, L.; Wang, Z.; Zhu, L. Long-term Landsat monitoring of mining subsidence based on spatiotemporal variations in soil moisture: A case study of Shanxi Province, China. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *102*, 102447. [[CrossRef](#)]
- Meng, Y.; Liu, X.; Wang, Z.; Ding, C.; Zhu, L. How can spatial structural metrics improve the accuracy of forest disturbance and recovery detection using dense Landsat time series? *Ecol. Indic.* **2021**, *132*, 108336. [[CrossRef](#)]

24. Berzas Nevado, J.J.; Rodríguez Martín-Doimeadios, R.C.; Guzmán Bernardo, F.J.; Jiménez Moreno, M.; Herculano, A.M.; do Nascimento, J.L.M.; Crespo-López, M.E. Mercury in the Tapajós River basin, Brazilian Amazon: A review. *Environ. Int.* **2010**, *36*, 593–608. [[CrossRef](#)] [[PubMed](#)]
25. Villas Bôas, R.C.; Beinhoff, C.; da Silva, A.R.B. *Mercury in the Tapajos Basin*; CYTED: Rio de Jeniro, Brazil, 2001; 198p.
26. Roulet, M.; Lucotte, M.; Canuel, R.; Farella, N.; De Freitas Goch, Y.G.; Pacheco Peleja, J.R.; Guimaraes, J.-R.D.; Mergler, D.; Amorim, M. Spatio-temporal geochemistry of mercury in waters of the Tapajós and Amazon rivers, Brazil. *Limnol. Oceanogr.* **2001**, *46*, 1141–1157. [[CrossRef](#)]
27. Lobo, F.L.; Costa, M.P.F.; Novo, E.M.L.M. Time-series analysis of Landsat-MSS/TM/OLI images over Amazonian waters impacted by gold mining activities. *Remote Sens. Environ.* **2015**, *157*, 170–184. [[CrossRef](#)]
28. Castello, L.; Macedo, M.N. Large-scale degradation of Amazonian freshwater ecosystems. *Glob. Chang. Biol.* **2016**, *22*, 990–1007. [[CrossRef](#)] [[PubMed](#)]
29. Genuer, R.; Poggi, J.-M.; Tuleau-Malot, C. VSURF: An R Package for Variable Selection Using Random Forests. *R J.* **2015**, *7*, 19–33. [[CrossRef](#)]
30. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
31. Roy, D.P.; Kovalskyy, V.; Zhang, H.K.; Vermote, E.F.; Yan, L.; Kumar, S.S.; Egorov, A. Characterization of Landsat-7 to Landsat-8 reflective wavelength and normalized difference vegetation index continuity. *Remote Sens. Environ.* **2016**, *185*, 57–70. [[CrossRef](#)]
32. Zhu, Z.; Wang, S.; Woodcock, C.E. Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4-7, 8, and Sentinel 2 images. *Remote Sens. Environ.* **2015**, *159*, 269–277. [[CrossRef](#)]
33. Flood, N. Seasonal Composite Landsat TM/ETM+ Images Using the Medoid (a Multi-Dimensional Median). *Remote Sens.* **2013**, *5*, 6481–6500. [[CrossRef](#)]
34. Cohen, W.B.; Yang, Z.; Healey, S.P.; Kennedy, R.E.; Gorelick, N. A LandTrendr multispectral ensemble for forest disturbance detection. *Remote Sens. Environ.* **2018**, *205*, 131–140. [[CrossRef](#)]
35. Farr, T.G.; Rosen, P.A.; Caro, E.; Crippen, R.; Duren, R.; Hensley, S.; Kobrick, M.; Paller, M.; Rodriguez, E.; Roth, L.; et al. The shuttle radar topography mission. *Rev. Geophys.* **2007**, *45*, 2004. [[CrossRef](#)]
36. Instituto Brasileiro de Geografia e Estatística (IBGE). Cartas e Mapas de Bases Cartograficas Continuas. 2020. Available online: <http://geoftp.ibge.gov.br/> (accessed on 4 May 2021).
37. Huete, A.; Didan, K.; Miura, T.; Rodriguez, E.P.; Gao, X.; Ferreira, L.G. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens. Environ.* **2002**, *83*, 195–213. [[CrossRef](#)]
38. Xu, D.; Wang, C.; Chen, J.; Shen, M.; Shen, B.; Yan, R.; Li, Z.; Karnieli, A.; Chen, J.; Yan, Y.; et al. The superiority of the normalized difference phenology index (NDPI) for estimating grassland aboveground fresh biomass. *Remote Sens. Environ.* **2021**, *264*, 112578. [[CrossRef](#)]
39. Xiao, X.; Boles, S.; Frohling, S.; Salas, W.; Moore Iii, B.; Li, C.; He, L.; Zhao, R. Observation of flooding and rice transplanting of paddy rice fields at the site to landscape scales in China using VEGETATION sensor data. *Int. J. Remote Sens.* **2002**, *23*, 3009–3022. [[CrossRef](#)]
40. Xu, H. Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *Int. J. Remote Sens.* **2006**, *27*, 3025–3033. [[CrossRef](#)]
41. Keeley, J.E. Fire intensity, fire severity and burn severity: A brief review and suggested usage. *Int. J. Wildland Fire* **2009**, *18*, 116–126. [[CrossRef](#)]
42. Kaptué, A.T.; Hanan, N.P.; Prihodko, L. Characterization of the spatial and temporal variability of surface water in the Soudan-Sahel region of Africa. *J. Geophys. Res.* **2013**, *118*, 1472–1483. [[CrossRef](#)]
43. Wang, C.; Chen, J.; Wu, J.; Tang, Y.; Shi, P.; Black, T.A.; Zhu, K. A snow-free vegetation index for improved monitoring of vegetation spring green-up date in deciduous ecosystems. *Remote Sens. Environ.* **2017**, *196*, 1–12. [[CrossRef](#)]
44. McFeeters, S.K. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *Int. J. Remote Sens.* **1996**, *17*, 1425–1432. [[CrossRef](#)]
45. Haralick, R.M.; Dinstein, I.; Shanmugam, K. Textural Features for Image Classification. *IEEE Trans. Syst. Man Cybern.* **1973**, *SMC-3*, 610–621.
46. Connors, R.W.; Trivedi, M.M.; Harlow, C.A. Segmentation of a high-resolution urban scene using texture operators (Sunnyvale, California). *Comput. Vis. Graph. Image Process.* **1984**, *25*, 273–310. [[CrossRef](#)]
47. Ramola, A.; Shakya, A.K.; Van Pham, D. Study of statistical methods for texture analysis and their modern evolutions. *Eng. Rep.* **2020**, *2*, e12149. [[CrossRef](#)]
48. Yang, X.; Tridandapani, S.; Beitler, J.J.; Yu, D.S.; Yoshida, E.J.; Curran, W.J.; Liu, T. Ultrasound GLCM texture analysis of radiation-induced parotid-gland injury in head-and-neck cancer radiotherapy: An in vivo study of late toxicity. *Med. Phys.* **2012**, *39*, 5732–5739. [[CrossRef](#)] [[PubMed](#)]
49. Kennedy, R.E.; Yang, Z.; Gorelick, N.; Braaten, J.; Cavalcante, L.; Cohen, W.B.; Healey, S. 5 LT-GEE Outputs | LT-GEE Guide [WWW Document]. Kennedy RE Yang Z Gorelick N Braaten J Cavalcante Cohen WB Heal. 2018 Implement. LandTrendr Algorithm Google Earth Engine. *Remote Sens.* **2018**, *10*, 691. [[CrossRef](#)]
50. Beven, K.J.; Kirkby, M.J. A physically based, variable contributing area model of basin hydrology/Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. *Hydrol. Sci. J.* **1979**, *24*, 43–69. [[CrossRef](#)]

51. Talukdar, S.; Singha, P.; Mahato, S.; Shahfahad Pal, S.; Liou, Y.A.; Rahman, A. Land-use land-cover classification by machine learning classifiers for satellite observations-A review. *Remote Sens.* **2020**, *12*, 1135. [CrossRef]
52. Chavent, M.; Genuer, R.; Saracco, J. Combining clustering of variables and feature selection using random forests. *Commun. Stat. Simul. Comput.* **2021**, *50*, 426–445. [CrossRef]
53. Belgiu, M.; Drăguț, L. Random forest in remote sensing: A review of applications and future directions. *ISPRS J. Photogramm. Remote Sens.* **2016**, *114*, 24–31. [CrossRef]
54. Friedman, J.H. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [CrossRef]
55. Greenwell, B.M. pdp: An R Package for Constructing Partial Dependence Plots. *R J.* **2022**, *9*, 421–436. [CrossRef]
56. RAISG. Amazon Geo-Referenced Socio-Environmental Information Network [WWW Document]. 2020. Available online: <https://mineria.amazoniasocioambiental.org/> (accessed on 6 April 2021).
57. Singh, P.S.; Singh, V.P.; Pandey, M.K.; Karthikeyan, S. Enhanced classification of hyperspectral images using improvised oversampling and undersampling techniques. *Int. J. Inf. Technol.* **2022**, *14*, 389–396. [CrossRef]
58. Matthew, W. Bias of the random forest out-of-bag (OOB) error for certain input parameters. *Open J. Stat.* **2011**, *1*, 205–211.
59. Svetnik, V.; Liaw, A.; Tong, C.; Wang, T. Application of Breiman’s random forest to modeling structure-activity relationships of pharmaceutical molecules. In Proceedings of the Multiple Classifier Systems: 5th International Workshop, MCS 2004, Cagliari, Italy, 9–11 June 2004; Springer: Berlin/Heidelberg, Germany, 2004; pp. 334–343. [CrossRef]
60. Joutsijoki, H.; Juhola, M. Kernel selection in multi-class support vector machines and its consequence to the number of ties in majority voting method. *Artif. Intell. Rev.* **2013**, *40*, 213–230. [CrossRef]
61. Jin, J.; Wang, Q. Assessing ecological vulnerability in western China based on Time-Integrated NDVI data. *J. Arid Land* **2016**, *8*, 533–545. [CrossRef]
62. Liu, S.; Wei, X.; Li, D.; Lu, D. Examining Forest Disturbance and Recovery in the Subtropical Forest Region of Zhejiang Province Using Landsat Time-Series Data. *Remote Sens.* **2017**, *9*, 479. [CrossRef]
63. Del Río-Mena, T.; Willemen, L.; Vrieling, A.; Nelson, A. Understanding Intra-Annual Dynamics of Ecosystem Services Using Satellite Image Time Series. *Remote Sens.* **2020**, *12*, 710. [CrossRef]
64. Asner, G.P. Automated mapping of tropical deforestation and forest degradation: CLASlite. *J. Appl. Remote Sens.* **2009**, *3*, 033543. [CrossRef]
65. Key, C.H.; Benson, N.C. Landscape Assessment (LA). Sampling and Analysis Methods. In USDA Forest Service—General Technical Report RMRS-GTR. 2006. Available online: https://gsp.humboldt.edu/OLM/Courses/GSP_216/labs/rmrs_gtr164_1_3_land_assess.pdf (accessed on 1 June 2021).
66. Caelli, T.; Julesz, B. Psychophysical Evidence for Global Feature Processing in Visual Texture Discrimination. *J. Opt. Soc. Am.* **1979**, *69*, 675–678. [CrossRef]
67. Bastarrika, A.; Chuvieco, E.; Martín, M.P. Mapping burned areas from Landsat TM/ETM+ data with a two-phase algorithm: Balancing omission and commission errors. *Remote Sens. Environ.* **2011**, *115*, 1003–1012. [CrossRef]
68. Claverie, M.; Ju, J.; Masek, J.G.; Dungan, J.L.; Vermote, E.F.; Roger, J.-C.; Skakun, S.V.; Justice, C. The Harmonized Landsat and Sentinel-2 surface reflectance data set. *Remote Sens. Environ.* **2018**, *219*, 145–161. [CrossRef]
69. Frampton, W.J.; Dash, J.; Watmough, G.; Milton, E.J. Evaluating the capabilities of Sentinel-2 for quantitative estimation of biophysical variables in vegetation. *ISPRS J. Photogramm.* **2013**, *82*, 83–92. [CrossRef]
70. Trivedi, M.; Marshall, M.; Estes, L.; de Bie, K.; Chang, L.; Nelson, A. Cropland Mapping in Tropical Smallholder Systems with Seasonally Stratified Sentinel-1 and Sentinel-2 Spectral and Textural Features. *Remote Sens.* **2023**, *15*, 3014. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.