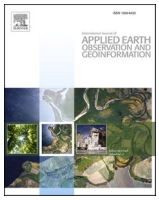
International Journal of Applied Earth Observation and Geoinformation 132 (2024) 104015

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

journal homepage: www.elsevier.com/locate/jag



Feasibility of multi-spectral and radar data fusion for mapping Artisanal Small-Scale Mining: A case study from Indonesia

Ilyas Nursamsi [a,*], Laura Jane Sonter [a,b], Matthew Scott Luskin [a,b], Stuart Phinn [a,c]

[a] School of the Environment, The University of Queensland, St Lucia, Brisbane, QLD, Australia
[b] Centre for Biodiversity and Conservation Science, The University of Queensland, Australia
[c] Remote Sensing Research Centre, The University of Queensland, St Lucia, Brisbane, QLD, Australia

ARTICLE INFO

Keywords:
Artisanal small-scale gold mining
Data fusion
Multi-spectral
SAR
Surveillance

ABSTRACT

Artisanal small-scale mining (ASM) is an environmentally damaging activity in many developing countries, particularly in the wet tropics, yet serves as a crucial economic resource for millions of people. The lack of effective mapping methods hinders quantifying the spatial extent of ASM and management efforts. This study presents a novel approach to integrate multi-spectral and imaging radar datasets within the Google Earth Engine (GEE) platform to map ASGM in a tropical rainforest. We used a case study of gold mining in central Kalimantan and diverse training and validation data sources. The methodology involved pre-processing multispectral and radar imagery, generating and standardizing covariates, applying feature-level data fusion for the Random Forest algorithm in GEE, and training and classifying data with optimized parameters through iterative loops. This approach achieved a classification accuracy of 81% in detecting ASM activities, surpassing the accuracy of a map constructed solely from Sentinel-2 multispectral data by 14%. Through the inclusion of evaluation metrics such as the f(β) score and Matthews Correlation Coefficient (MCC), our approach demonstrates its robustness in accurately identifying target instances, while reducing false positives and addressing imbalanced class sizes by 6.25% and 60%, respectively. Our model's efficacy underscores its potential to accurately map ASM at larger regional scales (104 – 10^6 km^2) in wet-tropical forests, while being scalable and resource-efficient. Opportunities to further improve this approach by mitigating false-positive errors involve integrating texture filtering with optical and radar data sets. Despite some inherent limitations, our approach overcomes some current challenges of mapping small-scale, but extensive, environmental changes in the wet tropics and thus advances improvements in the continual surveillance, management, and regulation of ASM and other activities that involve selective clearing.

1. Introduction

Artisanal Small-scale Gold Mining (ASGM) serves as a significant income source for underprivileged populations in developing countries, with an estimated global distribution of 16 million miners as of 2018 (Fritz et al. 2018). The escalating demand for gold, fuelled by its diverse applications in jewellery, investment, and technology, has spurred an increase in global gold production from approximately 2,445 MT in the year 2000 to around 3,300 MT in 2019 (Shafiee and Topal 2010). This surge has motivated low-capital miners to exploit lower-grade deposits situated beneath tropical forests (Swenson et al. 2011, Alvarez-Berríos and Aide 2015). Approximately 20 % of global gold production originates from ASGM operations in developing pantropical countries, which operate in over 1,200 identified hotspots (Bebbington et al. 2018, Foundation 2018, World-Bank-Group 2019).

Coinciding with this global trend, ASGM activity in Indonesia has grown rapidly over the past 20 years, with operations extending to 30 of its 34 provinces (Bebbington et al. 2018). Limited access to capital confines miners to horizontal extraction from alluvial sources, causing forest loss and environmental degradation (Kalamandeen et al. 2018, Macháček 2019). Deforestation from ASGM and its infrastructure (e.g., roads and temporary settlements) leads to erosion, sedimentation, and heavy metal pollution, harming humans, fish, and wildlife along river systems (Al-Hassan et al. 2019, Ofosu et al. 2020). Although the impacts are local, the increasing number of ASGM activities raises concerns about cumulative effects at regional scales, exacerbating existing

* Corresponding author.
E-mail address: i.nursamsi@uq.edu.au (I. Nursamsi).

https://doi.org/10.1016/j.jag.2024.104015
Received 20 March 2024; Received in revised form 26 June 2024; Accepted 1 July 2024
Available online 8 July 2024

environmental pressures from other human activities (Franks et al. 2010). Kalimantan (Indonesian Borneo) is a hotspot for heavy metal pollution, even within protected areas, reflecting a global pattern of ASGM overlapping with protected regions and impacting tropical South American forests (Villegas et al. 2013, Engstrand et al. 2024). Despite these threats, there is a lack of accurate and timely spatial information, hindering our understanding of ASGM's distribution and impacts.

Previous research on remote sensing (RS) techniques for ASGM mapping in wet tropical environments faces limitations with multispectral satellite imagery, including reduced usable data due to cloud cover, lack of unique spectral signatures for targets, limited detection capabilities of indices like NDVI, and inability to detect changes below closed canopies (Isidro et al. 2017). The elusive and temporary nature of ASGM, which often have a limited surface footprint (<1ha), further hampers effective RS surveillance (Hilson and McQuilken 2014, Isidro et al. 2017). Radar data are less sensitive to these specific limitations; however, they pose their challenges if not collected and corrected properly. These include speckle or signal noise issues, leading to increased measurement uncertainty and compromised classification accuracies (Janse van Rensburg and Kemp 2022, Alessi et al. 2023). Imaging radars are now accessible from multiple satellite sensors operated by governments and companies in analysis-ready formats. Integrating both multispectral and radar data can compensate for the drawbacks inherent in each, enabling multisource fusion development in RS methodologies such as multispectral-radar data fusion. Data fusion has proven effective in similar RS applications such as digital surface models (DSMs) and land-use/cover (LULC) mapping studies in other tropical environments.

The fusion of multispectral and radar datasets, however, requires large data storage and processing capabilities (Moomen et al. 2022). Fortunately, cloud-based geo-processing platforms such as Google's Earth Engine and Microsoft's Planetary Computer with global archives of fully corrected optical and radar data, facilitate scalability, data storage, and data processing demands (Mutanga and Kumar 2019). Additionally, these platforms provide an efficient, cost-effective way of accessing and analysing large quantities of satellite imagery, thereby eliminating the need for costly on-site hardware and software.

Here we outline how the synergy of multispectral-radar (feature-level) data fusion with cloud-based geo-processing platforms may enable accurate and efficient ASGM monitoring at regional scales (10^4 – 10^6 km^2). Using this approach, we aim to overcome the current challenges of mapping ASM in wet tropical forests while ensuring the approach is scalable and resource-efficient with a cloud-based platform. We provide a case study using 153,000 km^2 in Central Kalimantan Province, Indonesia, with satellite multispectral (Landsat-8 and Sentinel-2) and radar (Sentinel-1 and ALOS/PALSAR) datasets for the year 2022, processed within the GEE platform. By adopting a feature-level data fusion approach and capitalizing on the processing capabilities of GEE, we map ASGM across a large geographic extent.

2. Materials and methods

2.1. Reference data

We compiled reference sites of known ASGM from three distinct sources:

(1) manual interpretation of Google Earth imagery using known surface features of ASGM,

(2) reports from governmental and non-governmental organizations detailing confirmed ASGM sites, and.

(3) global digitized mining area datasets provided by Maus et al. (2022) and Tang and Werner (2023), which present an unsystematic manual interpretation of mining sites, including some ASGM sites.

We manually interpreted ASGM from Google Earth imagery, following Maus et al. (2022) and Tang and Werner (2023), identifying sites through visual recognition of features such as unvegetated, lunar-shaped open pits, craters, and crevasses, indicating surface soil removal, often with accompanying water-filled ponds for slushing and washing processes. We generated 546 polygon samples at verified ASGM locations, each with a minimum size of 2 ha (equivalent to 20x20 pixels at a 10 m resolution). We then divided these polygons into training and test datasets using the random selection function in GEE (ee.FeatureCollection.randomColumn()). This resulted in 100 polygons for ASGM class testing, 95 for non-ASGM class testing, 251 for ASGM class training, and 100 for non-ASGM class training.

2.2. Satellite image datasets and study region

This study integrates Landsat-8 and Sentinel-2 multispectral images with Sentinel-1 and ALOS/PALSAR radar images in the GEE environment. We selected cloud-less multispectral images captured during Indonesia's dry season (April 1, 2022, to October 31, 2022), imposing a strict threshold of less than 40 % cloud cover. We confined and clipped all data geographically to Central Kalimantan Province, Indonesia (Fig. 1). We retrieved Landsat-8 T1_SR (surface reflectance) (USGS 2016) and Sentinel-2 S2_SR (surface reflectance) products from both the 2a and 2b satellites (Drusch et al. 2012). Landsat-8, with a 30 m spatial resolution (except for bands 8, 11, and 12), and Sentinel-2, with varying resolutions (10 m for the blue, green, red, and near-infrared bands, 20 m for the red-edge and shortwave-infrared bands, and 60 m for all others). We selected bands 2-8a for Sentinel-2 and bands 2–6 Landsat 8.

Additionally, we acquired Sentinel-1's S1_GRD (ground range detected) data in Interferometric Wide (IW) swath mode, which covers VV (vertical–vertical) and VH (vertical-horizontal) polarization bands (Gorelick et al. 2017). Sentinel-1 has a spatial resolution of 10 m. We then incorporated JAXA ALOS/PALSAR ver. 2 data (for the year 2022), which covers HH (horizontal-horizontal) and HV (horizontal-vertical) polarizations (25 m resolution), to complete our dataset. Subsequently, we transformed the digital numbers (DN) values into backscattering coefficients (σ^0) in decibel units (dB) using a designated formula (Shimada et al. 2009, Shimada and Ohtaki 2010):

$$\sigma^0 = 10log_{10}(DN)^2 + CF$$

Here, σ^0 represents the backscattering coefficient, DN stands for the digital number in the original ALOS/PALSAR image, and CF denotes the calibration factor, consistently set at −83 dB for both HH and HV polarizations.

2.3. Methods overview

We initiated our analysis by applying pre-processing steps (sections 2.4.1 and 2.4.2) to all retrieved multispectral imagery (Fig. 2). We only conducted Lee speckle filtering and square cosine correction to de-speckle the image for the Sentinel-1 data (Lee et al. 2008, Van Tricht et al. 2018) since it had already been pre-processed using the Sentinel-1 toolbox from ESA, involving tasks such as updating orbit metadata, thermal noise removal, radiometric calibration, and terrain correction (Mullissa et al. 2021). We did not conduct any further pre-processing steps for the ALOS/PALSAR imagery since it has also been *ortho*-rectified and slope-corrected. (JAXA 2017). Subsequently, we generated covariates by computing multispectral- and radar-based indices using the respective bands (Section 2.4.3; Table 1). The chosen indices have a proven track record in previous studies for effectively identifying and mapping features similar to open-pit mining sites in forested environments (e.g., Almeida-Filho and Shimabukuro (2000), Forkuor et al. (2020), Gao et al. (2021), Zhang et al. (2021)). We then standardized these covariates to ensure uniformity of scale.

Pohl and Van Genderen (1998) outline three levels of data fusion: pixel-level, feature-level, and decision-level. This study utilized feature-level fusion, merging extracted features from various sensors into a multi-source feature stack. Following this, we then stacked all of the

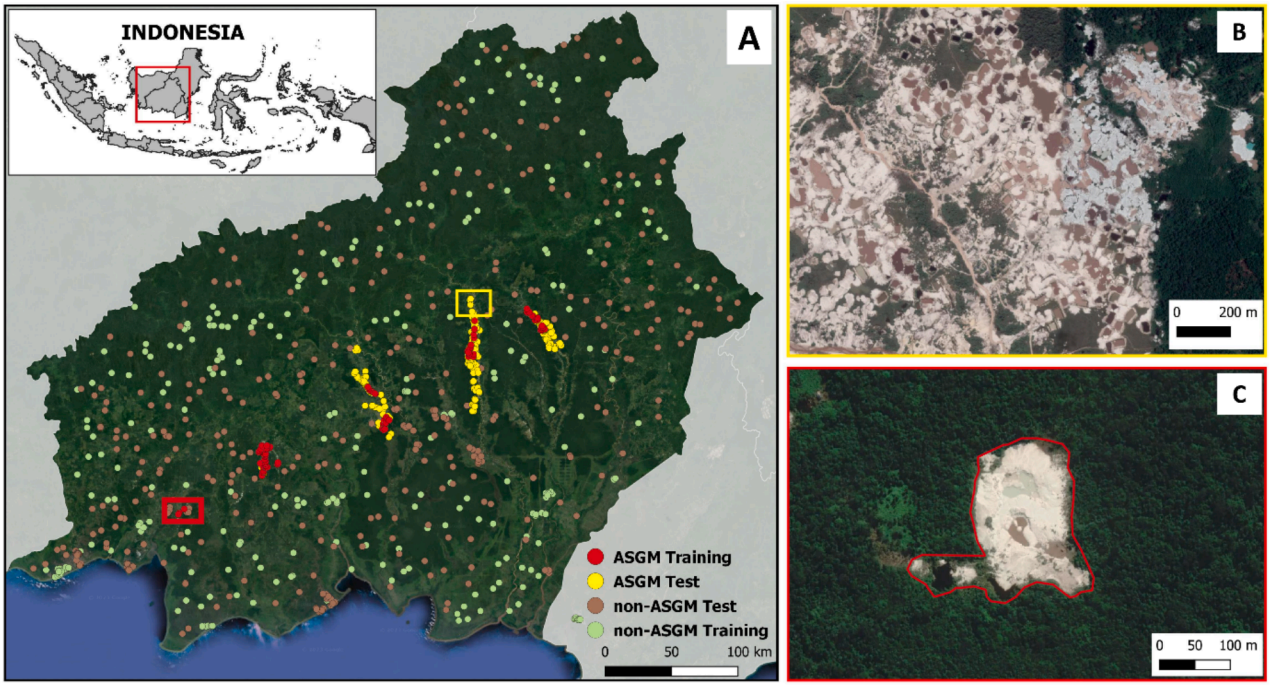


Fig. 1. Study area (Central Kalimantan province, Indonesia) and examples of ASGM surface footprints utilized for training and test data (Yellow and red box zoomed-in): A. Spatial distribution of training and test areas across the study region. B. Clusters of ASGM mining sites (Yellow box zoomed-in). C. Isolated individual ASGM mining site (red-box zoomed-in).

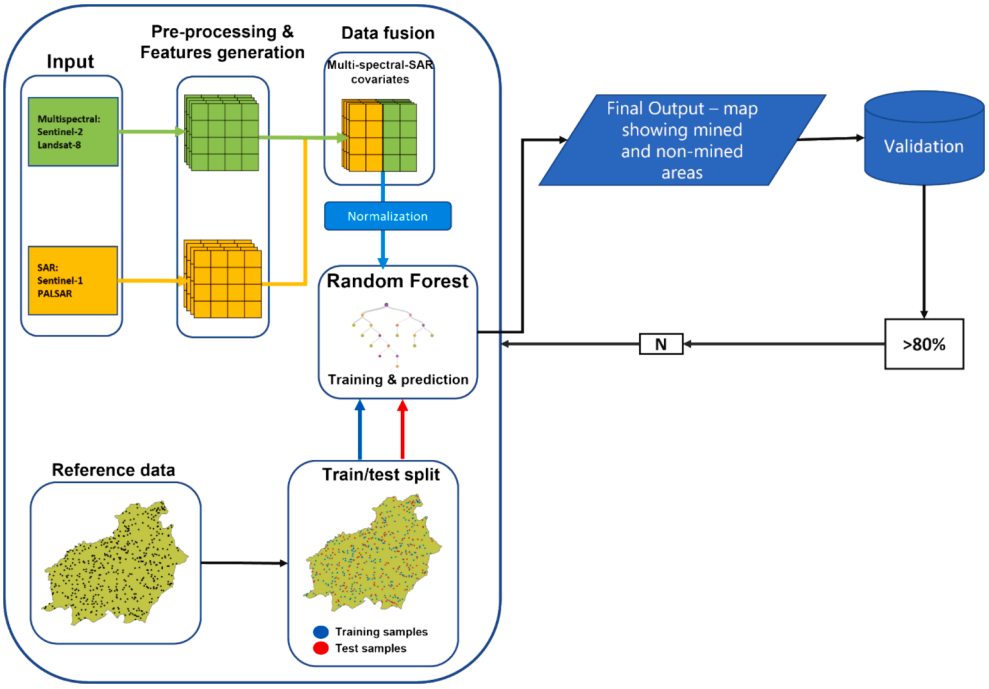


Fig. 2. Schematic overview of the methodological workflow.

bands and indices into a single feature vector (Pohl and Van Genderen 2016, Blum and Liu 2018) and used this feature vector in the Random Forest (RF) classification algorithm for subsequent training and classification steps, which included bootstrapped sampling, feature randomness, and the ensemble of decision trees (Breiman (2001)a).

We applied the RF algorithm in GEE (Breiman (2001)a), leveraging its non-parametric approach and ensemble of randomized decision trees. We chose RF over other deep-learning algorithms available in GEE for several reasons. Firstly, unlike deep learning models like Convolutional Neural Networks (CNNs), which require large, clean datasets and extensive computational resources, RF is computationally efficient and achieves robust results with fewer computational demands (Kulkarni and Sinha 2013). Secondly, RF's interpretability and adaptability are significant advantages. RF models provide insights into feature importance, helping us understand which variables most influence ASM detection, thereby informing targeted variable use. This contrasts with

Table 1
Multispectral- and radar-based indices used in ASGM mapping.

Index name	Rationale	References
Multispectral-based index		
Normalized Difference Spectral Vectors (NDSV)	Identify changes to the landscape, such as soil disturbance and the creation of waste rock piles associated with mining activities.	Angiuli and Trianni (2013)
Mine Solid Waste Index (MSWI)	Highlight solid waste generated by ASGM activities (e.g., tailings and waste rock piles).	Gao et al. (2021)
Morphological Mining Feature Index (MMFI)	Highlight mining features in remote sensing images (e.g., pits, tailings, waste rock piles).	Wu et al. (2019)
Bare Soil Index (BSI)	Identification of bare soil areas/mining open pit.	Ngom et al. (2020)
Normalized Difference Soil Index (NDSI)	Enhance the contrast between vegetation and mining open pits.	Madasa et al. (2021) and Leprieur et al. (2000)
Modified Soil Adjusted Vegetation Index (MSAVI)	Enhance the contrast between vegetation and mining open pits.	Mhangara et al. (2020)
Modified Bare Soil Index (MBSI)	Identification of bare soil areas/mining open pit accounting for atmospheric effects and shadows.	Zhang et al. (2018)
Normalized Difference Water Index (NDWI)	Mapping of soil moisture (detection of rivers and water bodies that are present in mining sites)	
Colour index (CI)	Detection of orange hues (e.g., lateritic cover)	Mathieu et al. (1998); Mathieu et al. (1998), Ngom et al. (2020)
Brightness Index (BI)	Detection of white rock-bearing minerals	Mathieu et al. (1998); Ngom et al. (2020)
Redness Index (RI)	Detection of orange hues (e.g., lateritic cover)	Mathieu et al. (1998); Ngom et al. (2020)
Radar-based index		
Normalized Difference Index (NDI)	Identify areas of soil disturbance, which can be an indicator of mining activity.	Almeida-Filho and Shimabukuro (2000)
Radar Forest Degradation Index (RFDI)	Detection of small-scale deforestation that may be associated with elusive mining activity.	Nicolau et al. (2021)
Dual-Polarization SAR Vegetation Index (DPSVI)	Detection of small-scale deforestation that may be associated with elusive mining activity.	De Luca et al. (2021)

the "black box" nature of deep learning models, which are often difficult to interpret (Biau and Scornet 2016, Shrestha and Mahmood 2019). Furthermore, RF's robustness in handling missing values and noisy data makes it well-suited to the inconsistent data quality typical in ASM studies. These attributes collectively position RF as the better option for this case study.

We extracted covariate feature vectors using 546 sampling polygons that were randomly divided into training and test datasets using the random selection function (ee.FeatureCollection.randomColumn()) in GEE to ensure unbiased selection (Section 2.1). Through iterative loops in the training and validation process, we explored various combinations of 'number of trees' and 'bag fraction' parameters (number of trees ranged from 10 to 150 with increments of 20, while Bag Fraction varied from 0.1 to 0.9 with increments of 0.1) to optimize the model, seeking the best values for these key parameters to achieve maximum accuracy (section 2.5). These iterations also involved extracting 'gini importance' values for covariate importance analysis, guiding the selection of the most influential covariates for a more concise model (Breiman (2001)b, Breiman (2004)). We then trained the final model and generated the final map using these optimized parameters and best covariates.

We validated the final ASGM map using 195 test data polygons from three distinct sources (section 2.1). The validation involved three accuracy assessment methods: (i) the confusion matrix, (ii) the F(β) score coefficient, and (iii) the Matthews Correlation Coefficient (MCC). To compare this fusion model's accuracy with the multispectral-only model, we mapped ASGM within the study area using an identical algorithm, optimized parameters, and training/test datasets, but solely with Sentinel-2 data (Section 2.6). Subsequently, we applied the same validation methods to compare the accuracies achieved by these two methodologies effectively.

2.4. Pre-processing and features generation

2.4.1. Cloud and cloud shadow masking

We implemented the Temporal Dark Outlier Mask (TDOM) (Housman et al. 2018) and the CFMASK algorithms (Zhu and Woodcock 2012) to reduce the effects of cloud and cloud shadows that plague imagery from tropical wet forests. The TDOM algorithm detects statistical outliers in relation to the sum of the infrared bands by identifying dark pixels within these bands that do not consistently exhibit darkness across past and/or future observations (Housman et al. 2018). We then created a cloud-masked Sentinel-2 image collection using the s2cloudless algorithm, which computes the cloud scores using the reflectance values across visible, near-infrared, and shortwave infrared bands (Zhu and Woodcock 2012, Gorelick et al. 2017) and filtering the images based on the defined region of interest and date range. We utilized the Quality Assessment (QA) band and custom GEE algorithms (Gorelick et al. 2017) to generate binary masks for cloud masking Landsat-8 images, which were then filtered and composited based on the dry season range (Zhu and Woodcock 2012). Temporal inconsistencies were minimized by compositing imagery from a single year (Gorelick et al. 2017).

2.4.2. Bidirectional reflectance distribution function (BRDF) correction

The Bidirectional Reflectance effect in remote sensing introduces variability in surface reflectance properties as a function of each pixel's solar illumination and sensor viewing geometries which vary over different months of each year. To detect any change in the land surface we need to normalise solar illumination and sensor viewing geometries (Nicodemus 1970, Roy et al. 2016). To address this issue, we utilized the BRDF correction method developed by Roy et al. (2016) for Landsat-8, while for Sentinel-2 BRDF correction, we employed the approaches proposed by Roy et al. (2017a) and Roy et al. (2017b).

2.4.3. Generating covariates to use as predictors of ASGM

Covariates in this study refer to the band layers and derived indices utilized as predictors of ASGM in the RF algorithm. The suite comprises individual bands and a combination of multispectral and radar-based indices (Table 1), totalling 45 covariates. These include spectral indices for Sentinel-2 and Landsat-8 (22 layers), radar-based indices for Sentinel-1 and ALOS/PALSAR (6 layers), and individual band layers: bands 2-8a for Sentinel-2, bands 2–6 for Landsat-8, bands VV and VH for Sentinel-1, and bands HH and HV for ALOS/PALSAR. We selected these indices based on their demonstrated effectiveness in discerning large-scale mining and features evocative of open-pit mining sites from previous studies. Notable among these indices are the Morphological Mining Feature Index (MMFI), which analyses landscape morphological characteristics to detect mining sites, and the Normalized Difference Index (NDI), used to identify surface soil disturbance as a mining activity indicator (Wu et al. 2019). We extracted covariate values using 546 sampling polygons (351 training and 195 test polygons; refer to section 2.1). Subsequently, we stacked these features into a single vector for deployment in the RF classification algorithm (Pohl and Van Genderen 2016).

2.5. Classification approach

We employed GEE's SMILE-RF algorithm for the classification (Breiman (2001)a) using a two-step approach. Initially, we fine-tuned adjustable hyperparameters to prevent overfitting and achieve the highest map accuracy through an exhaustive grid search, testing all possible combinations within specified ranges for each parameter (Zhou et al. 2020). This thorough analysis assessed the impact of two key parameters, Number of Trees and Bag Fraction, on model accuracy. The range for Number of Trees varied from 10 to 150 in increments of 20, while Bag Fraction ranged from 0.1 to 0.9 in increments of 0.1. Additionally, we performed covariate importance analysis using 'gini importance' to identify influential predictors, reducing unimportant covariates for a concise model (Breiman (2001)b, Breiman (2004)). Gini Importance measures the effectiveness of each predictor in splitting data into homogeneous subsets, quantifying their contribution to overall classification accuracy. Higher Gini Importance values indicate more influential variables (Strobl et al. 2007, Nembrini et al. 2018). Through iterative loops in the training and validation process, we explored various parameter combinations of these 'number of trees' and 'bag fraction,' while also extracting average 'gini importance' values.

After testing all possible hyper-parameter combinations, we identified the optimal set for the RF algorithm that yielded the highest accuracy, along with covariates showing significant gini importance values. We then trained the final model and generated the final ASGM map using the entire training dataset, the optimal hyper-parameters, and the top 25 covariates with the highest gini importance values (importance contributions exceeding 2.5). The final map was also resampled to a 10 m pixel size using the bicubic resampling method (Patil 2018).

2.6. Single sensor features versus SAR-multispectral fusion

We mapped ASGM separately using multispectral-only features and a fusion (of multispectral and radar) features to assess and compare their accuracy. For multispectral-only features, we selected 10 Sentinel-2 spectral bands and 11 multispectral-based indices derived from this Sentinel-2 (Table 1). Fusion features included all individual bands and both multispectral- and radar-based indices, as detailed in sections 2.2 and 2.4.3, resulting in two feature stacks (multispectral-only and fusion features). Using these sets, we produced two ASGM maps with the same hyper-parameterized best RF model (Section 2.5) and training/test datasets. Both maps were then validated using identical methods (Section 2.7).

2.7. Validation

We employed three different types of accuracy assessment to overcome the inherent class imbalance due to ASGM occurring at smaller extents than other LULC classes. These included: (1) the confusion matrix, (2) the f(β) score coefficient, and (3) the Matthews Correlation Coefficient (MCC). The confusion matrix for the final map was derived by GEE using the 195 test polygons (Section 2.1). The f(β) score, defined as the weighted harmonic mean between precision and recall (Couttenier et al. 2022), is particularly well-suited for evaluating binary classifiers in scenarios with class imbalance, as it effectively balances the trade-off between false positives and false negatives. A β value of 0.5 was chosen to emphasize precision, given that ASGM represents a small, minority land use class that can be easily misclassified as other categories. However, the f(β) score offers only a partial perspective of the predictive performance since it concentrates solely on positive measurements and disregards the true negative class in its computation (Hand and Christen 2018). To address this limitation, we used MCC to evaluate the binary classification quality of the model. The MCC presents a balanced assessment of positive and negative classifications and is also advantageous for evaluating models with imbalanced class sizes (Chicco and Jurman 2020). The MCC values range from −1 (entirely incorrect classification) to + 1 (perfect classification), with 0 signifying random classification.

3. Results

3.1. Regional-scale mapping result

The mapping results of the fusion model revealed a prevalence of ASGM across Central Kalimantan province, featuring substantial clusters of mining sites along major rivers and a widespread distribution of individual, cryptic, and secluded ASGM sites (Fig. 3). Our analysis distinguishes two primary types of ASGM sites. The first type consists of closely grouped mining sites with multiple individual pits adjacent to one another, forming a substantial surface footprint extending over 5 km wide along alluvial river zones (Fig. 3B, D, and F). The second type comprises isolated individual pits scattered within remote forest areas, with an average coverage area of less than 1 ha (Fig. 3C and E). These isolated ASGM activities are dispersed across the province, often concealed within dense forest canopies. Mapping these sites proves challenging due to frequent satellite imagery obstruction by vegetation and the coarse pixel resolution of freely available imagery. Additionally, the cryptic and secluded nature of these mining pits hinders detection through both satellite imagery and ground checks, impeding comprehensive data compilation for training and validation.

Mapping small surface features across expansive areas presents significant technical challenges, primarily due to spatial resolution limitations that cause imagery to lack the requisite level of detail for discriminating individual ASGM sites. While our fusion model, featuring a 10-meter resolution (Section 3.5), successfully maps ASGM sites as small as 0.01 ha, detecting sites smaller than this threshold may prove challenging. Moreover, ASGM features can spectrally resemble other land cover types, such as open bare soil and early-stage agriculture, necessitating supplementary discrimination techniques when relying solely on multispectral imagery. The integration of radar imagery into our model played a pivotal role in providing additional information to distinguish ASGM beyond spectral values, enabling differentiation from other land use and cover types (Forkuor et al. 2020, Alessi et al. 2023). Moreover, cloud cover can impede data acquisition, further complicating the task. Consequently, the fusion of multispectral and radar imagery, which is less affected by cloud cover, proves instrumental in mitigating these limitations and substantially enhancing the probability of successful ASGM detection and mapping across the study area.

3.2. Model parameterization and covariate importance

Our analysis identified the optimal values for the Number of Trees and Bag Fraction parameters, with a model featuring 100 Number of Trees and a Bag Fraction of 0.6 achieving the highest accuracy (Fig. 4a). Gini importance analysis revealed nearly equal roles for both multispectral and radar covariates in the classification process (Fig. 4b), suggesting that the incorporation of both types of remote sensing data could enhance classification accuracy. Moreover, the integration of mining-specific spectral indices, such as the Mine Soil Waste Index (MSWI) Gao et al. (2021), improved the differentiation of ASGM sites from other LULC with comparable spectral signatures, including bare soil. Additionally, the incorporation of radar-based indices, including the Radar Forest Degradation Index (RFDI), played a crucial role in ASGM detection. To streamline the model, we selected covariates with importance contributions exceeding 2.5, resulting in a final model with 25 significant covariates.

3.3. Performance metrics of fusion method

The mapping results highlight the model's capability to detect ASGM sites independently of non-mining land use, achieving an overall classification accuracy of 81 % (Table 2). A snippet map of ASGM sites

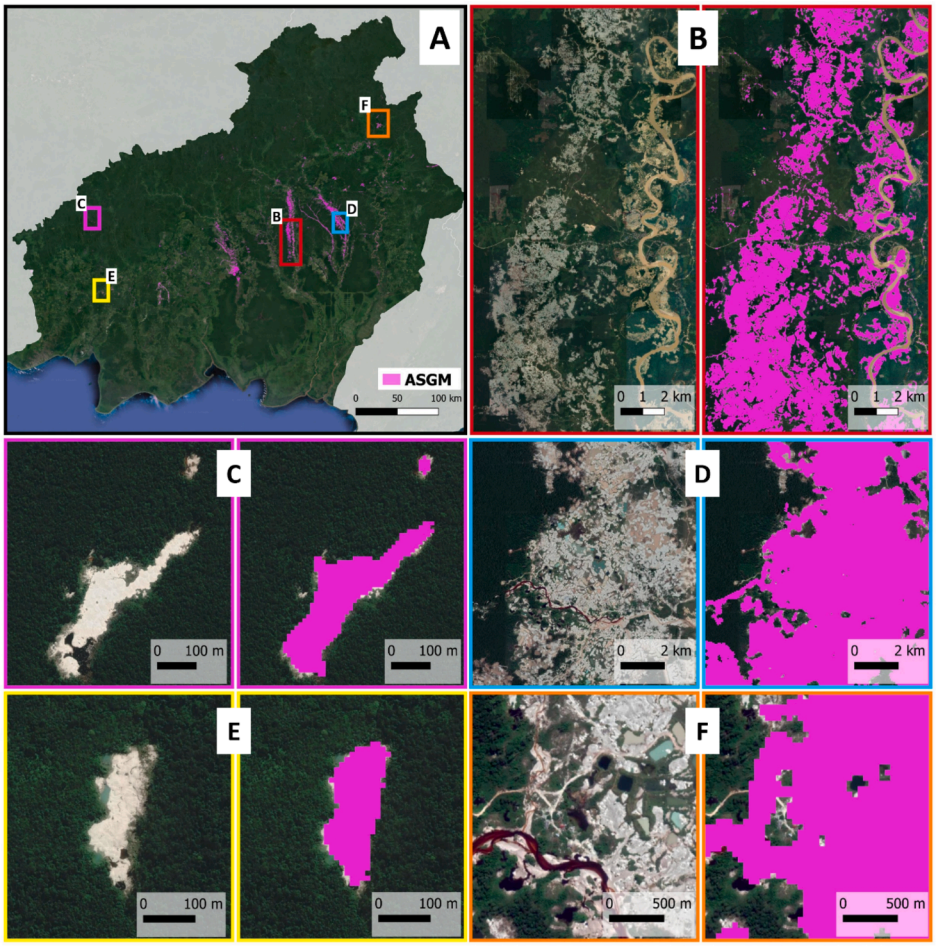


Fig. 3. ASGM map utilizing fusion features with examples of clustered and individual ASGM (color-coded). A. Distribution of mapped ASGM across Central Kalimantan Province. B, D, F. Examples of clustered ASGM along alluvial zones and the corresponding mapping results. C and E. Examples of individual isolated ASGM in forested areas.

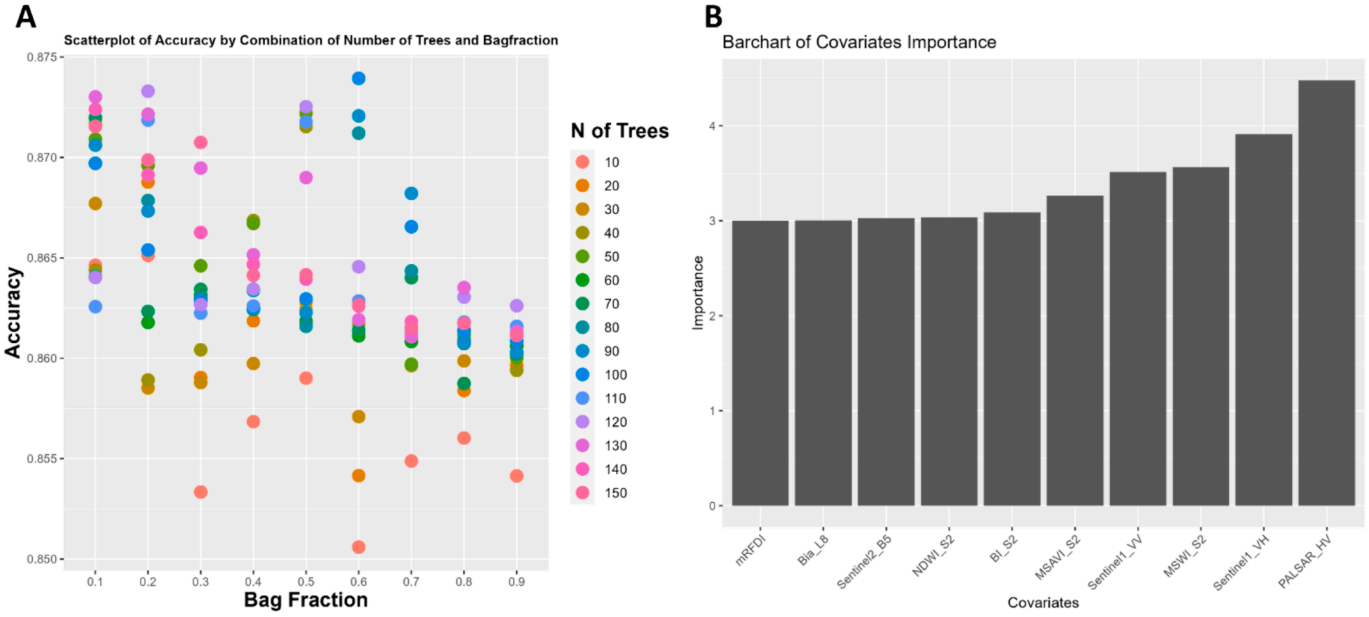


Fig. 4. Model parameterization test results demonstrate the highest accuracy achieved with 100 Number of Trees and a Bag Fraction of 0.6 (A), and the importance analysis of covariates highlights the significance of radar imageries and mining-specific indices (B).

Table 2
Performance matrix of the binary ASGM and non-ASGM mapping using data fusion.

		Validation data (n of pixels)		
		ASGM	Non-ASGM	
Classification	ASGM	387,597	66,403	Precision 0.85
	Non-ASGM	58,545	170,901	Negative Predictive Value 0.74
		Recall 0.86	Specificity 0.72	Accuracy 0.81
$f(\beta)_{score}$	$\frac{(\beta^2+1)\times precision\times recall}{(\beta^2 precision+recall)} \in [0,1] = 0.85$			
MCC_{score}	$\frac{(TP\times TN-FP\times FN)}{\sqrt{(TP+FP)\times(TP+FN)\times(TN+FP)\times(TN+FN)}} \in [-1,1] = 0.69$			

highlights the Upper Kahayan catchment – Central Kalimantan's largest ASGM hotspot – providing examples of mapping accuracy and errors (Fig. 5). The fusion model yields ASGM maps with a precision of 85 % and recall of 86 %, where precision reflects the model's ability to correctly differentiate ASGM, and recall indicates its capability to detect all mining sites. To assess binary classifiers in scenarios with class imbalance (i.e., the ASGM surface feature is minute compared to the total area of other land classes), we also utilized the common f(β) score as a performance measurement (Couttenier et al. 2022). The model's f (β) score of 0.85 indicates robust classification of the target class and low false positive errors given this class imbalance (Congalton and Green 2019, Powers 2020). The MCC score of 0.69 also indicates satisfactory accuracy in ASGM detection, considering the smaller ASGM class size than non-ASGM (Chicco and Jurman 2020, Tharwat 2020).

The fusion model works well to map congregated ASGM sites that are mostly concentrated in the alluvial zone as well as secluded individual ASGM within forest areas (Fig. 3). However, ASGM sites that directly border the river may introduce some positive errors as the area between the open pit mining sites and the river blends (Fig. 5). Another primary source of false-positive errors in our model, including bare soil, which exhibits identical surface characteristics with the typical ASGM operations. Moreover, the model introduces some false-negative errors, particularly those arising from ASGM sites concealed beneath vegetation present in the area, such small cryptic ASGM sites are difficult to detect

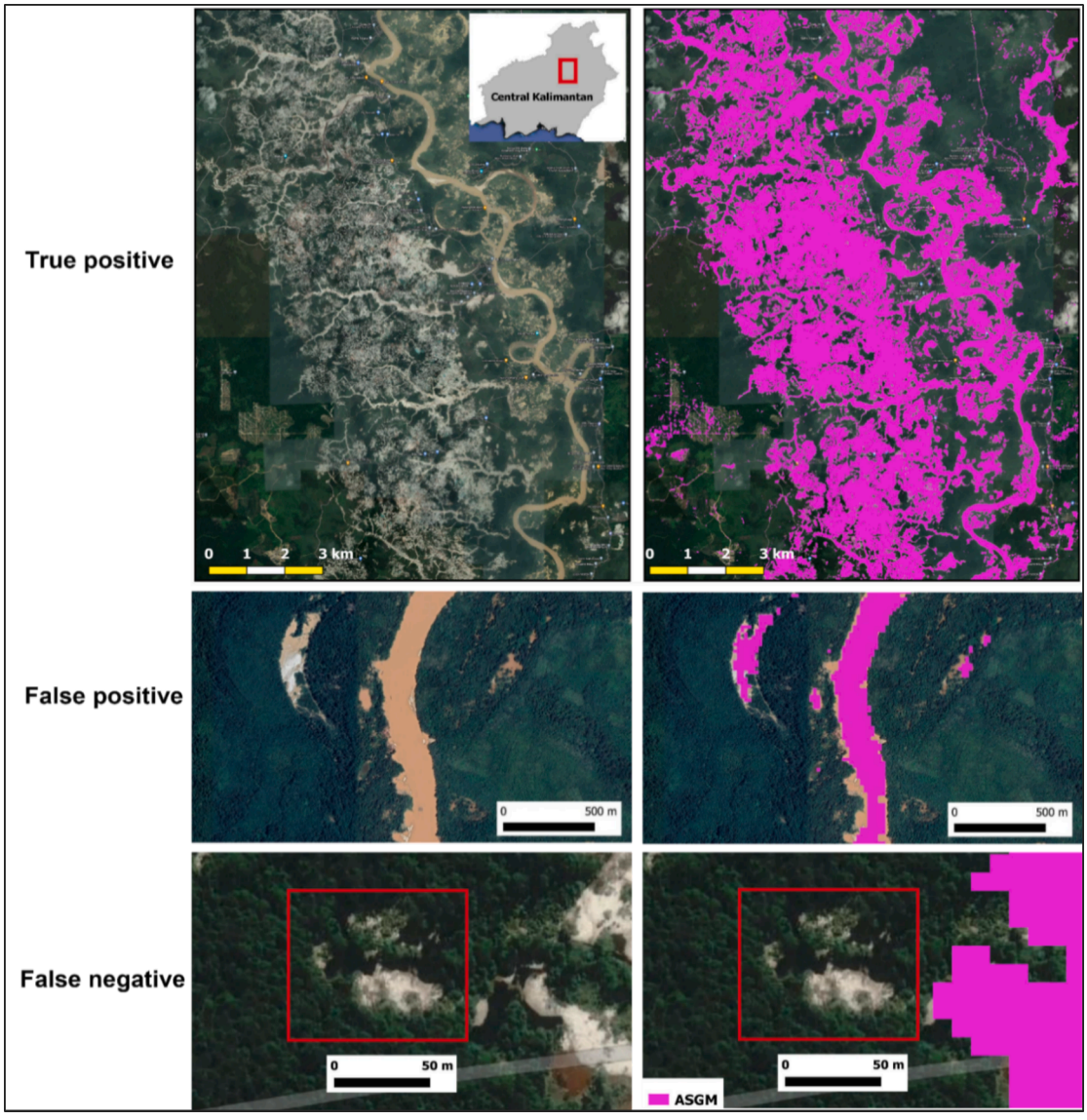


Fig. 5. Exemplar instances illustrating accurate and inaccurate mapping of ASGM sites in the Upper Kahayan catchment. The False positive error depicts a sediment-laden river incorrectly classified as ASGM, while the False negative error highlights a very small ASGM area (<1 ha) that the model failed to detect and classify.

by both multispectral and the imaging radar due to the dense vegetation concealers that even the L-band radar incapable to penetrate (Aoki et al. 2021).

3.4. Fusion model performances across ASGM size categories

Another noteworthy attribute of the model is its capacity to identify ASGM across a diverse range of surface footprint extents. Among the accurately mapped ASGM sites, the average size is approximately 0.5 ha, equivalent to nearly 50 pixels at the smallest spatial resolution of Sentinel-2 (10 m), with a standard deviation of 55.1 ha. We have stratified the mapped ASGM into four distinct size categories, encompassing micro ASM from 0.01 to 2.5 ha, small ASGM from 2.5 to 10 ha, medium ASGM from 10 to 55 ha, and extensive ASGM from above 55 ha (Fig. 6). Notably, the extensive ASGM zones (Fig. 6D) consist of smaller mining patches adjacent to each other, and the model effectively delineates these individual patches. Overall, these findings highlight the model's strong performance in detecting various forms of ASGM, even when the smallest site constitutes only a few pixels within images containing more than 350 million pixels (grain size 10 m).

3.5. Comparison to mapping employing single multispectral data type

The multispectral-only model, which relies solely on spectral-based information, effectively distinguishes large ASGM areas (>10 ha) but faces challenges in differentiating ASGM mining pits from bare-soil regions (e.g., early stages of palm plantation) and sediment-laden rivers, leading to notable false-positive errors (Table 3: Fig. 7). Moreover, it lacks the capability to detect secluded individual ASGM sites (<1 ha) due to mixed spectral values within the surrounding forest cover. In contrast, the fusion model, enriched with radar-based indices data, excels in distinguishing ASGM from other land cover with similar spectral signatures. It demonstrates the ability to detect smaller isolated ASGM sites (<1 ha) in the midst of the forest, leveraging the C-band and L-band radar's penetration of sparse vegetation canopies (Hirschmugl et al. 2020), independent of spectral information. However, the fusion model still exhibits, albeit to a lesser extent, false-positive errors, as it struggles to distinguish sediment-laden rivers from the tailing ponds of ASGM pits, particularly where the border between the ASGM sites and the river blends. Furthermore, the fusion model still contains some false-negative errors for ASGM located in very dense forests with an extent of less than 0.01 ha, primarily due to pixel size limitations and the C-band radar data's inability to penetrate very dense vegetation, even with the presence of L-band radar data, in the extremely dense primary forests of Kalimantan.

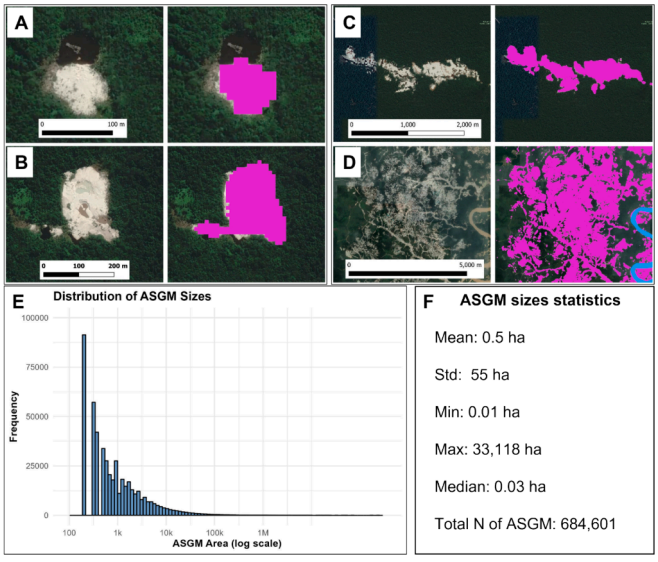


Fig. 6. ASGM size categories that were accurately detected. A. Micro ASGM (0.01 to 2.5 ha); B. Small ASGM (2.5 to 10 ha); C. Medium ASGM (10 to 55 ha); D. Extensive ASGM (>55 ha); E. Distribution of correctly mapped ASGM (in hectares) within the study area; F. Statistical outputs.

Table 3
Performance matrix of a ASGM mapping using multispectral-only.

		Validation data (n of pixels)		
		ASGM	Non-ASGM	
Classification	ASGM	355,681	86,403	Precision 0.80
	Non-ASGM	58,716	170,901	Negative Predictive Value 0.64
		Recall 0.79	Specificity 0.63	Accuracy 0.71
$f(\beta)_{score}$	$\frac{(\beta^2+1)\times precision \times recall}{(\beta^2 precision + recall)} \in [0,1] = 0.80$			
MCC_{score}	$\frac{(TP \times TN - FP \times FN)}{\sqrt{(TP+FP)\times(TP+FN)\times(TN+FP)\times(TN+FN)}} \in [-1,1] = 0.43$			

The multispectral-only model, built exclusively on Sentinel-2 satellite imagery, underperformed against the fusion model. This deficiency was underscored by precision and recall rates of 80 % and 79 % respectively (Table 3), coupled with a low f(β) score of 0.78, indicating a higher false positive rate, leading to misclassification of non-ASGM land covers as ASGM (Fig. 7). A Matthews Correlation Coefficient (MCC) score of 0.43 corroborated the single multispectral model's inferior accuracy in detecting ASGM, especially considering the smaller class size in relation to non-ASGM (Chicco and Jurman 2020, Tharwat 2020). Evidently, the fusion-based approach significantly surpasses the standalone multispectral model in detecting ASGM.

4. Discussion

Our study develops and validates a novel approach to detecting ASGM at regional scales within the challenging context of a wet-tropical environment in Kalimantan. This extensive mapping coverage capability offers a new level of insight into the collection of spatial distribution and extent of ASGM activities. Recognizing the typically scattered and remote nature of ASGM operations, our method may allow mapping and monitoring that was challenging in other areas, once tested and refined.

The Gini importance analysis demonstrated that there is a balanced contribution of multispectral and radar covariates in the classification process. Combining these layers improved the accuracy of identification compared to using only one type of satellite data and enabling the model to effectively distinguish elusive land-use such as ASGM. ASGM sites, typically surrounded by vegetation, impact vegetation extent and structure. Our findings indicate that although multispectral data, such as Sentinel-2, can identify certain vegetation disturbances, radar data, being an active sensor with longer wavelengths compared to optical sensors, provides additional benefits such as operating independently of daylight and weather conditions (Chaturvedi 2019, Alessi et al. 2023). This proves particularly beneficial in areas with substantial gaps in multispectral data due to persistent cloud coverage. Longer wavelengths also enhance radar penetration through surface features, reaching depths such as through topsoil or a forest canopy to understory structures, surpassing the capabilities of optical sensors (Moreira et al. 2013).

In our study, we combine Sentinel-1 (C-band) and ALOS/PALSAR (L-band) radar data, leveraging their respective strengths. ALOS/PALSAR, with a wavelength of 23.6 cm, operates day or night, penetrating clouds, haze, and smoke. Its sensitivity to forest structure and moisture makes it suitable for detecting disturbances in vegetation cover caused by ASGM activities. Additionally, radar signals at different frequencies interact differently with vegetation components. Sentinel-1′s shorter

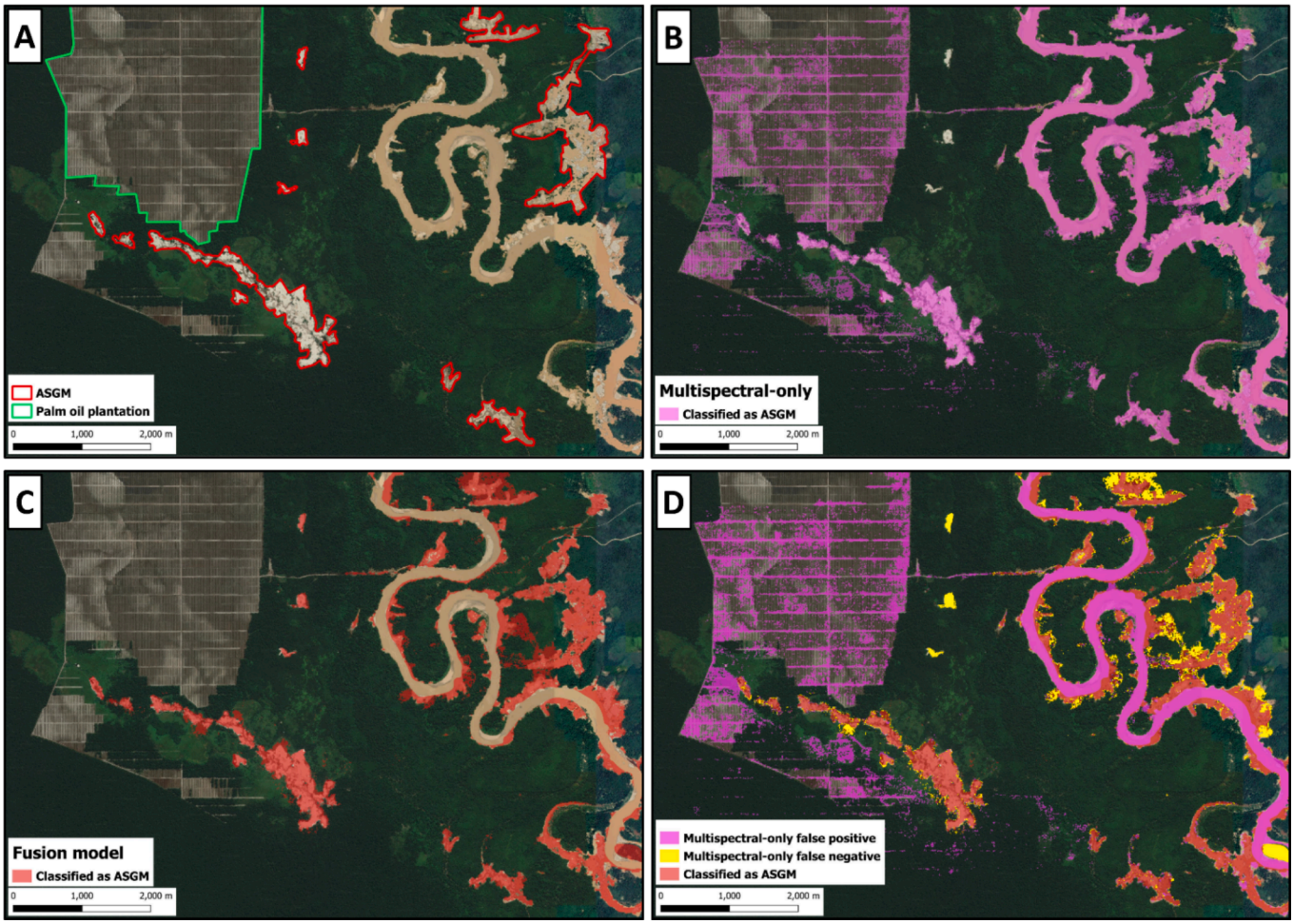


Fig. 7. Comparative mapping results in the Lower Kahayan Catchment: A. Google Earth Imagery illustrating two resembling surface features—ASGM and early-stage palm oil plantation. B. Multispectral-only ASGM-mapping output with false positives, mistakenly identifying sediment-laden rivers and palm plantations as ASGM and failing to map micro ASGM. C. Fusion model mapping output effectively distinguishing ASGM from sediment-laden rivers and accurately detecting micro ASGM. D. Overlapping classifications and associated errors reveal that the multispectral-only model produces significantly more false positives and false negatives compared to the more precise fusion model.

wavelengths interact with fine leaf and branch elements, resulting in canopy-level backscattering with limited penetration. In contrast, ALOS/PALSAR's longer wavelengths penetrate deeper, capturing backscatter from signal interactions with the ASGM surface footprint. The distinct detection signals from the C-band (Sentinel-1) and L-band (PALSAR) provide valuable information to distinguish ASGM from other forest disturbances, supporting our hypothesis that a multi-sensor approach using different SAR frequencies enhances ASGM sensing. The inclusion of ALOS/PALSAR and Sentinel-1 radar data as key predictors further enhances the model's performance, aligning with previous studies that emphasize the value of utilizing both C-band and L-band radar data to identify ASGM (Forkuor et al. 2020, Janse van Rensburg and Kemp 2022).

Our model surpasses the standalone multispectral model in overall classification accuracy, underscoring its capability for precise regional-scale ASGM mapping. Furthermore, by employing both the f(β) score and MCC in the evaluation process, we were able to confirm the model's robustness in accurately classifying target instances while minimizing false positives and addressing the challenges posed by imbalanced class sizes (Chicco and Jurman 2020, Couttenier et al. 2022). This underscores its value for both future ASGM research and on-ground applications. Moreover, our model detects ASGM activities across diverse surface footprints, from relatively small (<2.5 ha) to much larger (>55 ha) extents, outperforming the standalone model. Crucially, it exhibits proficiency in detecting and delineating individual mining patches, even when they occupy only a small fraction of the total image pixels.

Our study also highlights the significance of utilizing mining-specific spectral indices, such as the Mine Soil Waste Index (MSWI), for effective ASGM mapping. While traditional ASGM classification models often rely on established multi-band indices like NDVI and BSI (e.g., Ngom et al. (2020), Barenblitt et al. (2021), and Nyamekye et al. (2021)), they rarely integrate mining-centric indices such as MSWI and MMFI [38, 48]—indices originally crafted for large-scale mining detection. While the exploration of these indices is relatively recent, our findings suggest that the inclusion of these specific indices can bolster ASGM mapping accuracy. It should also be noted that the established index brings its unique challenges—for instance, the NDVI is influenced by various environmental factors, resulting in non-linear responses in estimates of vegetation amount due to, atmospheric conditions, varying background soil colours, closed canopies and presence of non-photosynthetic materials (Ding et al. 2017). Hence, the integration of mining-specific indices can supplement the model with a more comprehensive array of information to improve ASGM mapping accuracy.

These findings offer valuable directions for managing and monitoring ASGM activities. Integrating multi-spectral and radar data overcomes significant challenges that have traditionally hindered RS applications in wet-tropical environments (e.g., Pereira et al. (2013); Isidro et al. (2017)). The model's ability to detect and delineate

individual mining patches, irrespective of their size relative to the total image pixels, accentuates its applicability for comprehensive and precise regional-scale mapping of diverse ASGM forms, compared to some studies that have limited capabilities to detect ASGM patches smaller than 1 ha (e.g., Fonseca Gomez (2021); Nyamekye et al. (2021); Kimijima et al. (2021)). Additionally, using a cloud-based geo-processing platform addresses data storage and management issues, providing an economical solution for sustained regional-scale monitoring with sufficient accuracy. For example, regional-scale ASM mapping by Couttenier et al. (2022) requires approximately 6 GB of storage per 100 km x 100 km tile. This provides an economical solution for sustained regional-scale monitoring with sufficient accuracy – a critical requirement for developing nations hosting ASGM issues. As such, it presents a promising solution for accurate, ongoing surveillance and monitoring of ASGM activities at a large scale in wet-tropical forest contexts.

Despite the inherent limitations of each data type, their combined usage effectively counterbalances these drawbacks, resulting in improved potential for accurate ASGM mapping. However, certain challenges emerged during the model evaluation phase. Rivers with high sediment loads, specific logged areas, and shallow open soil were misclassified as ASGM due to spectral similarities, especially in the low green (0.53–0.59 μm) and high red (0.64–0.80 μm) wavelength spectrum and some soil-focused spectral indices like the Normalized Difference Soil Index (NDSI) (Ibrahim et al. 2020). These false positives are minimized by combining information from the distinct detection signals of the C-band (Sentinel-1) and L-band (PALSAR), which provide valuable information to distinguish ASGM from other forest disturbances. The combination of VV (Sentinel-1) and HH (PALSAR) polarization effectively differentiates roughness between these land cover classes (Forkuor et al. 2020, Janse van Rensburg and Kemp 2022). Additionally, the model faced difficulties in identifying ASGM sites concealed beneath vegetation (Fig. 5), despite its capacity to detect sites smaller than 1 ha. These challenges highlight the need for further refinement and optimization of the model to enhance its accuracy and effectiveness.

5. Conclusions and future work

In conclusion, our research demonstrates an initial capability for integrating passive and active satellite image data with the developed approach for accurate detection of ASGM sites in a wet-tropical environment. The findings emphasize the synergistic advantages of combining multispectral and radar data, overcoming limitations inherent to each data type, and improving classification accuracies despite challenges related to ASGM's minority status in land use classifications. The extensive regional coverage of the model provides spatially extensive, yet detailed insights into the spatial distribution and activities of ASGM. Our approach and the scales of information produced may significantly improve, ongoing surveillance, management, and regulation of ASGM activities and support the development of tailored, evidence-based strategies to mitigate ASGM's environmental and social impacts.

In transitioning from research to implementation, the approach presented in this study holds promise for mapping ASGM to larger extents and possibly to diverse environmental conditions. Enhancements can be made by incorporating additional approaches, such as texture filtering methods, to mitigate false-positive errors (Herold and Haack 2006, Shao et al. 2016). The primary sources of false-positive errors in our model, namely sediment-laden rivers and bare soil, exhibit distinct surface characteristics compared to the typical open-pit mining associated with ASGM. Therefore, surface texture differences, which can be derived from the radar data, could aid in their differentiation. Addressing false-negative errors, particularly those arising from ASGM sites concealed beneath vegetation, presents a more complex challenge due to the disturbances in surface features. However, to generate a more accurate map, a post-classification digitization process can be employed. This can involve the engagement of GIS technicians or trained local assistants with in-depth knowledge of the terrain and ASGM activities in the region, following the approach utilized by Aggrey et al. (2021) in which local individuals participated in the digitization of ASGM sites. The workflow of our model is also designed to be easily executed by local GIS technicians, promoting user-friendliness and scalability (Please refer to the data availability statement for source code). Future applications may encompass nationwide predictions or the analysis of historical or prospective trends.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

Ilyas Nursamsi: Writing – original draft, Methodology, Investigation, Data curation, Conceptualization. **Laura Jane Sonter:** Writing – review & editing, Conceptualization. **Matthew Scott Luskin:** Writing – review & editing, Conceptualization. **Stuart Phinn:** Writing – review & editing, Supervision, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

Aggrey, J.J., Ros-Tonen, M.A., Asubonteng, K.O., 2021. Using participatory spatial tools to unravel community perceptions of land-use dynamics in a mine-expanding landscape in Ghana. Environ. Manag. 68, 720–737.
Alessi, M.A., Chirico, P.G., Sunder, S., O'Pry, K.L., 2023. Detection and Monitoring of Small-Scale Diamond and Gold Mining Dredges Using Synthetic Aperture Radar on the Kadéï (Sangha) River. Central African Republic. Remote Sensing 15, 913.
Al-Hassan, S., Yaganuma, L., Odoi, B., 2019. The trajectory of lost mercury in artisanal and small scale gold extraction in Ghana. Ghana Mining Journal 19, 35–40.
Almeida-Filho, R., Shimabukuro, Y., 2000. Detecting areas disturbed by gold mining activities through JERS-1 SAR images, Roraima State, Brazilian Amazon. Int. J. Remote Sens. 21, 3357–3362.
Alvarez-Berríos, N.L., Aide, T.M., 2015. Global demand for gold is another threat for tropical forests. Environ. Res. Lett. 10, 014006.
Angiuli, E., Trianni, G., 2013. Urban mapping in Landsat images based on normalized difference spectral vector. IEEE Geosci. Remote Sens. Lett. 11, 661–665.
Aoki, Y., Furuya, M., De Zan, F., Doin, M.-P., Eineder, M., Ohki, M., Wright, T.J., 2021. L-band Synthetic Aperture Radar: Current and future applications to Earth sciences. Earth Planets Space 73, 1–4.
Barenblitt, A., Payton, A., Lagomasino, D., Fatoyinbo, L., Asare, K., Aidoo, K., Pigott, H., Som, C.K., Smeets, L., Seidu, O., 2021. The large footprint of small-scale artisanal gold mining in Ghana. Sci. Total Environ. 781, 146644.
Bebbington, D. H., R. Verdun, C. Gamboa, and A. J. Bebbington. 2018. Impacts of extractive industry and infrastructure on forests. Assessment and Scoping of Extractive Industries and Infrastructure in Relation to Deforestation: Amazonia.
Biau, G., Scornet, E., 2016. A random forest guided tour. TEST 25, 197–227.
Blum, R.S., Liu, Z., 2018. Multi-sensor image fusion and its applications. CRC Press.
Breiman, L., 2001. Random Forests. Machine Learning 45, 5–32.
Breiman, L., 2004. Consistency for a simple model of random forests. University of California at Berkeley, Technical Report, p. 670.
Chaturvedi, S.K., 2019. Study of synthetic aperture radar and automatic identification system for ship target detection. J. Ocean. Eng. Sci. 4, 173–182.
Chicco, D., Jurman, G., 2020. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC Genomics 21, 1–13.
Congalton, R.G., Green, K., 2019. Assessing the accuracy of remotely sensed data: principles and practices. CRC Press.
Couttenier, M., Di Rollo, S., Inguere, L., Mohand, M., Schmidt, L., 2022. Mapping artisanal and small-scale mines at large scale from space with deep learning. PLoS One 17, e0267963.
De Luca, G., Silva, J.M., Modica, G., 2021. A workflow based on Sentinel-1 SAR data and open-source algorithms for unsupervised burned area detection in Mediterranean ecosystems. Giscience & Remote Sensing 58, 516–541.

Ding, Q., Cheng, G., Wang, Y., Zhuang, D., 2017. Effects of natural factors on the spatial distribution of heavy metals in soils surrounding mining regions. Sci. Total Environ. 578, 577–585.
Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., 2012. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. Remote Sens. Environ. 120, 25–36.
Engstrand, R.C., Espejo, J.C., Silman, M.R., Asner, G.P., 2024. Repeated mining accounts for the majority of artisanal and small-scale gold mining activity in Southeastern Peru. Environ. Res. Lett. 19, 064036.
Fonseca Gomez, A., 2021. Detecting Artisanal Small-Scale Gold mines with LandTrendr multispectral and textural features at the Tapajós river basin. University of Twente, Brazil.
Forkuor, G., Ullmann, T., Griesbeck, M., 2020. Mapping and monitoring small-scale mining activities in Ghana using Sentinel-1 Time Series (2015–2019). Remote Sens. (basel) 12, 911.
Foundation, B., 2018. Mercury Country Situation Report Indonesia. BaliFokus Foundation South Jakarta, Indonesia.
Franks, D.M., Brereton, D., Moran, C.J., 2010. Managing the cumulative impacts of coal mining on regional communities and environments in Australia. Impact Assessment and Project Appraisal 28, 299–312.
Fritz, M., McQuilken, J., Collins, N., Weldegiorgis, F., 2018. Global Trends in Artisanal and Small-Scale Mining (ASM): A review of key numbers and issues. International Institute for Sustainable Development Winnipeg, MB, Canada.
Gao, S., Chen, Y., Li, K., Li, Y., Yu, J., Rao, R., 2021. Mapping opencast iron mine and mine solid waste based on a new spectral index from medium spatial resolution satellite data. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 14, 7788–7798.
Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sens. Environ. 202, 18–27.
Hand, D., Christen, P., 2018. A note on using the F-measure for evaluating record linkage algorithms. Stat. Comput. 28, 539–547.
Herold, N.D., Haack, B.N., 2006. Comparison and integration of radar and optical data for land use/cover mapping. Geocarto Int. 21, 9–19.
Hilson, G., McQuilken, J., 2014. Four decades of support for artisanal and small-scale mining in sub-Saharan Africa: a critical review. The Extractive Industries and Society 1, 104–118.
Hirschmugl, M., Deutscher, J., Sobe, C., Bouvet, A., Mermoz, S., Schardt, M., 2020. Use of SAR and optical time series for tropical forest disturbance mapping. Remote Sens. (basel) 12, 727.
Housman, I.W., Chastain, R.A., Finco, M.V., 2018. An evaluation of forest health insect and disease survey data and satellite-based remote sensing forest change detection methods: Case studies in the United States. Remote Sens. (basel) 10, 1184.
Ibrahim, E., Lema, L., Barnabé, P., Lacroix, P., Pirard, E., 2020. 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. 93, 102194.
Isidro, C., McIntyre, N., Lechner, A., Callow, I., 2017. Applicability of Earth Observation for Identifying Small-Scale Mining Footprints in a Wet Tropical Region. Remote Sens. (basel) 9.
Janse van Rensburg, G., Kemp, J., 2022. The Use of C-Band and X-Band SAR with Machine Learning for Detecting Small-Scale Mining. Remote Sens. (basel) 14, 977.
JAXA. 2017. 25m-Resolution PALSAR-2/PALSAR Mosaic and Global Forest/Non-Forest Map.
Kalamandeen, M., Gloor, E., Mitchard, E., Quincey, D., Ziv, G., Spracklen, D., Spracklen, B., Adami, M., Aragão, L.E., Galbraith, D., 2018. Pervasive rise of small-scale deforestation in Amazonia. Sci. Rep. 8, 1–10.
Kimijima, S., Sakakibara, M., Nagai, M., 2021. Detection of artisanal and small-scale gold mining activities and their transformation using earth observation, nighttime light, and precipitation data. Int. J. Environ. Res. Public Health 18, 10954.
Kulkarni, V.Y., Sinha, P.K., 2013. Random forest classifiers: a survey and future research directions. Int. J. Adv. Comput 36, 1144–1153.
Lee, J.-S., Wen, J.-H., Ainsworth, T.L., Chen, K.-S., Chen, A.J., 2008. Improved sigma filter for speckle filtering of SAR imagery. IEEE Trans. Geosci. Remote Sens. 47, 202–213.
Leprieur, C., Kerr, Y., Mastorchio, S., Meunier, J., 2000. Monitoring vegetation cover across semi-arid regions: comparison of remote observations from various scales. Int. J. Remote Sens. 21, 281–300.
Macháček, J., 2019. Typology of environmental impacts of artisanal and small-scale mining in African Great Lakes Region. Sustainability 11, 3027.
Madasa, A., Orimoloye, I.R., Ololade, O.O., 2021. Application of geospatial indices for mapping land cover/use change detection in a mining area. J. Afr. Earth Sc. 175, 104108.
Mathieu, R., Pouget, M., Cervelle, B., Escadafal, R., 1998. Relationships between satellite-based radiometric indices simulated using laboratory reflectance data and typic soil color of an arid environment. Remote Sens. Environ. 66, 17–28.
Maus, V., Giljum, S., da Silva, D.M., Gutschlhofer, J., da Rosa, R.P., Luckeneder, S., Gass, S.L., Lieber, M., McCallum, I., 2022. An update on global mining land use. Sci. Data 9, 1–11.
Mhangara, P., Tsoeleng, L., Mapurisa, W., 2020. Monitoring the development of artisanal mines in South Africa. J. South Afr. Inst. Min. Metall. 120, 299–306.
Moomen, A.-W., Lacroix, P., Benvenuti, A., Planque, M., Piller, T., Davis, K., Miranda, M., Ibrahim, E., Giuliani, G., 2022. Assessing the Applications of Earth Observation Data for Monitoring Artisanal and Small-Scale Gold Mining (ASGM) in Developing Countries. Remote Sens. (basel) 14, 2971.
Moreira, A., Prats-Iraola, P., Younis, M., Krieger, G., Hajnsek, I., Papathanassiou, K.P., 2013. A tutorial on synthetic aperture radar. IEEE Geosci. Remote Sens. Mag. 1, 6–43.
Mullissa, A., Vollrath, A., Odongo-Braun, C., Slagter, B., Balling, J., Gou, Y., Gorelick, N., Reiche, J., 2021. Sentinel-1 sar backscatter analysis ready data preparation in google earth engine. Remote Sens. (basel) 13, 1954.
Mutanga, O., and L. Kumar. 2019. Google earth engine applications. Page 591. MDPI.
Nembrini, S., König, I.R., Wright, M.N., 2018. The revival of the Gini importance? Bioinformatics 34, 3711–3718.
Ngom, N.M., Mbaye, M., Baratoux, D., Baratoux, L., Catry, T., Dessay, N., Faye, G., Sow, E., Delaître, E., 2020. Mapping artisanal and small-scale gold mining in Senegal using Sentinel 2 data. GeoHealth, 4:e2020GH000310.
Nicodemus, F.E., 1970. Reflectance nomenclature and directional reflectance and emissivity. Appl. Opt. 9, 1474–1475.
Nicolau, A.P., Flores-Anderson, A., Griffin, R., Herndon, K., Meyer, F.J., 2021. Assessing SAR C-band data to effectively distinguish modified land uses in a heavily disturbed Amazon forest. Int. J. Appl. Earth Obs. Geoinf. 94, 102214.
Nyamekye, C., B. Ghansah, E. Agyapong, and S. Kwofie. 2021. Mapping changes in artisanal and small-scale mining (ASM) landscape using machine and deep learning algorithms.-a proxy evaluation of the 2017 ban on ASM in Ghana. Environmental Challenges 3:100053.
Ofosu, G., Dittmann, A., Sarpong, D., Botchie, D., 2020. Socio-economic and environmental implications of Artisanal and Small-scale Mining (ASM) on agriculture and livelihoods. Environ Sci Policy 106, 210–220.
Patil, M., 2018. Interpolation techniques in image resampling. International Journal of Engineering & Technology 7, 567–570.
Pereira, L. d. O., C. d. C. Freitas, S. J. S. Sanı́ Anna, D. Lu, and E. F. Moran. 2013. Optical and radar data integration for land use and land cover mapping in the Brazilian Amazon. GIScience & remote sensing 50:301-321.
Pohl, C., Van Genderen, J.L., 1998. Review article multisensor image fusion in remote sensing: concepts, methods and applications. Int. J. Remote Sens. 19, 823–854.
Pohl, C., Van Genderen, J., 2016. Remote sensing image fusion: A practical guide. Crc Press.
Powers, D. M. 2020. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. arXiv preprint arXiv:2010.16061.
Roy, D.P., Zhang, H., Ju, J., Gomez-Dans, J.L., Lewis, P.E., Schaaf, C., Sun, Q., Li, J., Huang, H., Kovalskyy, V., 2016. A general method to normalize Landsat reflectance data to nadir BRDF adjusted reflectance. Remote Sens. Environ. 176, 255–271.
Roy, D.P., Li, J., Zhang, H.K., Yan, L., Huang, H., Li, Z., 2017a. Examination of Sentinel-2A multi-spectral instrument (MSI) reflectance anisotropy and the suitability of a general method to normalize MSI reflectance to nadir BRDF adjusted reflectance. Remote Sens. Environ. 199, 25–38.
Roy, D.P., Li, Z., Zhang, H.K., 2017b. Adjustment of Sentinel-2 multi-spectral instrument (MSI) Red-Edge band reflectance to Nadir BRDF adjusted reflectance (NBAR) and quantification of red-edge band BRDF effects. Remote Sens. (basel) 9, 1325.
Shafiee, S., Topal, E., 2010. An overview of global gold market and gold price forecasting. Resour. Policy 35, 178–189.
Shao, Z., Zhang, Y., Zhang, L., Song, Y., Peng, M., 2016. Combining spectral and texture features using random forest algorithm: Extracting impervious surface area in Wuhan. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 41, 351–358.
Shimada, M., Isoguchi, O., Tadono, T., Isono, K., 2009. PALSAR radiometric and geometric calibration. IEEE Trans. Geosci. Remote Sens. 47, 3915–3932.
Shimada, M., Ohtaki, T., 2010. Generating large-scale high-quality SAR mosaic datasets: Application to PALSAR data for global monitoring. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 3, 637–656.
Shrestha, A., Mahmood, A., 2019. Review of deep learning algorithms and architectures. IEEE Access 7, 53040–53065.
Strobl, C., Boulesteix, A.-L., Zeileis, A., Hothorn, T., 2007. Bias in random forest variable importance measures: Illustrations, sources and a solution. BMC Bioinf. 8, 1–21.
Swenson, J.J., Carter, C.E., Domec, J.-C., Delgado, C.I., 2011. Gold mining in the Peruvian Amazon: global prices, deforestation, and mercury imports. PLoS One 6, e18875.
Tang, L., Werner, T.T., 2023. Global mining footprint mapped from high-resolution satellite imagery. Communications Earth & Environment 4, 134.
Tharwat, A., 2020. Classification assessment methods. Applied Computing and Informatics 17, 168–192.
USGS. 2016. Using the USGS Landsat8 product. US Department of the Interior-US Geological Survey–NASA.
Van Tricht, K., Gobin, A., Gilliams, S., Piccard, I., 2018. Synergistic use of radar Sentinel-1 and optical Sentinel-2 imagery for crop mapping: A case study for Belgium. Remote Sens. (basel) 10, 1642.
Villegas, C., A. B. Turay, and D. Sarmu. 2013. cAn ArtisAnAl mining & conserVAtion co-eXist? A Case Study of Artisanal Mining in and adjacent to Sierra Leone's Gola Rainforest National Park and Recommendations on the Way Forward. Produced for the ASM-PACE Programme. Available at: www. asm-pace. org.
World-Bank-Group, 2019. Forest-Smart Mining: Identifying Good and Bad Practices and Policy Responses for Artisanal and Small-Scale Mining in Forest Landscapes. World Bank.
Wu, B., Zhao, Y., Fang, C., 2019. Detection of Spatiotemporal Changes of Surface Mining Area in Changting County Southeast China. In: IGARSS 2019–2019 IEEE International Geoscience and Remote Sensing Symposium. IEEE, pp. 1606–1609.
Zhang, Y., Ling, F., Wang, X., Foody, G.M., Boyd, D.S., Li, X., Du, Y., Atkinson, P.M., 2021. Tracking small-scale tropical forest disturbances: Fusing the Landsat and Sentinel-2 data record. Remote Sens. Environ. 261, 112470.

Zhang, S., Yang, K., Li, M., Ma, Y., Sun, M., 2018. Combinational biophysical composition index (CBCI) for effective mapping biophysical composition in urban areas. IEEE Access 6, 41224–41237.

Zhou, B., Okin, G.S., Zhang, J., 2020. Leveraging Google Earth Engine (GEE) and machine learning algorithms to incorporate in situ measurement from different times for rangelands monitoring. Remote Sens. Environ. 236, 111521.

Zhu, Z., Woodcock, C.E., 2012. Object-based cloud and cloud shadow detection in Landsat imagery. Remote Sens. Environ. 118, 83–94.